IRT in Mplus

November 14, 2013

In this note we describe several of the IRT modeling features implemented in Mplus, namely the the item characteristic curves, the item information curves, the total information curve, item difficulty parameter and item discrimination parameter.

1 ICC curves

1.1 Logit Link, ML/MLR/MLF Estimators

Let U_i be a categorical indicator for a latent factor f in the presence of a categorical latent class variable C. The item characteristic curves (ICC) for the item U_i, given that $C = k$ are computed as follows using the logistic model. If the category j is the first category

$$P_{ijk}(f) = P(U_i = j|f, C = k) = \frac{1}{1 + Exp(-\tau_{ijk} + \lambda_{ik} f)}. \tag{1}$$

If the category j is the last category

$$P_{ijk}(f) = P(U_i = j|f, C = k) = 1 - \frac{1}{1 + Exp(-\tau_{ij-1k} + \lambda_{ik} f)}. \tag{2}$$

If the category j is a middle category

$$P_{ijk}(f) = P(U_i = j|f, C = k) =$$

$$\frac{1}{1 + Exp(-\tau_{ijk} + \lambda_{ik} f)} - \frac{1}{1 + Exp(-\tau_{ij-1k} + \lambda_{ik} f)}. \tag{3}$$

In the presence of other covariates/other latent variables X the formulas are modified as follows. If the category j is the first category

$$P_{ijk}(f) = P(U_i = j|f, C = k, X = x) = \frac{1}{1 + Exp(-\tau_{ijk} + \lambda_{ik} f + \beta_{ik} x)}. \quad (4)$$

If the category j is the last category

$$P_{ijk}(f) = P(U_i = j|f, C = k, X = x) =$$

$$1 - \frac{1}{1 + Exp(-\tau_{ij-1k} + \lambda_{ik} f + \beta_{ik} x)}. \quad (5)$$

If the category j is a middle category

$$P_{ijk}(f) = P(U_i = j|f, C = k, X = x) =$$

$$\frac{1}{1 + Exp(-\tau_{ijk} + \lambda_{ik} f + \beta_{ik} x)} - \frac{1}{1 + Exp(-\tau_{ij-1k} + \lambda_{ik} f + \beta_{ik} x)}. \quad (6)$$

1.2 Probit Link, ML/MLR/MLF Estimators

Let Φ be the standard normal cumulative distribution function. The ICC curves are given as follows. If the category j is the first category

$$P_{ijk}(f) = P(U_i = j|f, C = k, X = x) = \Phi(\tau_{ijk} - \lambda_{ik} f - \beta_{ik} x). \quad (7)$$

If the category j is the last category

$$P_{ijk}(f) = P(U_i = j|f, C = k, X = x) =$$

$$1 - \Phi(\tau_{ij-1k} - \lambda_{ik} f - \beta_{ik} x). \quad (8)$$

If the category j is a middle category

$$P_{ijk}(f) = P(U_i = j|f, C = k, X = x) =$$

$$\Phi(\tau_{ijk} - \lambda_{ik} f - \beta_{ik} x) - \Phi(\tau_{ij-1k} - \lambda_{ik} f - \beta_{ik} x). \quad (9)$$

1.3 Probit Link, WLS/WLSM/WLSMV/ULS Estimators, Theta Parametrization

In this situation the model does not include latent categorical variable C however multiple group models are included. Let G denote the group variable. With the Theta parametrization the residual parameter θ_{ik} is an actual parameter in the model. For basic models this parameter is fixed to 1 since it will not be identified without model restrictions, however for multiple group and growth models the parameter could be identified. If these parameters are not printed in the results section that means that they are fixed to 1. The ICC curves are given as follows. If the category j is the first category

$$P_{ijk}(f) = P(U_i = j|f, G = k, X = x) = \Phi\left(\frac{\tau_{ijk} - \lambda_{ik} f - \beta_{ik} x}{\sqrt{\theta_{ik}}}\right). \qquad (10)$$

If the category j is the last category

$$P_{ijk}(f) = P(U_i = j|f, G = k, X = x) =$$
$$1 - \Phi\left(\frac{\tau_{ij-1k} - \lambda_{ik} f - \beta_{ik} x}{\sqrt{\theta_{ik}}}\right). \qquad (11)$$

If the category j is a middle category

$$P_{ijk}(f) = P(U_i = j|f, G = k, X = x) =$$
$$\Phi\left(\frac{\tau_{ijk} - \lambda_{ik} f - \beta_{ik} x}{\sqrt{\theta_{ik}}}\right) - \Phi\left(\frac{\tau_{ij-1k} - \lambda_{ik} f - \beta_{ik} x}{\sqrt{\theta_{ik}}}\right). \qquad (12)$$

1.4 Probit Link, WLS/WLSM/WLSMV/ULS Estimators, Delta Parametrization

With the Delta parametrization the θ_{ik} are not actual parameters but are dependent parameters that are obtained from the following equation

$$\theta_{ik} = \Delta_{ik}^{-2} - Var(\lambda_{ik} f)$$

where Δ_{ik} are actual parameters that can be either free or fixed. Again the Δ_{ik} are typically not identifiable and are fixed to 1, however in growth and multiple group models the parameter can be free and identified. When the Δ_{ik} parameters are not present in the results, they are fixed to 1. The θ_{ik} parameters are always reported in the results section and are typically smaller than 1. For example when the Δ_{ik} parameters are fixed to 1 the θ_{ik} are smaller than 1. The ICC curves are given as in the previous section.

2 IIC curves

The item information curves (IIC) for a categorical indicator U_i and a latent factor f in class $C = k$ (or group k) is computed as in Samejima (1974). Define for $1 \leq j \leq l-1$

$$Q_{ijk} = \sum_{r=1}^{j} P_{irk}. \tag{13}$$

and $Q_{i0k} = 0$, $Q_{ilk} = 1$. The IIC is defined as follows

$$I_{ik}(f) = \sum_{r=1}^{l} \frac{(\partial P_{irk}/\partial f)^2}{P_{irk}}. \tag{14}$$

For the ML/MLF/MLR estimators with the logit link functions the IIC curve is given by

$$I_{ik}(f) = \lambda_{ik}^2 \sum_{r=1}^{l} \frac{(Q_{irk}(1-Q_{irk}) - Q_{i,r-1,k}(1-Q_{i,r-1,k}))^2}{P_{irk}}. \tag{15}$$

For binary items the above formula reduces to

$$I_{ik}(f) = \lambda_{ik}^2 P_{i1k}(1-P_{i1k}). \tag{16}$$

For the ML/MLF/MLR estimators with the probit link functions, we use the logit to probit approximation and give the IIC curve by

$$I_{ik}(f) = 3.29 \cdot \lambda_{ik}^2 \sum_{r=1}^{l} \frac{(Q_{irk}(1-Q_{irk}) - Q_{i,r-1,k}(1-Q_{i,r-1,k}))^2}{P_{irk}}. \tag{17}$$

For the WLS/WLSM/WLSMV/ULS estimators with the probit link functions and either theta or delta parametrization the IIC curve is given by

$$I_{ik}(f) = 3.29 \cdot \frac{\lambda_{ik}^2}{\theta_{ik}} \sum_{r=1}^{l} \frac{(Q_{irk}(1-Q_{irk}) - Q_{i,r-1,k}(1-Q_{i,r-1,k}))^2}{P_{irk}}. \tag{18}$$

The total information function is obtained by adding all item information functions and the prior information

$$I_k(f) = 1/\psi + \sum_i I_{ik}(f). \tag{19}$$

where ψ is the variance of the factor f. The term $1/\psi$ is minus the second derivative of the log-likelihood of the prior. The meaning of prior here is simply the part of the likelihood that specifies the factor as a $N(0,1)$ latent variable or more generally as a $N(\alpha,\psi)$ latent variable. The above formulas are obtained using the logit link function and are exact in this case. They are simply an approximation for the probit link function. The constant 3.29 used with the probit link function is simply the $\pi^2/3$ constant which is needed to adjust the scale of the loadings.

The information function $I(f)$ can be used to calculate approximate standard errors for the factor score estimates

$$SE(f) = \frac{1}{\sqrt{I(f)}}.$$

This is because $I(f)$ is the expected information function. To obtain the standard errors for the factor score in the presence of missing data the total information function in the above formula is replaced by the sum of the IIC for the indicators that are present for that observation. These factor score standard errors can also be obtained in Mplus directly using the SAVEDATA command and the ML/MLF/MLR estimators with numerical integration from the estimated posterior distribution of the factor. The two methods differ to some extent but generally yield approximately the same results. A third method for computing the factor score standard errors is with the Bayes estimator where the posterior distribution for each factor can also be estimated.

3 IRT Parameterization

For binary items with a single factor we provide the parameter estimates also in the traditional IRT scale. Let the factor mean be α and the factor variance be ψ. Thus $f = \alpha + \sqrt{\psi}\theta$ where θ is the IRT standard normal latent variable with mean 0 and standard deviation 1. For the ML/MLF/MLR estimators with the logit link function

$$P(U_i = 1|f) = \frac{1}{1 + Exp(\tau_{ik} - \lambda_{ik} f)} = \frac{1}{1 + Exp(-a_{ik}(\theta - b_{ik}))} \tag{20}$$

where a_{ik} is the item discrimination parameter and b_{ik} is the item difficulty parameter. These parameters are computed as follows

$$a_{ik} = \lambda_{ik}\sqrt{\psi} \tag{21}$$

$$b_{ik} = \frac{\tau_{ik} - \lambda_{ik}\alpha}{\lambda_{ik}\sqrt{\psi}}. \tag{22}$$

For the other estimations, links and parametrization the IRT parametrization is obtained by the same approach. The resulting formulas for b_{ik} is the same as (22), while the parameter a_{ik} is obtained as follows. For the ML/MLF/MLR estimators with the probit link function

$$a_{ik} = \lambda_{ik}\sqrt{\psi}. \tag{23}$$

For the WLS/WLSM/WLSMV/ULS estimators with the probit link functions and theta parametrization

$$a_{ik} = \frac{\lambda_{ik}\sqrt{\psi}}{\sqrt{\theta_{ik}}} \tag{24}$$

For the WLS/WLSM/WLSMV/ULS estimators with the probit link functions and delta parametrization

$$a_{ik} = \frac{1}{\sqrt{\Delta_{ik}^{-2}\lambda_{ik}^{-2}\psi^{-1} - 1}} \tag{25}$$

The standard errors of these parameters are computed by the delta method.

4 References

The following references can be used for additional information on the IRT model.

Baker, F.B. & Kim, S.H. (2004). Item response theory. Parameter estimation techniques. Second edition. New York: Marcel Dekker.

Bock, R.D. (1997). A brief history of item response theory. Educational Measurement: Issues and Practice, 16, 21-33.

du Toit, M. (2003). IRT from SSI. Lincolnwood, IL: Scientific Software International, Inc. (BILOG, MULTILOG, PARSCALE, TESTFACT)

Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Mahwah, NJ: Erlbaum.

Hambleton, R.K. & Swaminathan, H. (1985). Item response theory. Boston: Kluwer-Nijhoff.

MacIntosh, R. & Hashim, S. (2003). Variance estimation for converting MIMIC model parameters to IRT parameters in DIF analysis. Applied Psychological Measurement, 27, 372-379.

Muthen, B. (1985). A method for studying the homogeneity of test items with respect to other relevant variables. Journal of Educational Statistics, 10, 121-132. (#13)

Muthen, B. (1988). Some uses of structural equation modeling in validity studies: Extending IRT to external variables. In H. Wainer & H. Braun (Eds.), Test Validity (pp. 213-238). Hillsdale, NJ: Erlbaum Associates. (#18)

Muthen, B. (1989). Using item-specific instructional information in achievement modeling. Psychometrika, 54, 385-396. (#30)

Muthen, B. (1994). Instructionally sensitive psychometrics: Applications to the Second International Mathematics Study. In I. Westbury, C. Ethington, L. Sosniak & D. Baker (Eds.), In search of more effective mathematics education: Examining data from the IEA second international mathematics study (pp. 293-324). Norwood, NJ: Ablex. (#54)

Muthen, B. & Asparouhov, T. (2002). Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus. Mplus Web Note #4 (www.statmodel.com).

Muthen, B., Kao, Chih-Fen & Burstein, L. (1991). Instructional sensitivity in mathematics achievement test items: Applications of a new IRT-based detection technique. Journal of Educational Measurement, 28, 1-22. (#35)

Muthen, B. & Lehman, J. (1985). Multiple-group IRT modeling: Applications to item bias analysis. Journal of Educational Statistics, 10, 133-142. (#15)

Samejima, F. (1972) A general model for free-response data. Psychometrika Monograph Supplement, No.18

Samejima, F. (1974). Normal ogive model on the continuous response level in the multidimensional latent space. Psychometrika, 39, 111-121.

Takane, Y. & DeLeeuw, J. (1987). On the relationship between item response theory and factor analysis of discretized variables. Psychometrika, 52, 393-408.