6-26-2018

Fitting the Rasch Model under the Logistic Regression Framework to Reduce Estimation Bias

Tianshu Pan

Pearson, tianshu.pan@pearson.com

Follow this and additional works at: https://digitalcommons.wayne.edu/jmasm

Part of the Applied Statistics Commons, Social and Behavioral Sciences Commons, and the Statistical Theory Commons

Fitting the Rasch Model under the Logistic Regression Framework to Reduce Estimation Bias

Tianshu Pan
Pearson
San Antonio, TX

The purpose of this study was to show how and why the Rasch model can be fitted under the logistic regression framework. Then a penalized maximum likelihood (Firth, 1993) for logistic regression models can be used to reduce ML biases when fitting the Rasch model. These conclusions are supported by a simulation study.

Keywords: The Rasch model, logistic regression, maximum likelihood, penalized maximum likelihood

Introduction

The Rasch model (Rasch, 1960) has been widely used in psychological and educational assessments. Those who know the binary logistic regression and the Rasch models might notice the similarity between them, i.e. both have the mathematical expression of the logit or logistic function. Accordingly, Wright (1993) used a Rasch model to do logistic regression for discrete-time survival analysis. Uekawa (2005) further used an example to show that the parameters of a binary logistic regression and the Rasch models are one-to-one correspondent. However, those studies identified only the similarity of model expressions between the logistic regression and the Rasch models, and did not show whether they are equivalent or how to make them equivalent.

Kamata (1998, 2001) first connected the standard Rasch model with a special multilevel logistic regression model, and found that they also have similar mathematical expressions. However, person ability is fixed in the Rasch model but

is random in the multilevel model (Kamata, 1998, 2001; De Boeck & Wilson, 2004). Thus the multilevel logistic regression and the Rasch models are not equivalent.

Skrondal and Rabe-Hesketh (2004) argued that the Rasch model is a special case of the fixed-effect logistic regression under the conditional maximum likelihood. However, the logistic regression models are usually fitted using the maximum likelihood (ML).

This article will discuss the relationship between the logistic regression and the Rasch model under ML. In the following sections, the article will first illustrate how and why the standard Rasch model is equivalent to a special logistic regression model under ML. Then, because the penalized maximum likelihood (PML) can reduce the ML bias of fitting logistic regression models (Firth, 1993), it may be applied directly to reduce the ML bias of fitting the Rasch model. Lastly, a simulation study is used to show that the logistic regression and the Rasch model software gives comparable parameter estimates using ML, and Firth's PML can reduce ML bias in the estimation of the Rasch model.

Equivalence between the Rasch Model and Logistic Regression

Suppose I persons take J dichotomous-scored items in a test, and a standard Rasch model is used to estimate item and person parameters; the model will then be specified as follows:

$$\Pr\left(y_{ij}=1\right)=\frac{\exp\left(\theta_i-b_j\right)}{1+\exp\left(\theta_i-b_j\right)}$$

where y_{ij} is the score of person i on item j, person parameter θ_i is the ability of person i, item parameter b_j is the difficulty of item j, $i = 1,\ldots, I$, and $j = 1,\ldots, J$. The equation can be rearranged as

$$\log\frac{\Pr\left(y_{ij}=1\right)}{1-\Pr\left(y_{ij}=1\right)}=\theta_i-b_j \tag{1}$$

If $I + J$ dummy variables are used to indicate the scores of different persons on different items, the following logistic regression model can fit the scores of all persons on all items:

$$\log \frac{\Pr\left(y_{ij}=1\right)}{1-\Pr\left(y_{ij}=1\right)} = \sum_{n=1}^{I}\beta_{1n}x_{1n,ij} + \sum_{m=1}^{J}\beta_{2m}x_{2m,ij} \tag{2}$$

where $x_{1n} = 1$ if $n = i$, $x_{2m} = 1$ if $m = j$, and are otherwise 0, and β_{1n} and β_{2m} are, respectively, the coefficients of person and item variables. So for any given i and j, $\Sigma_n\beta_{1n}x_{1n} + \Sigma_m\beta_{2m}x_{2m} = \beta_{1i} + \beta_{2j}$. Then

$$\log \frac{\Pr\left(y_{ij}=1\right)}{1-\Pr\left(y_{ij}=1\right)} = \beta_{1i} - \beta_{2j} \tag{3}$$

By comparing equation (1) with (3), coefficients β_{1i} and β_{2j} in the logistic regression model are found to be respectively correspondent to parameters θ_i and $-b_j$ in the Rasch model.

Therefore a one-to-one correspondence holds between the parameters of the two models. How about their parameter estimates? Usually ML is used to estimate logistic regression models. Wright and collaborators (e.g., Wright & Douglas, 1977; Wright & Panchapakesan, 1969; Wright & Stone, 1979) described how to use ML to fit the Rasch model, which is implemented in WINSTEPS (Linacre, 2008). Because they have one-to-one correspondent parameters, ML should not give meaningfully different estimates for them.

To estimate the logistic regression model shown by equation (2) using ML, a likelihood or score equation is given as follows (Agresti, 2002):

$$g\left(\beta_k\right) = \sum_{i,j}\left(y_{ij} - \pi_{ij}\right)x_{k,ij} = 0 \tag{4}$$

where $\pi_{ij} = \Pr(y_{ij} = 1)$ and $k = 11, 12,\ldots, 1I, 21, 22,\ldots, 2J$. The first digits 1 and 2 are used to differentiate the person and item parameters, respectively. Then for the person coefficients,

$$g\left(\beta_{1n}\right) = \sum_{i,j}\left(y_{ij} - \pi_{ij}\right)x_{1n,ij} = \sum_{i,j}y_{ij}x_{1n,ij} - \sum_{i,j}\pi_{ij}x_{1n,ij} = \sum_{j}y_{nj} - \sum_{j}\pi_{nj} \tag{5}$$

because all independent variables are dummy variables. By the same way, for the item coefficients,

$$g\left(\beta_{2m}\right)=\sum_{i,j}\left(y_{ij}-\pi_{ij}\right)x_{2m,ij}=\sum_{i,j}y_{ij}x_{2m,ij}-\sum_{i,j}\pi_{ij}x_{2m,ij}=\sum_{i}y_{im}-\sum_{i}\pi_{im} \qquad (6)$$

Namely, parameter estimates are the solutions of the following equations:

$$\begin{cases} \sum_{j}y_{nj}-\sum_{j}\pi_{nj}=0 \\ \sum_{i}y_{im}-\sum_{i}\pi_{im}=0 \end{cases} \qquad (7)$$

The first terms of the equations are actually the total of observed scores of a person and an item, respectively, and the second ones are the expected person or item scores under the Rasch framework.

According to Wright and collaborators (Wright & Douglas, 1977; Wright & Stone, 1979), parameter estimates of the Rasch model are solutions of the following equations:

$$\begin{cases} r_{n}-\sum_{j}\pi_{nj}=0 \\ t_{m}-\sum_{i}\pi_{im}=0 \end{cases} \qquad (8)$$

where r_n is the total of scores person n obtains on all items and t_m is the total of scores all persons obtain on item m. Using the notations of equation (7), $r_n = \Sigma_j y_{nj}$ and $t_m = \Sigma_i y_{im}$. Therefore the logistic regression model specified above and the Rasch model have the same ML score equations.

However, the model specified in equation (2) can be fitted only with some constraint. In the Rasch model or WINSTEPS, a sum-to-zero constraint is imposed on item parameters, i.e., $\Sigma_j \beta_j = 0$ (Wright & Douglas, 1977; Wright & Stone, 1979). If the same constraint is imposed on equation (2), then

$$\sum_{m=1}^{J}\beta_{2m}=0 \Rightarrow \beta_{2J}=-\sum_{m=1}^{J-1}\beta_{2m} \qquad (9)$$

Substituting the above equation into equation (2), and the following equation is obtained:

$$\log \frac{\Pr\left(y_{ij}=1\right)}{1-\Pr\left(y_{ij}=1\right)} = \sum_{n=1}^{I} \beta_{1n} x_{1n,ij} + \sum_{m=1}^{J-1} \beta_{2m} x_{2m,ij} - \sum_{m=1}^{J-1} \beta_{2m} x_{2J,ij} \qquad (10)$$

So

$$\log \frac{\Pr\left(y_{ij}=1\right)}{1-\Pr\left(y_{ij}=1\right)} = \sum_{n=1}^{I} \beta_{1n} x_{1n,ij} + \sum_{m=1}^{J-1} \beta_{2m} \left(x_{2m,ij} - x_{2J,ij}\right) \qquad (11)$$

Let $z_{m,ij} = x_{2m,ij} - x_{2J,ij}$. As $x_{2m,ij} = 1$ if $m = j$ as noted before, $x_{2J,ij} = 1$ only if $j = J$. Thus $z_{m,ij} = 1$ if $m = j$; $z_{m,ij} = -1$ if $j = J$. The above equation can then be reformulated as follows:

$$\log \frac{\Pr\left(y_{ij}=1\right)}{1-\Pr\left(y_{ij}=1\right)} = \sum_{n=1}^{I} \beta_{1n} x_{1n,ij} + \sum_{m=1}^{J-1} \beta_{2m} z_{m,ij} \qquad (12)$$

where $x_{1n,ij} = 1$ if $n = i$, $z_{m,ij} = 1$ if $m = j$, and $z_{m,ij} = -1$ if $j = J$; otherwise they are 0. These coded z variables for items have values 1, 0, and -1. This is called effect coding and imposes a sum-to-zero constraint on the model coefficients (Rutherford, 2001). Equation (12) has I dummy variables for persons but $J - 1$ effect-coded variables for items. The coefficient or parameter of the last item can be obtained through the sum-to-zero constraint, i.e., equation (9).

First, after the last item parameter is obtained, the logistic regression model, shown by equation (12), has parameters for all persons and items which are in one-to-one correspondence with the parameters in a standard Rasch model. Second, the two models have the same ML score equations as discussed before. Third, by the effect-coded item variables, a sum-to-zero constraint is imposed on item parameters of equation (12). Therefore the Rasch model is equivalent to this special logistic regression model shown by equation (12) under ML. Their parameter estimates should be very similar except for the signs of item parameter estimates because β_{2j} corresponds to $-b_j$ as previously noted. If equation (12) is then specified in the logistic regression computer programs, e.g., SAS LOGISTIC procedure (SAS Institute, 2011a, 2011b) and SPSS LOGISTIC REGRESSION command (SPSS Inc., 2005), the results should be comparable with WINSTEPS's except for the signs of item parameters.

Reducing ML Bias

When ML is applied in item response theory (IRT) (Lord, 1980), it is also called the joint maximum likelihood (JML). "Joint" means the method estimates person and item parameters simultaneously (Drasgow, 1989). The method for the Rasch model proposed by Wright and collaborators (Wright & Douglas, 1977; Wright & Stone, 1979) can also be called JML. But JML is specially used for IRT models. It has some slight differences from the regular ML. For example, the computation algorithm in JML proposed by Wright and collaborators is slightly different from what is used in the regular ML fitting the logistic regression models. They used different initial values and approaches to satisfy the sum-to-zero constraint.

However, related research has shown that JML or ML estimates are inconsistent (Ghosh, 1995; Wright & Douglas, 1977), i.e., they are biased, and the biases cannot be eliminated when the sample size increases. Wright and Douglas (1977) provided a corrective approach, i.e., JML estimates are multiplied by $(L-1)/L$, where L is the smaller of the average person or item response count. However, the correction appears to contain puzzling assumptions and to rest on inadequate logic (Jansen, van den Wollenberg, & Wierda, 1988). Corrective approaches also generally require the existence of a finite estimate or they may reduce bias only in an asymptotic sense (Firth, 1993). Firth then suggested using PML to reduce ML bias preventively when fitting the logistic regression models. If the Rasch model can be estimated under the logistic regression framework, it can also reduce the ML bias of fitting the Rasch model.

PML has been earlier applied in IRT. Wang and Wang (2001) showed that a weighted likelihood estimation of person parameters proposed by Warm (1989) is a special case of Firth's PML. Kosmidis (2007) suggested that Firth's PML can be applied to the Rasch model and to the two-parameter logistic (2PL) IRT Model. But he did not provide the detailed derivation or proof, and a special computer program may also be required to implement this method. However, as mentioned, if the Rasch model is able to be fitted under the logistic regression framework, all existing applications of Firth's PML for the logistic regression can directly be used to the Rasch model with no extra derivation or proof.

ML bias can be reduced by introducing a small bias into the score function (Firth, 1993). Regularly, the ML estimate is derived as a solution to the score equation

$$U(\theta) = \nabla l(\theta) = 0 \tag{13}$$

where $l(\theta)$ is the log-likelihood function. A modified score function is then

$$U^*(\theta) = U(\theta) - i(\theta)b(\theta) \tag{14}$$

where $-i(\theta) = U'(\theta)$ is the local gradient and $b(\theta)$ is the bias. If θ is the canonical parameter of an exponential family model, a modified log-likelihood function is

$$l^*(\theta) = l(\theta) + \log|i(\theta)|^{1/2} \tag{15}$$

where $|i(\theta)|^{1/2}$ is the penalty function. Suppose there are s observations and t variables, y is the dependent variable, x the independent variable, and β the parameters in the logistic regression model. The usual score (gradient) equation

$$g(\beta_t) = \sum_s (y_s - \pi_s)x_{st} = 0 \tag{16}$$

is modified as

$$g(\beta_t) = \sum_s \left[y_s - \pi_s + h_s \left(\frac{1}{2} - \pi_s \right) x_{st} \right] = 0 \tag{17}$$

where h_s is the ith diagonal element of the hat matrix $\mathbf{W}^{1/2}\mathbf{X}(\mathbf{X'WX})^{-1}\mathbf{X'W}^{1/2}$ and $\mathbf{W} = \text{diag}[\pi_s(1 - \pi_s)]$ (Heinze & Schemper, 2002). The Hessian matrix is not modified by this penalty. The method is implemented in the SAS LOGISTIC procedure (SAS Institute, 2011b). In the logistic regression, estimated standard errors of PML estimators could still be obtained as the square roots of diagonal elements of the inverse of information matrix as regular ML does (Firth, 1993). Kosmidis (2007) suggested using the square roots of the inverse of diagonal elements of the information matrix for the Rasch or 2PL models. Furthermore, Kosmidis (2007) pointed out that estimated standard errors can be obtained directly by the value of the information matrix at the last iteration in ML, but it would yield an underestimation of the standard errors in PML; they should be obtained from a separate evaluation.

Simulation Study and Results

A simulation study was implemented as suggested by Harwell, Stone, Hsu, and Kirisci (1996). Simulated data were generated based on the standard Rasch model, shown in equation (1). Both difficulty parameter and person ability were generated from the standard normal distribution. Wright and Stone (1979) implied that 20 items and 200 examinees were enough to obtain adequate parameter estimates using the Rasch model. So the data were simulated to have item responses of 200, 500, and 1,000 simulees to 20-, 40-, and 60-item-long tests. These different sample sizes and test lengths were completely crossed and formed nine test conditions. Each condition had 1,000 replications. The simulated data were generated using SAS (SAS Institute, 2011a), and analyzed by WINSTEPS, the SAS LOGISTIC procedure using ML and PML (SAS-ML; SAS-PML), respectively. For the purpose of comparison, the Newton-Raphson method and the convergence criterion, 0.0001, were used in both the SAS LOGISTIC procedure and WINSTEPS.

When using the SAS LOGISTIC procedure to fit the Rasch model, the model shown by equation (12) needs to be specified in the procedure. It should be noted that the model has no intercept, and the event category '1' is fitted. The dependent variable is the scores of persons on test items, and its independent variables are the I dummy variables for persons and the $J - 1$ effect-coded variables for items. If a simulated data set has 1,000 simulees, then 1,000 dummy variables are needed to be specified. It takes a long time to estimate so many parameters. Fortunately, it is unnecessary to create so many dummy variables because the same parameter estimates are given to persons who take the same item set and receive the same total of scores in the Rasch model. Practically, the dummy variables for persons can be created for all observed total scores and item sets instead of all persons. Namely, dummy variables are created to differentiate persons by their total scores and assigned item sets together.

In this study, the accuracy of the three methods was evaluated using the root mean square error (RMSE) between item difficulty parameter estimates and their true values. The following comparisons were made.

- Item and person parameter estimates obtained from WINSTEPS, SAS-ML, and SAS-PML, respectively, compared with true values of item parameters simulated.
- Comparisons of parameter estimates between the three methods.

Table 1. Root mean square errors of item parameter estimates

Test Length	Sample Size	True value vs.			WINSTEPS vs.		SAS-ML vs. -PML
		WINSTEPS	SAS-ML	SAS-PML	SAS-ML	SAS-PML	
20	200	0.3191	0.3191	0.3145	0.0001	0.0141	0.0141
20	500	0.2964	0.2964	0.2949	0.0001	0.0055	0.0055
20	1000	0.2797	0.2797	0.2789	0.0001	0.0027	0.0028
40	200	0.2579	0.2579	0.2544	0.0001	0.0132	0.0133
40	500	0.2185	0.2185	0.2174	0.0001	0.0052	0.0053
40	1000	0.1983	0.1983	0.1978	0.0001	0.0026	0.0026
60	200	0.2305	0.2306	0.2275	0.0001	0.0128	0.0129
60	500	0.1879	0.1879	0.1869	0.0001	0.0052	0.0052
60	1000	0.1723	0.1723	0.1719	0.0001	0.0026	0.0026

Note: SAS-ML = SAS LOGISTIC procedure using maximum likelihood; SAS-PML = SAS LOGISTIC procedure using penalized maximum likelihood

Table 2. Root mean square errors of person parameter estimates

Test Length	Sample Size	N	True value vs.			WINSTEPS vs.		SAS-ML vs. -PML
			WINSTEPS	SAS-ML	SAS-PML	SAS-ML	SAS-PML	
20	200	199037	0.3191	0.3191	0.3145	0.0001	0.0141	0.0141
20	500	497532	0.2964	0.2964	0.2949	0.0001	0.0055	0.0055
20	1000	995265	0.2797	0.2797	0.2789	0.0001	0.0027	0.0028
40	200	199886	0.2579	0.2579	0.2544	0.0001	0.0132	0.0133
40	500	499732	0.2185	0.2185	0.2174	0.0001	0.0052	0.0053
40	1000	999534	0.1983	0.1983	0.1978	0.0001	0.0026	0.0026
60	200	199971	0.2305	0.2306	0.2275	0.0001	0.0128	0.0129
60	500	499953	0.1879	0.1879	0.1869	0.0001	0.0052	0.0052
60	1000	999886	0.1723	0.1723	0.1719	0.0001	0.0026	0.0026

Note: SAS-ML = SAS LOGISTIC procedure using maximum likelihood; SAS-PML = SAS LOGISTIC procedure using penalized maximum likelihood

Table 1 shows the RMSEs of item parameter estimates of the three methods. By this table, Firth's PML reduced ML bias because SAS-PML had the smallest RMSE under each condition. Among the three methods, estimates from WINSTEPS and SAS-ML were almost identical. RMSE between them was smaller than 0.0001.

Table 2 shows the RMSEs of person parameter estimates of the three methods after excluding simulees obtaining extreme (zero or perfect) scores because JML or ML cannot provide finite estimates to parameters of those persons. The results are similar to the ones in Table 1. SAS-PML still had the smallest RMSE; RMSEs between WINSTEPS and SAS-ML were still smaller than 0.0001.

Therefore this simulation study provides further evidence that the Rasch model can be fitted under the logistic regression framework using ML and the logistic regression software similar to the SAS LOGISTIC procedure, and Firth's PML reduced ML or JML biases of fitting the Rasch model. But the study found that some tiny differences existed between the estimates of WINSTEPS and SAS-ML, i.e., JML and ML. It may be a result of slight differences between JML in WINSTEPS and the regular ML in the SAS LOGISTIC procedure as mentioned earlier.

Discussion

The paper further showed that the standard Rasch model is equivalent to a logistic regression model specially specified under ML. At least their parameter estimates are equivalent under ML. The Rasch model can be fitted under the logistic regression framework using ML, and the ML estimates are comparable with what the Rasch software WINSTEPS gives using JML. But it is inappropriate to say that the Rasch model is a special case of logistic regression. It is because of the following:

- This study showed only that the Rasch model is equivalent to a special logistic regression model under ML or PML. The Rasch model can be fitted using other methods, e.g., the marginal maximum likelihood (Bock & Aitkin, 1981).
- They have different standard errors for both item and person parameter estimates. In the logistic regression, standard errors are calculated from the square root of diagonal elements of the inverse of an information matrix, but in the Rasch model, they are actually obtained from the square root of the inverse of diagonal elements of the information matrix. The algorithm of the Rasch model actually simplifies the logistic regression's.
- In the Rasch framework, every item parameter has its own standard error although one has none in the logistic regression model because it is calculated through the sum-to-zero constraint.
- The Rasch model provides the infit and outfit statistics for each item or person, but the logistic regression has no such fit statistics.

In contrast with ML, not only can PML reduce ML bias, but it can also generate a finite parameter estimate to an item or a person obtaining an extreme score. Heinze and Schemper (2002) have shown that Firth's method always yields finite estimates of parameters under complete or quasi-complete separation. Then

Firth's PML can directly estimate parameters of all items or persons together simultaneously and have no convergent problem whether they have extreme scores or not. Table 3 shows the RMSEs of WINSTEPS and SAS-PML estimating parameters of the simulees who received extreme scores.

WINSTEPS adjusts an extreme score and makes it a little less than perfect or a little more than zero because the parameter of a person with an extreme score is inestimable using ML. By default, the adjustment is 0.3 (Linacre, 2008). By Table 3, using the adjustment, WINSTEPS estimates for extreme scores had smaller RMSEs than SAS-PML when the test had 20 items; SAS-PML performed better when the test became longer. But it seems that the number of extreme scores influences the accuracy of SAS-PML when estimating parameters of persons with extreme scores. Table 3 shows that the more extreme scores appeared, the greater RMSE of SAS-PML became. But more evidence is needed to draw a final conclusion because the number of simulees with extreme scores was relatively small in the simulation. In other IRT software, e.g., SAS PROC IRT (SAS Institute, 2011a, 2011b) and IRTPRO (Cai, Thissen, & du Toit, 2011), expected a posteriori (EAP), and maximum a posteriori (MAP) can be used to estimate person parameters. PML may also be compared with EAP and MAP in future studies.

Table 3. Root mean square errors of parameter estimates for persons obtaining extreme scores

Test Length	Sample Size	N	True value vs.		WINSTEPS vs. SAS-PML
			WINSTEPS	SAS-PML	
20	200	963	2.5153	2.5492	0.4857
20	500	2468	2.5090	2.9582	0.6938
20	1000	4735	2.4865	3.3935	1.0764
40	200	114	2.5772	2.1255	0.4868
40	500	268	2.5731	2.1985	0.4562
40	1000	466	2.5305	2.2964	0.4406
60	200	29	2.5139	2.0188	0.5092
60	500	47	2.4846	2.0037	0.4935
60	1000	114	2.5407	2.1254	0.4697

Note: SAS-PML = SAS LOGISTIC procedure using penalized maximum likelihood

References

Agresti, A. (2002). *Categorical data analysis*. New York: Wiley-Interscience. doi: 10.1002/0471249688

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46*(4), 443-459. doi: 10.1007/bf02293801

Cai, L., Thissen, D., & du Toit, S. H. C. (2011). IRTPRO: Flexible, multidimensional, multiple categorical IRT modelling [Computer software]. Lincolnwood, IL: Scientific Software International.

De Boeck, P., & Wilson, M., (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York, NY: Springer. doi: 10.1007/978-1-4757-3990-9

Drasgow, F. (1989). An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model. *Applied Psychological Measurement, 13*(1), 77-90. doi: 10.1177/014662168901300108

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika, 80*(1), 27-38. doi: 10.1093/biomet/80.1.27

Ghosh, M. (1995). Inconsistent maximum likelihood for the Rasch model. Statistics & Probability Letters, 23(2), 165-170. doi: 10.1016/0167-7152(94)00109-l

Harwell, M., Stone, C. A., Hsu, T.-C., & Kirisci, L. (1996). Monte Carlo studies in item response theory. *Applied Psychological Measurement, 20*(2), 101-125. doi: 10.1177/014662169602000201

Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine, 21*(1), 2409-2419. doi: 10.1111/j.0006-341x.2001.00114.x

Jansen, P. G., van den Wollenberg, A. L., & Wierda, F. W. (1988). Correcting unconditional parameter estimates in the Rasch model for inconsistency. *Applied Psychological Measurement, 12*(3), 297-306. doi: 10.1177/014662168801200307

Kamata, A. (1998). *Some generalizations of the Rasch model: An application of the hierarchical generalized linear model* (Unpublished doctoral dissertation). Michigan State University, East Lansing, MI.

Kamata, A. (2001). Item analysis by the hierarchical generalized linear model. *Journal of Educational Measurement, 38*(1), 79-93. doi: 10.1111/j.1745-3984.2001.tb01117.x

Kosmidis, I. (2007). *Penalized likelihood for a non-linear Rasch model* (Unpublished manuscript). Centre for Research in Statistical Methodology, the University of Warwick, Coventry, UK.

Linacre, J. M. (2008). *A user's guide to WINSTEPS-MINISTEP: Rasch-model computer programs*. Chicago, IL: winsteps.com.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum. doi: 10.4324/9780203056615

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen, Denmark: Danmarks Pædagogiske Institut.

Rutherford, A. (2001). *Introducing ANOVA and ANCOVA: A GLM approach*. Thousand Oaks, CA: SAGE.

SAS Institute. (2011a). SAS/STAT® 9.3 [Computer program]. Cary, NC: SAS Institute.

SAS Institute. (2011b). *SAS/STAT® 9.3 user's guide*. Cary, NC: SAS Institute.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.

SPSS Inc. (2005). *SPSS 14.0 command syntax reference*. Chicago, IL: SPSS Inc.

Uekawa, K. (2005). *Why can we do Rasch model using HLM?* (Unpublished manuscript). Retrieved from http://www.estat.us/sas/RaschModelEssay.doc

Wang, S., & Wang, T. (2001). Precision of Warm's weighted likelihood estimation of ability for a polytomous mode in CAT. *Applied Psychological Measurement, 25*(4), 317-331. doi: 10.1177/01466210122032163

Warm, T. A. (1989). Weighted likelihood estimation of ability in the item response theory. *Psychometrika, 54*(3), 427-450. doi: 10.1007/bf02294627

Wright, B. D. (1993). Discrete-time survival analysis: A Rasch-model logistic regression. *Rasch Measurement Transactions, 7*(3), 307. Retrieved from https://www.rasch.org/rmt/rmt73g.htm

Wright, B. D., & Douglas, G. A. (1977). Best procedures for sample-free item analysis. *Applied Psychological Measurement, 1*(2), 281-294. doi: 10.1177/014662167700100216

Wright, B. D., & Panchapakesan, N. (1969). A procedure for sample-free item analysis. *Educational and Psychological Measurement, 29*(1), 23-48. doi: 10.1177/001316446902900102

Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago, IL: MESA Press.