1. Introduction

An item response model attempts to describe the joint distribution, \( \text{pr}(X = x) \), of \( J \) dichotomously scored \((1 = \text{correct}, 0 = \text{incorrect})\) test items

\[ X = (X_1, X_2, \ldots, X_J) \]

in terms of a simple structure, namely

\[
\text{pr}(X = x) = \prod_{j=1}^{J} r_j(u)^{x_j} (1 - r_j(u))^{1-x_j} dF(u)
\]

where \( F(\cdot) \) is a cumulative probability distribution, and

\( r_j(u) \) is a monotone nondecreasing function of \( u \) with

\[ 0 \leq r_j(u) \leq 1 \quad \text{for all } u \text{ and for } j = 1, 2, \ldots, J. \]

Item response models have been extensively studied by, for example, Andersen (1980), Birnbaum (1968), Bock and Lieberman (1970), Cressie and Holland (1983), Lord (1952, 1980), Molenaar (1983), Rasch (1960), Tjur (1982), and others.

In practice, a parametric form is usually assumed for the \( r_j(u) \)'s and perhaps also for \( F(\cdot) \). Here, however, we are not concerned with the specific form of the model, but rather with testing the assumption of the existence of some functions \( \{r_j(u), j = 1, 2, \ldots, J\} \) and \( F(u) \) such that (1.1) and (1.2) hold.

Condition (1.1) states that the item responses \( X \) are conditionally independent given a latent variable \( U \). Although (1.1) is commonly called local independence, the term conditional independence is arguably more descriptive and more appropriate, particularly...
because the term local independence is often used to describe independence that is local in time or space (e.g., Schweder, 1970). The monotonicity of the item characteristic curves \( r_j(u) \) in (1.2) is implicit in most commonly used item response models. For brevity, conditions (1.1) and (1.2) will be called conditional independence and monotonicity or CI&M. Holland (1981, Theorem 3) developed a number of conditions which the distribution of observable item responses, \( p_r(X = x) \), must satisfy if conditions (1.1) and (1.2) are to hold. These conditions provide a basis for testing the CI&M assumptions, but they can be difficult to apply when large numbers of items are involved. In §2, we extend Holland’s Theorem to include conditions on subscores and other monotone item summaries, thereby permitting us to study larger numbers of items simultaneously. In §3, we briefly examine the theoretical properties of the performance of certain of these tests. Several examples are given in §4.

2. A Theorem: CI&M Imply Nonnegative Conditional Covariances Between Monotone Functions of Item Responses

2.1 Monotone Nondecreasing Item Summaries: Some Examples

A function of item responses, \( g(x) \), is said to be monotone nondecreasing if for all \( j \), and for all \( x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_J \),

\[
g(x_1, x_2, \ldots, x_{j-1}, 0, x_{j+1}, \ldots, x_J) \leq g(x_1, x_2, \ldots, x_{j-1}, 1, x_{j+1}, \ldots, x_J)
\]

that is, if an additional correct response (i.e., \( x_j \) becomes 1) does not decrease \( g(x) \). Some examples of monotone nondecreasing functions are:

(a) an item response, e.g.,
\[
g(x) = x_1;
\]

(b) the number correct,
\[
g(x) = \sum_{j=1}^{J} x_j;
\]

(c) subscores, e.g., the number correct among the first five items,
\[
g(x) = \sum_{j=1}^{5} x_j;
\]

(d) the “all correct” function,
\[
g(x) = 1 \text{ if } x_j = 1 \text{ for } j = 1, 2, \ldots, J, \text{ and } g(x) = 0 \text{ otherwise};
\]

(e) the “at least m correct” function
\[
g(x) = 1 \text{ if } \sum_{j=1}^{J} x_j \geq m, \text{ and } g(x) = 0 \text{ otherwise};
\]

(f) positively weighted scores,
\[
g(x) = \sum_{j=1}^{J} w_j x_j \text{ with fixed } w_j \geq 0;
\]

(g) positively weighted quadratics,
\[
g(x) = \sum_{i \leq j} w_{ij} x_i x_j \text{ with fixed } w_{ij} \geq 0;
\]