STOCHASTIC ORDER IN DICHOTOMOUS ITEM RESPONSE MODELS FOR FIXED, ADAPTIVE, AND MULTIDIMENSIONAL TESTS

WIM J. VAN DER LINDEN
UNIVERSITY OF TWENTE

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. 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 a series of theorems and corollaries which apply to dichotomous IRT models. One part of the results holds for unidimensional models with fixed item parameters. The other part holds for models with random item parameters as used, for example, in adaptive testing or for tests with multidimensional abilities.

Key words: item response theory (IRT), stochastic order, computerize adaptive testing (CAT), multidimensional tests, classical test theory.

Suppose an educational or psychological test consists of a set of \( n \) dichotomously scored items indexed by \( i = 1, \ldots, n \). Responses to item \( i \) are denoted by a random variable \( Y_i \) with realization \( y_i \), where \( y_i = 1 \) denotes a correct response and \( y_i = 0 \) an incorrect response. In addition, it is assumed that the examinees responding to the test items have an ability that can be represented by a (latent) unidimensional variable \( \theta \). 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 response function to model the probabilities with which an examinee responds to an item. Let \( p_i(\theta) \) be defined as the two-parameter logistic (2-PL) function

\[
p_i(\theta) = \frac{1}{1 + \exp(-a_i(\theta - b_i))}, \quad -\infty < \theta < \infty, -\infty < b_i < \infty, \text{ and } a_i > 0,
\]

where \( b_i \) and \( a_i \) are usually interpreted as the difficulty and discriminating power of item \( i \), respectively. Then, the probability that an examinee with ability level \( \theta \) produces a correct response to the item \( i \) is modeled as

\[
\text{Prob} \{ Y_i = 1 | \theta \} = p_i(\theta) = \frac{1}{1 + \exp(-a_i(\theta - b_i))}.
\]

The equation in (2) is an example of the response function representation of an IRT model. Alternatives to the two-parameter logistic function are the more sparsely parameterized Rasch or one-parameter logistic (1-PL) (Fischer & Molenaar, 1995) and the

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 the referees for their comments on the previous version of the manuscript, as well as to the Greek fisherman who picked up the only copy of the set of handwritten notes for this paper from the harbor of Karpathos.

Requests for reprints should be sent to W. J. van der Linden, Department of Educational Measurement and Data Analysis, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands. Email: vanderlinden@edte.utwente.nl.

0033-3123/98/0900-9647$00.75/0
© 1998 The Psychometric Society
Birnbaum or three-parameter logistic (3-PL) function (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. Its 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 probability mass functions (pmfs) for the distribution of \(Y_i\). This family can be denoted as \(\{f_i(y_i|\theta)\}; -\infty < \theta < \infty\), where

\[
f_i(y_i|\theta) = p_i(\theta)^y (1 - p_i(\theta))^{1-y},
\]

and \(p_i(\theta)\) is defined by (1). This representation focuses on the conditional probability distribution of \(Y_i\) given \(\theta\). It is standard in texts on the statistical treatment of the estimation of the values of the item and/or ability parameters. Under the assumption of conditional independence, 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 cumulative distribution functions (cdfs) \(\{F_i(y_i|\theta)\}; -\infty < \theta < \infty\), where

\[
F_i(y_i|\theta) = \sum_{u=0}^{y_i} f_i(u|\theta),
\]

and \(f_i(u|\theta)\) is given by (3). Note that this representation in fact concentrates on the incorrect response whereas the one in (1) concentrates on the correct response. The cdf representation is the one addressed in the current paper. In particular, the interest is in the property of stochastic order in this representation.

Note that the three representations constitute a hierarchy because the cdf representation uses the pmf representation, which, in turn, is based on the response function representation. Further, it is important to note a few 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. Also, the logistic function in the first representation is monotonically increasing in \(\theta\) whereas the family of cdfs in the third representation is monotonically nonincreasing in \(\theta\) 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); Hemker, Sijtsma, Molenaar, and Junker (1996); Hemker, Sijtsma, Molenaar, and Junker (1997); Holland (1981, 1990); Holland and Rosenbaum (1986); Huyhn (1994); Junker (1991, 1993); Mokken (1971, 1997); Mokken and Lewis (1982); Molenaar (1997); Ramsay (1991, 1997); Rosenbaum (1984, 1985); Stout (1987, 1990); Sijtsma (1988); Sijtsma and Junker (1994); and Sijtsma and Meijer (1992). It is believed to be useful to examine 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 examination may help to reveal certain structures in the data which 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. An example of such diagnostics is pro-
vided in the early work of Mokken (1971) as well as the follow up by Holland and Rosenbaum (1986) and Rosenbaum (1984, 1985). These authors derived an important result for conditional covariances between item response variables which is reviewed in Theorem 1 below.

The first part of this paper consists of a systematic treatment of the notion of stochastic order in families of (dichotomous) (multivariate) random variables such as those defined in (4). Several properties of these families will be introduced as a series of lemmas with proofs. The first lemmas are not new but review known results in a concise form; exact proofs of these results can be found in the references given. The main results in this paper are formulated in the next lemmas as well as in a series of theorems and corollaries that specialize to IRT. One part of these results holds for the conventional case of a unidimensional test with a fixed design. The other part specifies the conditions under which the results hold if the ability structure underlying the test is multidimensional or if the test items are assigned randomly to the examinees, for example, as in adaptive testing.

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. However, a more comprehensive treatment of the notion of stochastic order lacks in these texts, which typically only give the definitions of a monotone likelihood ratio and stochastic order (see below) and then leave Lemmas 1 and 2 (see below) as an exercise to the reader. Because of the relevance of the concept of stochastic order for dichotomous IRT, one of the goals of this section of the paper is to try to fill this void. In particular, it is examined 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.

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 support of their pdfs. For simplicity, the notation \( f(\cdot) \) and \( F(\cdot) \) will be used for all pdfs and cdfs throughout the paper. Likewise, \( q(\cdot) \) will be used as a common notation for algebraic functions. With a few exceptions, \( Y \) is a generic symbol for an observable random variable and \( \theta \) a generic symbol for the (real-valued) parameter of interest. In the framework of IRT, \( Y \) can be thought of as an observed score on a test, a subtest, or an item whereas \( \theta \) is the latent variable measured by the test; however, the theory has larger generality. In a Bayesian fashion, all parameters are considered as realizations of random variables upon which the distribution of the observed random variable is conditioned. Thus, \( f(y|\theta) \) is the pdf associated with the distribution of a random variable \( Y \) with parameter \( \theta \). Likewise, \( \{Y|\theta\} \) represents a family of random variables indexed by parameter \( \theta \).

**Definition 1 (Monotone likelihood ratio).** A family of (conditional) density functions \( \{f(y|\theta)\} \) has a monotone likelihood ratio (MLR) in \( y \) with respect to \( \theta \) if for any \( \theta_1 > \theta_0 \)

\[
\frac{f(y|\theta_1)}{f(y|\theta_0)}
\]

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 $\{Y|\theta\}$ is stochastically ordered (SO) in $\theta$ if for all $y$ the cumulative distribution functions $\{F(y|\theta)\}$ do not increase in $\theta$ (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 $\{Y|\theta\}$ is SO if $Y$ and $\Theta$ 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, as is well known, MLR is stronger than SO (see Lemma 4 below). MLR is a useful property in statistical inference whereas the assumption of SO in $\theta$ is often made in statistical modeling because it is weaker and has the nice interpretation of a family of cdfs being similarly ordered across all possible parameter values for each possible value of its argument.

As explained in the introduction, a dichotomous IRT model can be represented by a family of cdfs $\{F(y_i|\theta): -\infty < \theta < \infty\}$ fully determined by probabilities $\{f(y_i|\theta): -\infty < \theta < \infty\}$ that are modeled as a increasing function of $\theta$. Since $\{F(y_i|\theta)\}$ is decreasing in $\theta$, this family is SO in $\theta$. Also, because the response variables $Y_i$ are dichotomous, it holds that $\{f(y_i|\theta)\}$ has MLR in $Y_i$ w.r.t. $\theta$ (Lemma 6). Finally, the usual assumption of local independence between response variables for different items guarantees the conditional independence required in some of the lemmas and theorems below.

**Expected Values**

Note that if the above two definitions hold, they also hold for $\theta$ and/or $Y$ replaced by nondecreasing functions $\varphi_1(\theta)$ and $\varphi_2(Y)$. The well-known Lemmas 1 and 2 below are based on a multivariate version of this property. The lemmas apply to IRT, for example if $Y_i, i = 1, \ldots, n$, are chosen to be the responses to $n$ test items and $\varphi(Y_1, \ldots, Y_n)$ the (possibly positively weighted) number-right score.

**Lemma 1.** Let $\{\{Y_i|\theta\}: i = 1, \ldots, n\}$ be independently distributed with densities $f(Y_i|\theta)$ and let $\varphi(y_1, \ldots, y_n)$ be a function nondecreasing in each $y_i$. If $\{f(y_i|\theta)\}$ has MLR in $y_i$ w.r.t. $\theta$ for all $i$, then $E[\varphi(Y_1, \ldots, Y_n)|\theta]$ is a nondecreasing function of $\theta$ (Lehmann, 1986, p. 85, Lemma 2(i)).

**Lemma 2.** Under the same conditions as in Lemma 1, if $\{\{Y_i|\theta\}: i = 1, \ldots, n\}$ is SO in $\theta$, then $E[\varphi(Y_1, \ldots, Y_n)|\theta]$ is a nondecreasing function of $\theta$ (Lehmann, 1986, p. 116, exercise 5).

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

**Inverse Conditioning**

The question can be raised when the properties of MLR and SO for a family of conditional variables, $\{Y|\theta\}$, imply MLR and/or SO for the inverse family, $\{\Theta|Y\}$. 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. \( \{ f(y|\theta) \} \) has MLR in \( y \) w.r.t. \( \theta \) if and only if \( \{ f(\theta|y) \} \) has MLR in \( \theta \) w.r.t. \( y \) (Chuang, Chen & Novick, 1981, Theorem 1).

Though this property of symmetry seems to support the intuition of MLR as “positive dependency” 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.

Lemma 4. If \( \{ f(y|\theta) \} \) has MLR in \( y \) w.r.t. \( \theta \), then \( \{ Y|\theta \} \) is SO in \( \theta \) and \( \{ \Theta|y \} \) is SO in \( y \) (Lehmann, 1986, p. 85, Lemma 2(ii)).

**Dichotomous Variables**

Since the response variables in the IRT models considered in this paper are dichotomous, it is important to review the following properties of SO in families of dichotomous variables:

Lemma 5. If \( Y \) is dichotomous, then \( \{ f(y|\theta) \} \) has MLR in \( y \) w.r.t. \( \theta \) if and only if \( \{ Y|\theta \} \) is SO in \( \theta \) (Grayson, 1988, Theorem 2 for \( n = 1 \); Holland & Rosenbaum (1986, sect. 5).

Lemma 6. If \( Y \) is dichotomous and \( \{ Y|\theta \} \) is SO in \( \theta \), then \( \{ \Theta|y \} \) is SO in \( y \).

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

The last two lemmas follow immediately from Lemma 3 and 5. Conditions under which observed scores in IRT models for polytomous items have MLR are presented in Hemker, Sijtsma, Molenaar and Junker (1996).

**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 the families \( \{ Z|y \} \) 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

\[
f(z|x) = \int f(z, y|x) \, dy
\]

\[
= \int f(z|y, x) f(y|x) \, dy
\]

\[
= \int f(z|y) f(y|x) \, dy.
\]

Thus,
Table 1

Numerical example showing that SO is not transitive

<table>
<thead>
<tr>
<th></th>
<th>$\Theta = 1$</th>
<th>$\Theta = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y = 1$</td>
<td>.20</td>
<td>.10</td>
</tr>
<tr>
<td>$Y = 2$</td>
<td>.05</td>
<td>.20</td>
</tr>
<tr>
<td>$Z = 1$</td>
<td>.10</td>
<td>.10</td>
</tr>
<tr>
<td>$Z = 2$</td>
<td>.15</td>
<td>.10</td>
</tr>
</tbody>
</table>

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

$F(z|y)$ is nonincreasing in $y$ and $\{Y|x\}$ is SO in $x$. It follows from Lemma 2 that $F(z|x)$ does not increase in $x$, and thus that $\{Z|x\}$ is SO in $x$. □

The lemma has the following analogy in linear correlations: If $\rho(X, Y) > 0$ and $\rho(Y, Z) > 0$, then $\rho(X, Z) > 0$ if the partial correlation $\rho(Z, X|Y) = 0$. Here it is used to proof the following result which applies to IRT when $Y_1$ and $Y_2$ are chosen to be observed scores on different (sets of) test items. Practical examples will be given in Corollary 1 below.

**Lemma 9.** If $\{Y_1|\theta\}$ and $\{Y_2|\theta\}$ are SO in $\theta$, $Y_1$ is dichotomous, and $Y_1$ and $Y_2$ are independent given $\theta$, then $\{Y_2|Y_1\}$ is SO in $y_1$.

**Proof.** Lemmas 6 and 8. □

The example in Table 1 shows that SO is not transitive. As $P(Y_1 = 1|\Theta = 1) = 0.30/0.50 > P(Y_1 = 1|\Theta = 2) = 0.20/0.50$ and $P(Y_2 = 1|Y_1 = 1) = 0.30/0.50 > P(Y_2 = 1|Y_1 = 2) = 0.25/0.50$, it follows that $F_{Y_1|\theta}(y_1|1) \geq F_{Y_1|\theta}(y_1|2)$ and $F_{Y_2|Y_1}(y_2|1) \geq F_{Y_2|Y_1}(y_2|2)$ for all $y_1$ and $y_2$, respectively. However, $F_{Y_2|\theta}(1|1) < F_{Y_2|\theta}(1|2)$ because $P(Y_2 = 1|\Theta = 1) = 0.25/0.50 < P(Y_2 = 1|\Theta = 2) = 0.30/0.50$.

**Multivariate Conditioning Variables**

Families of conditional distributions with more than one conditioning variable are introduced and the question is raised under what conditions the property of SO is maintained if the transition to a single conditioning variable is made. As will become clear later in this paper, the question is relevant for the treatment of stochastic order in IRT models for multivariate abilities or models with a stochastic item difficulty parameter as in adaptive testing. For simplicity, only the case of two conditioning variables is discussed but generalization to larger numbers of conditioning variables is readily obtained. The two parameters or conditioning variables are denoted as $(\xi_1, \xi_2)$. This notation encompasses the case
where one parameter is the ability variable and the other the difficulty of the items, $(\xi_1, \xi_2) = (\theta, b)$, as well as the case in which the two parameters denote two ability variables, $(\xi_1, \xi_2) = (\theta_1, \theta_2)$.

The family $\{Y|\xi_1, \xi_2\}$ is defined to be SO in $\xi_1$ and $\xi_2$ if $F(y|\xi_1, \xi_2)$ is nonincreasing in $\xi_1$ for all $\xi_2$ and in $\xi_2$ for all $\xi_1$. The following lemma identifies a condition under which SO is transferred to $\{Y|\xi_1\}$:

**Lemma 10.** Let $Y$ be a random variable with density function $f(y)$. Further, $\{Y|\xi_1, \xi_2\}$ is assumed to be SO in $\xi_1$ and $\xi_2$. Then $\{Y|\xi_1\}$ is SO in $\xi_1$ if $\{\Xi_2|\xi_1\}$ is SO in $\xi_1$.

**Proof.** The lemma is proved as follows:

$$f(y|\xi_1) = \int f(y, \xi_2|\xi_1) \, d\xi_2$$

$$= \int f(y|\xi_1, \xi_2) f(\xi_2|\xi_1) \, d\xi_2.$$  

Thus,

$$F(y|\xi_1) = \int F(y|\xi_1, \xi_2) f(\xi_2|\xi_1) \, d\xi_2.$$  

Since $F(y|\xi_1, \xi_2)$ is nonincreasing in $\xi_2$ and $\{\Xi_2|\xi_1\}$ is SO in $\xi_1$, it follows from Lemma 2 that $\{Y|\xi_1\}$ is SO in $\xi_1$. \[\square\]

Note that the condition of $\{\Xi_2|\xi_1\}$ being SO in $\xi_1$ implies that Lemma 10 holds if $\Xi_1$ and $\Xi_2$ are independent. A direct proof of this implication can be obtained through the use of the weighted mean-value theorem for integrals (Apostol, 1967, sec. 3.19).

The lemma thus shows that one can proceed from a multivariate to a marginal condition, if the multivariate condition demonstrates SO itself. The lemma was presented in van der Linden (1998) and van der Linden and Vos (1996) as a tool for analyzing the monotonicity of decision rules in testing for selection, mastery, and placement. Note that for $\theta_1$ and $\theta_2$ being independent, Lemma 8 is a special case of Lemma 10.

**Multivariate Distributions**

A multivariate family of random variables $\{Y_1, \ldots, Y_n|\theta\}$ is defined to be SO in $\theta$ if, for all $(y_1, \ldots, y_2)$, $\{F(y_1, \ldots, y_n|\theta)\}$ does not increase in $\theta$.

The example in Table 2 shows that SO in a series of families of univariate distribution functions does not imply multivariate SO. For example, $F_{Y_1|\theta}(1|1) = (0.25 + 0.15)/0.50 > F_{Y_1|\theta}(1|2) = (0.10 + 0.00)/0.50$. The same relation holds for $F_{Y_2|\theta}(1|\theta)$. However, $F_{Y_1, Y_2|\theta}(2, 1|1) = 0.05/0.50 < F_{Y_1, Y_2|\theta}(2, 1|2) = 0.10/0.50$.

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

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

**Proof.** The lemma follows immediately from the fact that the univariate cdfs are nonnegative and do not increase in $\theta$. \[\square\]
Table 2

**Numerical example showing that univariate SO does not imply multivariate SO**

<table>
<thead>
<tr>
<th></th>
<th>$\Theta = 1$</th>
<th>$\Theta = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y_1 = 1$</td>
<td>$Y_1 = 2$</td>
<td>$Y_1 = 1$</td>
</tr>
<tr>
<td>$Y_2 = 1$</td>
<td>.25</td>
<td>.05</td>
</tr>
<tr>
<td>$Y_2 = 2$</td>
<td>.15</td>
<td>.05</td>
</tr>
</tbody>
</table>

The reverse implication, however, does hold generally:

**Lemma 12.** If $\{Y_1, \ldots, Y_n|\theta\}$ is SO in $\theta$, then any subset of $Y$ variables is SO in $\theta$.

**Proof.** Clearly, for any $\theta_1 > \theta_0$,

$$F(y_1, \ldots, y_n|\theta_1) \leq F(y_1, \ldots, y_n|\theta_0)$$

for all values of $(y_1, \ldots, y_n)$. Let $(y_{i_1}, \ldots, y_{i_q})$ be any permutation of the vector of $y$ variables, $(y_{i_1}, \ldots, y_{i_p})$ its subvector with the first $p$ variables and $(y_{i_{p+1}}, \ldots, y_{i_q})$ its complement. Then,

$$F(y_{i_1}, \ldots, y_{i_p}|\theta_1) = \lim_{y_{i_{p+1}}, \ldots, y_{i_q} \to y_{i_{p+1}}, \ldots, y_{i_q}|\theta_1} F(y_{i_1}, \ldots, y_{i_q}|\theta_1)$$

$$\leq \lim_{y_{i_{p+1}}, \ldots, y_{i_q} \to y_{i_{p+1}}, \ldots, y_{i_q}|\theta_0} F(y_{i_1}, \ldots, y_{i_q}|\theta_0)$$

$$= F(y_{i_1}, \ldots, y_{i_p}|\theta_0)$$

for all values of $(y_{i_1}, \ldots, y_{i_q})$. $\square$

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

1. $\{\Phi_1, \Phi_2, \Phi_3|\theta\}$ is SO in $\theta$;
2. $\{\Phi_1, \Phi_2|\varphi_3\}$ is SO in $\varphi_3$;
3. $\{\Phi_j|\varphi_3\}, j = 1, 2, \text{ is SO in } \varphi_3$;
4. $\{\Phi_j|\varphi_3\}, j \neq k = 1, \ldots, 3, \text{ is SO in } \theta \text{ for all values of } \varphi_k$;
5. $\{\Theta|\varphi_j, \varphi_3\}, j = 1, 2, \text{ is SO in } \varphi_3 \text{ for all values of } \varphi_j$;
6. $\{\Phi_j|\varphi_k, \varphi_3\}, j \neq k = 1, 2, \text{ is SO in } \varphi_3 \text{ for all values of } \varphi_k$. 


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

1. It is first proved that \( \{\Phi_j|\theta\} \) is SO in \( \theta \). The functions \( \varphi \) are nondecreasing in each of their variables \( y \). The same holds for the (composite) functions \( f(\varphi|\theta) = I(\varphi > \varphi^*)\varphi(\varphi|\theta) \) for any value \( \varphi^* \), where \( I(\varphi > \varphi^*)\varphi(\varphi|\theta) \) is the indicator function that assigns the value 1 to \( \varphi > \varphi^* \) and the value 0 otherwise. From Lemma 2, it follows that \( E[I(\varphi > \varphi^*)\varphi(\varphi|\theta)] \) is nondecreasing in \( \theta \). This expectation is equal to \( 1 - F(\varphi^*|\theta) \), and therefore \( \{\Phi_j|\theta\} \) is SO in \( \theta \). Because \( \{\Phi_j|\theta\}, j = 1, \ldots, 3, \) are also independent, Lemma 11 gives the required result.

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

\[
F(\varphi_1, \varphi_2|\varphi_3) = \int F(\varphi_1, \varphi_2|\varphi_3, \theta) f(\theta|\varphi_3) \, d\theta
\]

From Lemmas 13.1 and 12 it follows that \( F(\varphi_1, \varphi_2|\theta) \) does not increase in \( \theta \) and that \( \{\Phi_3|\theta\} \) is SO in \( \theta \). If \( \varphi_3 \) is dichotomous, Lemma 6 shows that \( \{\Theta|\varphi_3\} \) is SO in \( \varphi_3 \). If \( \varphi_3 \) is a nondecreasing function of \( \sum_{i=q+1} y_i \), Lemma 6 has to be replaced by Grayson’s (1988; see also Hynh, 1994) result of MLR for the family of density functions associated with \( \{\sum_{i=q+1} Y_i|\theta\} \) that \( \{\Theta|\varphi_3\} \) is also SO in \( \theta \). In either case, Lemma 2 gives us the desired result.

3. Lemmas 13(2) and 12.

4. As \( \Phi_j \) and \( \Phi_k \) are independent given \( \theta \),

\[
F(\varphi_j|\theta, \varphi_k) = F(\varphi_j|\theta),
\]

and the result follows immediately.

5. If \( \varphi_3 \) is dichotomous, the result follows directly from Lemmas 13(4) and 6. If \( \varphi_3 \) is a nondecreasing function of \( \sum_{i=q+1} y_i \), Lemma 6 has to be replaced by Grayson’s result referred to above.

6. It holds that

\[
F(\varphi|\varphi_k, \varphi_3) = \int F(\varphi, \theta|\varphi_k, \varphi_3) \, d\theta
\]

From Lemma 13(1) it follows that \( F(\varphi_j|\theta) \) is not increasing in \( \theta \) whereas Lemma 13(5) shows that \( \{\Theta|\varphi_k, \varphi_3\} \) is SO in \( \varphi_3 \) for all values of \( \varphi_k \). Thus, Lemma 2 gives the desired result.

Applications to IRT

In this section, several results for response variables with distributions described by an IRT model are derived from the previous lemmas. The first theorem reviews properties of the conditional expectations and covariances between response variables whose distributions are determined by a single ability parameter. As before, these variables are denoted as \( \{Y_i|\theta\}, i = 1, \ldots, n \). These properties have been presented earlier in different contexts,
under different names, and sometimes only for special cases of the conditioning function \( \varphi_3 \) assumed below. They are formulated here as a coherent theorem with a new proof in which the previous lemmas are combined with well-known properties of expectations and covariances of random variables. The second theorem shows how the results can be extended to response variables with distributions determined by two parameters, which are denoted as \( \{ Y_i|\xi_1, \xi_2 \}, i = 1, \ldots, n \), again. Two additional theorems and corollaries specialize the results to models with multiple ability parameters as well as models with an ability and an item parameter. The latter category encompasses the cases of tests with a fixed item parameter (e.g., a standardized test) or test designs in which the parameter is random (e.g., domain-referenced or adaptive testing). For the currently popular form of adaptive testing with the maximum-information criterion for item selection and the expected a posteriori (EAP) ability estimator, it is proved for the 1-PL model that this form of testing shares the property of monotonic functions of the observed responses being SO in the ability parameter with tests with a fixed design.

**Theorem 1.** Let \( \{ Y_i|\theta \}, i = 1, \ldots, n, \) be independent and SO in \( \theta \), and let

\[
\begin{align*}
\varphi_1 &= \varphi_1(y_1, \ldots, y_p), \\
\varphi_2 &= \varphi_2(y_{p+1}, \ldots, y_q) \\
\varphi_3 &= \varphi_3(y_{q+1}, \ldots, y_n),
\end{align*}
\]

be nondecreasing in each of their arguments. If \( \varphi_3 \) is (1) dichotomous or a (2) nondecreasing function of \( \sum_{i=q+1}^n y_i \) with each \( y_i \) dichotomous, then:

1. \( E(\varphi_j|\varphi_3), j = 1, 2, \) is a nondecreasing function of \( \varphi_3 \);
2. \( \text{Cov}(\Phi_1, \Phi_2|\varphi_3) \geq 0; \)
3. \( \text{Cov}(\Phi_j, \Phi_k|\varphi_3) \geq 0, j, k = 1, \ldots, 3; j \neq k. \)

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

1. Lemmas 13(3) and 2.
2. Note that

\[
\text{Cov}(\Phi_1, \Phi_2|\varphi_3) = \text{Cov}((\Phi_1, E(\Phi_2|\Phi_1)|\varphi_3)).
\]

Let \( \tau(\varphi_1, \varphi_2) = E(\Phi_2|\varphi_1, \varphi_3). \) Lemmas 13(6) and 2 show that \( \tau \) is a nondecreasing function of \( \varphi_3. \) Using the fact that \( \text{Cov}(U, V) = E(U - E(U)V, following an argument in Casella and Berger (1990, sec. 4.7.2), it is proved that

\[
\begin{align*}
\text{Cov}(\Phi_1, \tau(\Phi_1, \varphi_3)|\varphi_3) &= E(\Phi_1 \tau(\Phi_1, \varphi_3)|\varphi_3) - E(\Phi_1|\varphi_3) E(\tau(\Phi_1, \varphi_3)|\varphi_3) \\
&= E[(\Phi_1 - E(\Phi_1)) \tau(\Phi_1, \varphi_3)|\varphi_3] \\
&\geq 0. \tag{1}
\end{align*}
\]

Because

\[
E[(\Phi_1 - E(\Phi_1)) \tau(\Phi_1, \varphi_3)|\varphi_3] \\
= E[(\Phi_1 - E(\Phi_1)) \tau(\Phi_1, \varphi_3) I_{(-\infty,0)}(\Phi_1 - E(\Phi_1))]|\varphi_3 \\
+ E[(\Phi_1 - E(\Phi_1)) \tau(\Phi_1, \varphi_3) I_{[0,\infty)}(\Phi_1 - E(\Phi_1))]|\varphi_3, \tag{2}
\]

\( \tau \) is nondecreasing, and observing the signs of the terms, the sum in (2) is not smaller than

\[
\begin{align*}
E[(\Phi_1 - E(\Phi_1)) \tau(\Phi_1, \varphi_3), \varphi_3) I_{(-\infty,0)}(\Phi_1 - E(\Phi_1))|\varphi_3] \\
+ E[(\Phi_1 - E(\Phi_1)) \tau(\Phi_1, \varphi_3), \varphi_3) I_{[0,\infty)}(\Phi_1 - E(\Phi_1))|\varphi_3] \\
= \tau E(\Phi_1|\varphi_3, \varphi_3) E(\Phi_1 - E(\Phi_1)|\varphi_3). \tag{3}
\end{align*}
\]
The inequality in (1) follows because the last factor in (3) is equal to zero for all values of \( \varphi_3 \).

3. It holds that

\[
\text{Cov} (\Phi_j, \Phi_k) = E(\text{Cov} (\Phi_j, \Phi_k|\Theta)) + \text{Cov} (E(\Phi_j|\Theta), E(\Phi_k|\Theta)).
\]

From part (2) of this theorem it follows that the first term is the expected value of a nonnegative statistic. As \( E(\Phi_j|\Theta) \) and \( E(\Phi_k|\Theta) \) are nondecreasing in \( \Theta \), 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 below when applications to IRT are discussed more directly. As already noted, individual parts of the theorem can be found in other places in the psychometric literature where they were established using different methods of proof than the one based on the lemmas and properties of covariances above, and sometimes for special versions of \( \varphi_3 \). The covariance property in the third part of the theorem was given earlier in Mokken (1971) and Holland (1981). Mokken based his proof on the notion of similarly ordered functions. Esary, Proschan, and Walkup (1967) define the covariance property in the third part of the theorem using the notion of associated random variables but give no conditions under which the property of association holds. Ahmed, León, and Proschan (1981, sect. 3.5) derive association between \( \Phi_1 \) and \( \Phi_2 \) under the same conditions as used here. Ellis (1993) proofs the third part of the theorem to be valid for any fixed subpopulation of examinees, provided that local homogeneity holds. Important references are Rosenbaum (1984) and Holland and Rosenbaum (1986) where a version of the second part of the theorem is given not based on the assumption on \( \varphi_3 \) made here. Finally, Junker (1993) establishes the first part of the theorem as the property of manifest monotonicity.

As noted above, the theorem involves several properties of response data that can be observed in large samples. The most important properties are summarized in Corollary 1. Some of these properties are also listed elsewhere (Rosenbaum, 1984; Sijtsma & Junker, 1994).

**Corollary 1.** For any dichotomous IRT model with a single ability parameter it holds that:

1. conditional item \( \pi \)-values given the (number-right) score on another item or another set of items are nondecreasing functions of the conditioning score;
2. the probability of passing a cutoff score on a subtest is a nondecreasing function of the number-right score on another subtest;
3. if \( \pi_i^H \) and \( \pi_i^L \) are the \( \pi \)-values of item \( i \) in subpopulations with a high and low number-right score on the other items, respectively, it holds that \( D = \pi_i^H - \pi_i^L \) is nonnegative;
4. all correlations between item scores are nonnegative;
5. all item-rest correlations (item scores discrimination indices) are nonnegative;
6. all previous properties hold in any subpopulation defined by number-right scores on other items or subtests;
7. all dichotomous conditioning variables may be dichotomizations of weighted number-right 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 con-
stist 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 status discrimination index 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. However, as will be shown below (Theorem 4), these analyses are not sufficient to distinguish between unidimensional and multidimensional IRT models.

The IRT model addressed in Theorem 1 and Corollary 1 are only assumed to have one parameter representing the ability of the examinee. The critical feature of the next theorem is the presence of more than one parameters needed to characterize the distributions of the responses. Due to this property, the results apply to response models that in addition to a person parameter have a parameter representing the difficulty of the item. The regular logistic models reviewed in the introductory part of this paper do have this parameter. Besides, the theorem applies to models with multiple ability parameters. Examples of models in the latter category are the logistic multidimensional IRT model (e.g., Reckase, 1997) and the Rasch model with a multivariate ability distribution presented in Glas (1992).

All assumptions made in the previous theorem are assumed to hold in the next theorem. The use of double indices is only to allow for later reference to the rows and columns of an item × person matrix with response data.

**Theorem 2.** Let $\{Y_{ij}\}$ be independent and SO in $\xi_1$ and $\xi_2$. $P$ and $Q$ are defined to be the sets of indices of two disjoint subsets of variables of $\{Y_{ij}; 1 = i, \ldots, n, j = 1, \ldots, m\}$. Let $\varphi_P = \varphi_P(\cdot)$ and $\varphi_Q = \varphi_Q(\cdot)$ be two functions nondecreasing in each of the variables with indices in $P$ and $Q$, respectively. It is assumed that $\varphi_Q(\cdot)$ is either dichotomous or nondecreasing in $(i,j) \in OY_{ij}$ with each $Y_{ij}$ dichotomous. Finally, $\{\xi_2\}$ and $\{\xi_3\}$ are assumed to be SO in $\xi_1$ and $\xi_2$, respectively. It holds that:

1. $E(\Phi_K|\xi_v), K = P, Q$, is a nondecreasing function of $\xi_v$ \( v = 1, 2; \)
2. $E(\Phi_P|\xi_v, \varphi_Q)$ is a nondecreasing function of $\xi_v$ \( v = 1, 2, \) for all values of $\varphi_Q$.

**Proof.** The two parts of this theorem follow from the previous theorem combined with the following lemmas:

1. Lemmas 13(1), 10, 12, and 2.
2. Theorem 2(1) and Lemma 13(4) (conditional independence).

First the case of IRT models with a person and an item parameter is addressed, where the item parameter is used to represent the difficulty of the item. At the end of this section the case of multiple ability parameters is returned to.

For models with an ability and a difficulty parameter, the theorem implies that the expected sum of item scores in any part of the data matrix is ordered in either parameter provided the parameters are independent or stochastically ordered themselves. Also, the same feature holds for the expected column and row sums of the data matrix. Thus, the theorem reveals sufficient conditions under which the row and column sums are ordered by the ability and the item parameter. These conditions have remained implicit in traditional Gutmann scalogram analyses where this order is typically taken for granted when analyzing the matrix with response data. The following corollary summarizes the results:
**Corollary 2.** If \( \{F(y|\theta, b)\} \) represents an IRT model with ability parameter \( \theta \) and difficulty parameter \( b \), then \( \{Y|\theta\} \) has the property of SO in \( \theta \) if: (a) the values of \( b \) are the same fixed constants for each examinee; or (b) \( b \) is a random parameter with an identical distribution for each examinee.

An example of the first design with the same fixed values for the item parameters is the well-known case of a standardized test administered to a sample of examinees. An example of the second design is a test sampled at random from an item pool and then administered to a sample of examinees (domain-referenced testing). In both cases the result in Theorem 2 holds because the ability and item parameters are independent and independence implies (weak) stochastic order.

Another example of a random test design is computerized adaptive testing. In this case the assumption of independence between the values of the ability and difficulty parameters is unlikely to hold since adaptive procedures invariably use item selection rules in which more able examinees tend to get more difficult items. However, the same feature suggests that the use of such rules may lead to distributions of the values of the difficulty parameter that are SO in the ability parameter. The following theorem shows that for a currently popular form of adaptive testing the difficulty parameter does have this property of SO in \( \theta \) and hence that any nondecreasing function of the response vector is SO in \( \theta \) as well. In this theorem, it is assumed that the pool of dichotomous items is large enough so that constraints on the availability of the values of the item parameter do not exist.

**Theorem 3.** For a fixed-length adaptive test from a 1-PL item pool based on maximum-information item selection and expected a posteriori (EAP) estimation of the ability parameter and for any prior for \( \theta \), it holds that any nondecreasing function of the response vector is SO in \( \theta \).

**Proof.** Let \( \varepsilon_k = -b_k \) be the value of the easiness parameter of the \( k \)th item in the adaptive test. Let \( \hat{\theta}^{k-1} \) be the ability estimator after \( k - 1 \) items have been administered and \( Y_1 = y_1, \ldots, Y_{k-1} = y_{k-1} \) the responses to these items. As the EAP estimator is defined as the expected value of the posterior distribution, it holds that \( \hat{\theta}^{k-1} = E(\theta|Y_1, \ldots, y_{k-1}) \). According to the maximum-information principle, the next item is selected such that \( \varepsilon_k = \hat{\theta}^{k-1} \). Therefore, \( \varepsilon_k = \varepsilon_k(y_1, \ldots, y_{k-1}) = E(\theta|Y_1, \ldots, y_{k-1}) \). However, since \( \{\theta|y_1, \ldots, y_{k-1}\} \) is SO in \( y_1, \ldots, y_{k-1} \) (Lemma 6), it follows that \( \varepsilon_k(y_1, \ldots, y_{k-1}) \) is nondecreasing in each of its arguments. Because \( \{Y_1, \ldots, Y_{k-1}|\theta\} \) is SO in \( \theta \), it follows that \( \{\varepsilon_k(Y_1, \ldots, y_{k-1})|\theta\} \) is SO in \( \theta \) (Lemma 2), and so does the difficulty parameter. Now, Lemma 10 applies and the response variables \( Y_i \) are SO in \( \theta \), and so is any nondecreasing function of these variables. This result holds for any prior \( f(\theta) \) because Lemma 6 does.

Observe that the result holds independently of the value of the parameter of the initial item in the test. The use of the EAP estimator in this theorem suggests application in the context of a full Bayesian approach to adaptive testing, with priors both for the ability and item parameters. The statistical framework needed for this approach is offered in Tsutakawa and Johnson (1990). In such an approach, the fact that the theorem holds for any prior for \( \theta \) allows for the possibility to chose the priors for the ability and item parameters to be independent. In Bayesian adaptive testing, independence of these priors is required to allow us to ignore the item selection mechanism and make full inference with respect to the ability estimator from the response vector (Mislevy & Wu, 1988).

The following theorem identifies a condition under which response variables in models with a two-dimensional ability structure are SO in the individual ability parameters:
Theorem 4. If \( \{F(y|\theta_1, \theta_2)\} \) represents a two-dimensional IRT model with ability parameters \( \theta_1 \) and \( \theta_2 \) and \( \{\Theta_2|\theta_1\} \) is a location family with a location parameter \( \varphi(\theta_1) \) that is a nondecreasing function of \( \theta_1 \), it holds that \( \{Y|\theta_1\} \) is SO in \( \theta_1 \).

Proof. The fact that location families have the property of SO is well documented (e.g., Lehmann, 1986, pp. 84–85). Thus, \( \{\Theta_2|\theta_1\} \) is SO in \( \varphi(\theta_1) \). Since \( \varphi(\theta_1) \) is nondecreasing in \( \theta_1 \), \( \Theta_2 \) is increasing in \( \theta_1 \). Hence, the condition in Lemma 10 is met, and \( \{Y|\theta_1\} \) is SO in \( \theta_1 \).

An important application is the case of the two-dimensional logistic IRT model (Reckase, 1997) with a bivariate normal ability distribution with nonnegative correlation and common conditional variance, where \( \varphi(\theta_1) \) is the regression function of \( \Theta_2 \) on \( \theta_1 \). Since the regression function is nondecreasing, the response variables are SO in \( \theta_1 \) and the same holds for the class of functions of these variables addressed in Lemma 13. In fact, because under these conditions the regression function of \( \Theta_1 \) on \( \theta_2 \) is also nondecreasing, the model has these properties in either ability variable.

Theorem 4 also shows why the classical item analysis procedures suggested by Corollary 1 can be used to detect items not fitting IRT models but may fail to detect violation of unidimensionality. Interestingly, whether or not these procedures do have the power to detect multidimensionality depends on the empirical ability distribution. If this distribution has the property of SO for at least one variable in another, the response variables are also SO in the latter. As a consequence, the regression and covariance properties between the classical statistics in Corollary 1 still hold, and multidimensionality will remain unnoticed. However, as Lemma 10 gives only a sufficient condition for SO, it is unclear what happens if none of the ability variables is SO in any of the others.

Conclusion

The main finding in this paper is that properties of stochastic order in response models with one parameter do not automatically generalize to the case of models with multiple parameters. If the response model has both ability and item parameters, it is the design of the test that determines whether generalization is possible. As shown in this paper, the case of the same item parameter values for each examinee (classical standardized test) involves responses and test scores that are SO both in the ability and difficulty parameters. The same holds for a stochastic testing design with the item parameter values drawn from a common distribution for all examinees (domain-referenced testing). However, if the item parameter values are stochastic and determined adaptively, a more complicated case arises. The values of the item parameters are now determined by such factors as the ability estimator, item selection criterion, choice of initial item rules, and stopping rule. Only one case could be identified in which the response vector is SO in \( \theta \) (Theorem 3).

If the model has multiple ability parameters, the question whether the properties of SO generalize to each individual ability parameter is no longer a matter determined by the design of the test but by the multivariate ability distribution of the examinees (Theorem 4). It becomes then a matter of empirical fact whether or not the condition of the ability variables being SO in each other hold. A statistical test for this condition can be based on the class of models with multivariate normal ability presented in Glas (1992).

References


*Manuscript received 3/20/96*

*Final version received 11/25/97*