Rasch Maximum Likelihood Estimation for Theta and W-scores with Panel Study of Income Dynamics Woodcock-Johnson Revised Achievement Raw Scores

Ezekiel J Dixon-Román, University of Pennsylvania
Rasch Maximum Likelihood Estimation for Theta and W-scores with Panel Study of Income Dynamics Woodcock-Johnson Revised Achievement Raw Scores

Ezekiel J. Dixon-Román

University of Pennsylvania

September 29, 2014

Acknowledgements: This was supported by a grant from the American Educational Research Association, which receives funds for its “AERA Grants Program” from the National Science Foundation and the National Center for Education Statistics of the Institute of Education Sciences (U.S. Department of Education) under NSF Grant #DRL-0634035. I would like to thank Charles Lewis for helpful comments. All correspondence should be sent to Ezekiel Dixon-Román, Penn School of Social Policy & Practice, University of Pennsylvania, 3701 Locust Walk, Philadelphia, PA 19104-6214, email: ezekiel@sp2.upenn.edu.
Rasch Maximum Likelihood Estimation for Theta and W-scores with Panel Study of Income Dynamics Woodcock-Johnson Revised Achievement Raw Scores

This appendix explains the estimation of the Rasch maximum likelihood estimated thetas. I begin by explaining (with equations) Rasch maximum likelihood estimation (MLE) for theta values and what BILOG, in fact, does. Given this, I briefly explain why a program was written and used in True Basic in order to estimate MLE theta values.

Suppose we have a test consisting of \( n \) items, responses to each of which are scored 0 for incorrect and 1 for correct. Let \( y_i \) denote the score for item \( i \). Moreover, suppose that each item is considered to measure a latent variable, denoted by \( \theta \), and that the responses to the items are conditionally independent, given \( \theta \). If the conditional probability that \( y_i = 1 \) given \( \theta \) is denoted by \( P_i(\theta) \), then the joint probability of observing a response vector \( y \) may be written as

\[
\Pr(Y = y|\theta) = \prod_{i=1}^{n} P_i(\theta)^{y_i} [1 - P_i(\theta)]^{1-y_i}.
\]  

The Rasch model for binary items specifies

\[
P_i(\theta) = \frac{\exp(\theta - \beta_i)}{1 + \exp(\theta - \beta_i)},
\]

where \( \beta_i \) denotes the “difficulty” of item \( i \). Suppose that the \( \beta_i \) are “known” (actually, estimated from responses to the test items from a sample of test takers). Then the maximum likelihood estimate for \( \theta \) based on a response vector \( y \) is the value that maximizes the “likelihood” given in Equation 1, using the definition for \( P_i(\theta) \) given in Equation 2. It turns out that this estimate is the solution to the equation

\[
\sum_{i=1}^{n} P_i(\theta) = \sum_{i=1}^{n} y_i = s,
\]
where \( s \) denotes the “number correct” score on the test. Equation 3 will have a unique solution (denoted by \( \hat{\theta} \)) for all values of \( s \) between 1 and \( n - 1 \). However, there is no solution when \( s = 0 \) or \( s = n \), so typically a large negative value is assigned to \( \hat{\theta} \) when \( s = 0 \) and a corresponding large positive value is used when \( s = n \). (We would choose the value of \( \hat{\theta} \) used when \( s = 0 \) to be smaller than \( \hat{\theta} \) for \( s = 1 \) and choose the value of \( \hat{\theta} \) used when \( s = n \) to be larger than \( \hat{\theta} \) for \( s = n - 1 \).)

The item response theory software BILOG may be used to solve Equation 3, given item difficulty estimates and response vectors. Apparently, however, it truncates the values for \( \hat{\theta} \) to lie in the interval \((-4.0, 4.0)\). This may be reasonable when the scale for \( \theta \) has been set so that, for some population of interest, the mean of \( \theta \) is 0 and the standard deviation is 1. However, for the Rasch model as given in Equation 2, the units have already been set (by taking the item “slopes” all equal to 1), so the standard deviation of \( \theta \) may be quite different from 1. If this standard deviation is substantially larger than 1, then values for \( \hat{\theta} \) outside the interval \((-4.0, 4.0)\) would not be unusual, even if the mean of \( \theta \) has been set to 0. To deal with this problem, I used True Basic in order to solve Equation 3. Unlike BILOG, it obtains distinct \( \hat{\theta} \) values for all scores from 0 to \( n \).

The \( W \)-scores were then calculated by multiplying the Rasch MLE theta by a constant (i.e., 9.1024) then adding a constant (i.e., 140) as suggested by Woodcock and Dahl (1971).
References