The Interaction Model

Shelby J. Haberman
Educational Testing Service

13.1 Introduction

The interaction model, a generalization of the Rasch model (RM) for binary responses, retains many of the attractive features of the RM but does not assume local independence. Like the RM, the interaction model has simple sufficient statistics and a relatively straightforward interpretation. Computation of conditional maximum-likelihood estimates is a task of comparable difficulty to the corresponding computation for the Rasch model.

The interaction model can be used to test the validity of the RM (Rasch, 1960) by use of conventional conditional likelihood-ratio tests, and the interaction model can also be used to examine the size of the error of the RM in an information-theoretic sense (Gilula & Haberman, 1994, 1995). The interaction model has interest in its own right as an alternative to the 2PL model, which requires much less computation in large samples.

In Section 13.2, the interaction model is defined, and its relationship to the RM and to common loglinear models is discussed. In Section 13.3, computational methods are considered for the interaction model. Section 13.4 examines use of the interaction model to test the RM. Section 13.5 illustrates results by use of a multiple-choice examination with 45 items and 8,686 examinees. The examination is from the Praxis series of examinations for teacher training and certification.

13.2 Basic Properties of the Interaction Model

To define the interaction model, consider a test with \( I \geq 3 \) items and \( N \geq I \) examinees. For examinee \( v, 1 \leq v \leq N \), and item \( i, 1 \leq i \leq I \), let \( x_{vi} \) be 1 if the response to item \( i \) is correct, and let \( x_{vi} \) be 0 otherwise. Let \( x_v \) be the \( I \)-dimensional response vector with coordinates \( x_{vi}, 1 \leq i \leq I \). Let \( \theta_v \) be a one-dimensional measure of the ability of examinee \( v \), and assume that the \( \theta_v \) are random variables and the \( x_v \) are random vectors. Let the pairs \((x_v, \theta_v)\),
$1 \leq v \leq N$, be independent and identically distributed. Let the common distribution function of $\theta_v$ be $F$, let $p(x)$ be the probability that $x_v = x$ for $x$ in the set $\Gamma$ of $I$-dimensional vectors with coordinates 0 or 1, and let $p(x|\theta)$ be the conditional probability that $x_v = x$ given that $\theta_v = \theta$. Let $x_v = \sum_{i=1}^I x_{vi}$ be the number of items correctly answered by examinee $v$, and for $0 \leq r \leq I$, let $\Gamma(r)$ be the set of $x$ in $\Gamma$ such that $x = \sum_{i=1}^I x_i = r$, so that $x_v$ is in $\Gamma(r)$ if $r_v = x_v = r$. For $0 \leq r \leq I$, let $p^R(r)$ be the probability that $r_v = r$, and let $p_j$ be the probability that $x_{vi} = 1$ for $1 \leq i \leq I$. For $I$-dimensional vectors $a$ and $b$ with respective coordinates $a_i$ and $b_i$, $1 \leq i \leq I$, let

$$a'b = \sum_{i=1}^I a_ib_i.$$  

In the RM, for unknown real $\beta_i$, $1 \leq i \leq I$, it is assumed that the conditional probability that $x_{vi} = 1$ given that $\theta_v = u$ is $\{1 + \exp(-u + \beta_i)\}^{-1}$, and it is assumed that the $x_{vi}$, $1 \leq i \leq I$, are conditionally independent given $\theta_v$. It follows that

$$\log p_{\theta}(x) = \alpha(\beta, \theta) + \sum_{i=1}^I x_i(\theta - \beta_i) = \alpha(\beta, \theta) + x\theta - (\beta'x), \quad (13.1)$$

where $\beta$ is the $I$-dimensional vector with coordinates $\beta_i$, $1 \leq i \leq I$, and

$$\alpha(\beta, \theta) = -\sum_{i=1}^I \log[1 + \exp(\theta - \beta_i)]. \quad (13.2)$$

To identify parameters, the convention may be adopted that $\beta_1 = 0$. As is well known, if, for any $I$-dimensional vector $a$,

$$s_r(a) = \sum_{x \in \Gamma(r)} \exp(-a'x),$$

then the conditional probability that $x_v = x$ in $\Gamma(r)$ given that $r_v = r$ and $\theta_v = \theta$ is

$$p(x|r) = p(x)/p^R(r) = \frac{\exp(-\beta'x)}{s_r(\beta)},$$

so that $x_v$ and $\theta_v$ are conditionally independent given $r_v$ (Andersen, 1972).

In the interaction model, an additive interaction term is added for each pair of items, so that for unknown $\beta_i$ and $\gamma_i$, $1 \leq i \leq I$,

$$\log p_{\theta}(x) = \alpha(\beta, \gamma, \theta) + \sum_{i=1}^I x_i(\theta - \beta_i) + \sum_{i=2}^I \sum_{j=1}^{i-1} (\gamma_i + \gamma_j)x_ix_j$$

$$= \alpha(\beta, \gamma, \theta) + x\theta - \beta'x + (x - 1)\gamma'x, \quad (13.3)$$

where $\gamma$ is the $I$-dimensional vector of $\gamma_i$, $1 \leq i \leq I$, and