MAXLOG: A computer program for the estimation of the parameters of a multidimensional logistic model

ROBERT L. MCKINLEY and MARK D. RECKASE
American College Testing Program, Iowa City, Iowa

Latent trait theory has proved to be a powerful tool for mental measurement. Latent trait models, such as those proposed by Birnbaum (1968) and Rasch (1960), have been widely applied. (For a discussion of the advantages of latent trait theory and applications of latent trait models, see Hambleton, Swaminathan, Cook, Eignor, & Gifford, 1978.) One serious limitation of these models has been the requirement that the tests to which they are applied be unidimensional. It is difficult to construct a unidimensional test, and often, a unidimensional test is not desired, as is often the case with achievement testing.

Reckase and McKinley (Note 1) provide a theory for the use of latent trait models with multidimensional tests and propose an extension of the two-parameter logistic (2PL) latent trait model for use with dichotomously scored multidimensional tests. The MAXLOG program is designed to perform maximum likelihood estimation of the parameters of that model.

The Model. The program estimates the parameters of the model given by

$$P(u_{ij}) = \frac{\exp[(d_i + A_i\theta_j')u_{ij}]}{1 + \exp(d_i + A_i\theta_j')}, \quad (1)$$

where θ_j is a column vector of ability parameters for Examinee j , A_i is a row vector of item discrimination parameters for Item i , d_i is an item difficulty parameter for Item i , and $P(u_{ij})$ is the probability of Response u_{ij} to Item i by Examinee j . The θ and A vectors are of order m , where m is the number of dimensions to be modeled.

Estimation Procedure. The estimation procedure is designed to maximize the criterion function given by

$$L = \ln \prod_{j=1}^N \prod_{i=1}^n P(u_{ij}), \quad (2)$$

where L is the natural logarithm of the response matrix likelihood function given the current parameter estimates, N is the number of examinees, and n is the number of items.

This work was performed under Contract N0014-81-K0817 with the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research. The authors' mailing address is: American College Testing Program, P.O. Box 168, Iowa City, Iowa 52243.

The procedure is organized into cycles, with each cycle having two stages. In the first stage, the ability parameters are estimated, and in the second stage, the item parameters are estimated. Each item is treated independently of all other items, and each examinee is treated independently of all other examinees.

The estimation procedure is as follows. Let U represent the response string for the item or examinee of interest, and let Z represent a vector of parameters for that item or examinee. Estimates of Z are obtained by taking the first derivative of the likelihood function for U given Z , setting it equal to zero, and solving for Z . For an item, U represents the responses to the item by all the examinees. For an examinee, U represents the responses to all the items by that examinee. For an item, Z is composed of d and A , whereas for an examinee, Z is composed of θ .

The procedure results in a set of equations that are nonlinear in Z . Therefore, estimates are obtained using an iterative procedure based on the Newton-Raphson technique, which requires the matrix of second derivatives with respect to Z . Using initial estimates for Z , new estimates are obtained using

$$Z = Z_0 + [H^{-1}(\delta L/\delta Z)]|_{Z_0}, \quad (3)$$

where Z_0 is a vector of initial estimates of Z , Z is a vector of updated estimates, and H is the matrix of second derivatives. Each time that updated estimates are obtained, they are used as initial estimates for the next iteration. Iteration continues until the estimates become stable or until a maximum number of iterations is reached. Estimates have become stable when they change less than a specified amount during an iteration.

Each time a cycle is completed, the criterion function given by Equation 2 is evaluated and compared to the results from the previous cycle. Cycling continues until the value of the criterion function increases by less than a specified value or until a maximum number of cycles is reached.

Input. There are two types of input to the program. One type includes information such as a title for the run, input formats for the response data, and a number of control parameters. The minimum input of this type is the number of examinees and the number of items. All other control parameters have default values. Options available to the user allow the selection of whether item parameters, person parameters, or both item and person parameters will be estimated. Also, the user may specify the maximum number of cycles and iterations, as well as criterion values for determining whether iteration or cycling should be terminated before the maximum number is reached.

The other type of input includes the response data and any initial parameter values that might be required.

For example, if only item parameters are to be estimated, values for the ability parameters must be supplied; if only ability parameters are to be estimated, values for the item parameters must be supplied. If both item and person parameters are to be estimated, initial item discrimination estimates must be supplied. Varimax rotated factor loadings work well as initial item discrimination parameter estimates, but less informative initial estimates will also work.

Output. Output includes a listing of the user's input, information for evaluating the stability of the estimates (whether the estimation procedure converged), and the final estimates obtained and the standard errors of estimates for each estimated parameter. In addition, the results of an analysis of residuals are printed.

Language. The program is written in FORTRAN IV and has been tested on an IBM 370. No external routines are required except those contained in the standard FORTRAN library.

Availability. A copy of the program and a user's guide are available at no cost from Robert L. McKinley, ACT National Office, P.O. Box 168, Iowa City, Iowa

52243. Requests should be accompanied by a tape; otherwise, a fee for the cost of a tape will be charged.

REFERENCE NOTE

1. Reckase, M. D., & McKinley, R. L. *Some latent trait theory in a multidimensional latent space*. Paper presented at the 1982 Item Response Theory/Computerized Adaptive Testing Conference, Wyzata, Minnesota, July 1982.

REFERENCES

- BIRNBAUM, A. Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores*. Reading, Mass: Addison-Wesley, 1968.
- HAMBLETON, R. K., SWAMINATHAN, H., COOK, L. L., EIGNOR, D. R., & GIFFORD, J. A. Developments in latent trait models, technical issues, and applications. *Review of Education Research*, 1978, **48**, 467-510.
- RASCH, G. *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research, 1960.

(Accepted for publication October 28, 1982.)