A BINOMIAL TEST FOR HIERARCHICAL DEPENDENCY

ROLAND B. GUAY AND GEORGE P. MCCABE

DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY

A binomial test for hierarchical dependency is presented. The null hypothesis is that all members of a population who possess a certain skill are a subset of the members who possess another skill. This hypothesis is basic to the writings of several prominent theorists, such as Gagné and Piaget. The model, assumptions, formula derivations, and procedures for the test are explained. An illustrative example is also provided.

Key words: hierarchical dependency; binomial test.

The notion of hierarchical dependencies is basic to the writings of several prominent theorists in education and psychology, for example, Gagné (1965) and Piaget (1963). The existence of a hierarchical dependency suggests that mental skills are acquired in an invariant sequence among all people, that is, all people must be able to display a lower-order prerequisite skill before being able to display a higher-order dependent skill. This notion may be stated by the hypothesis that, within a population, all members who possess a certain skill are a subset of the members who possess another skill.

Typically, research designed to assess the validity of posited hierarchical dependencies has used one of the following statistical procedures: Gagné and Paradise's (1961) index of positive transfer, Guttman's (1944) coefficient of reproducibility, or the phi-correlation coefficient. The major deficiencies of these statistical procedures for testing hypotheses of hierarchical dependencies are summarized in an article by White (1974). In a different article, White recommends the use of the test of inclusion (White & Clark, 1973) to replace existing statistical procedures.

White and Clark's (1973) test of inclusion requires the measurement of the posited higher-order and lower-order skills for a sample of subjects using two or more questions for each skill. The hierarchical dependency hypothesis being evaluated by the test of inclusion is rejected when the number of people observed who answer all of the higher-order skill questions correctly and concurrently answer all of the lower-order skills incorrectly exceeds a critical value. This critical value is estimated from the observed data assuming the dependence hypothesis under investigation is true.

Dayton and Macready (1976) develop probabilistic models for studying hierarchies. Their models involve slightly different assumptions than the present work and require iterative solutions to likelihood equations. In contrast, the technique presented here is comparatively simple and the computations required are trivial.

Many authors have written on topics related to hierarchies from different points of view. The text by Hildebrand, Laing, and Rosenthal (1977) presents a general framework for prediction analysis using cross-classified data. Bart and Krus (1973) introduce the idea of a tolerance level of disconfirmatory response patterns while Goodman (1959) and Sagi

For our procedure, measures of the posited higher-order and lower-order skills are obtained using one or more questions for each skill. The test involves the computation of expected values using a model developed to describe the distribution of the population assuming a particular dependency hypothesis. The test procedure is reduced to a one-sided binomial test.

Model and Assumptions

Suppose that a study is designed to investigate whether a hierarchical dependency exists between a skill A and a skill B. Assume that the literature suggests that skill A is a prerequisite to the development of skill B. To test the hypothesis that skill B is hierarchically dependent on skill A, a random sample of N subjects is drawn from the population of interest. The subjects answer I questions measuring skill A and J questions for skill B. These measures may be either selection- or supply-type.

The measures can be represented as

\[ X_1, X_2, \ldots, X_i, \ldots, X_I \] for skill A,

and

\[ Y_1, Y_2, \ldots, Y_j, \ldots, Y_J \] for skill B.

The value "1" denotes a correct response and the value "0" an incorrect response.

We assume that (a) if the subjects have skill A, they will be able to answer all the questions for skill A correctly (i.e., \( X_i = 1 \) for \( i = 1, \ldots, I \)); and (b) if they have skill B, they will answer all the questions for skill B correctly (i.e., \( Y_j = 1 \) for \( j = 1, \ldots, J \)). This assumption plays a crucial role in the analysis which follows. Situations for which this assumption does not appear to be valid lie outside the scope of the present work.

If a subject does not have skill A, the probabilities associated with the response to the \( i \)-th skill A question are given by

\[ P(X_i = 1) = p_{Ai}, \]

and

\[ P(X_i = 0) = 1 - p_{Ai} = q_{Ai}, \quad \text{for} \ i = 1, \ldots, I. \]

Similarly, if a subject does not have skill B, then

\[ P(Y_j = 1) = p_{ Bj}, \]

and

\[ P(Y_j = 0) = 1 - p_{ Bj} = q_{ Bj}, \quad \text{for} \ j = 1, \ldots, J. \]

In addition, if a subject does not have skill A, responses to the skill A questions are assumed to be independent. Similarly for skill B. If neither skill is possessed, the \( X \)'s and \( Y \)'s are assumed to be independent.

The values of \( p_{Ai} \) and \( p_{Bj} \) are assumed to be known. For multiple choice items, reasonable values are obtained by assuming that all choices are equally likely. In some cases, data on persons known to lack the skills may be available for estimation of these quantities. It will be shown that the test statistic does not depend upon the \( p_{Ai} \). However,