THE DEVELOPMENT OF A ZERO-INFLATED RASCH MODEL

Sungyeun Kim [a,*] and Guemin Lee [b]

Abstract. The purpose of this study was to develop a zero-inflated Rasch (ZI-Rasch) model, a combination of the Rasch model and the ZIP model. The ZI-Rasch model was considered in this study as an appropriate alternative to the Rasch model for zero-inflated data. To investigate the relative appropriateness of the ZI-Rasch model, several analyses were conducted using PROC NLMIXED procedures in SAS under various simulation conditions. Sets of criteria for model evaluations (-2LL, AIC, AICC, and BIC) and parameter estimations (RMSE, and r) from the ZI-Rasch model were compared with those from the Rasch model. In the data-model fit indices, regardless of the simulation conditions, the ZI-Rasch model produced better fit statistics than did the Rasch model, even when the response data were generated from the Rasch model. In terms of item parameter λ estimations, the ZI-Rasch model produced estimates similar to those of the Rasch model.

1. Introduction

The zero-inflated Poisson (ZIP) model has been used in various fields, such as public health, medicine, and criminology, with many zeros in the data ([4], [8], [10], [15]). It is also not uncommon to encounter data with many zeros in education fields. The American Invitational Mathematics Examination (AIME) serves as an example as an illustration of zero-inflated data. The AIME is a three-hour examination with 15 questions given to those who rank in the top 5% at the high-school mathematics level. Because each answer is one integer number ranging from 0 to 999, guessing is not a factor influencing student test scores. One point is given to each correct answer, and no partial credit is considered. Thus, the AIME scores range from 0 to 15 overall. According to the results from 2002 to 2005, the mean scores ranged

from 2.195 to 2.741 and the median scores were 2 and 3 ([1], [2]). Many students failed to gain even a single point, hence the term zero-inflated data. In contrast, Mansell and Curtis [9] reported that that the "top private school dumps too easy the General Certificate of Secondary Education (GCSE)". In such a case, the number of incorrect answers at high-ability levels among test takers can be close to zero. Hence, zero-inflated data can arise if we count the number of incorrect answers.

The family of Rasch measurement models, sharing similar Poisson characteristics, is used in many fields despite the fact that these models do not provide a good fit for data with excess numbers of zeros. Ogasawara [11] termed this phenomenon as overdispersion or extra-Poisson variation. In fact, the equal mean and variance property in the Poisson model could be violated for data with excessive numbers of zeros. In the context of applying the Poisson model, overdispersion leads to underestimation of standard errors. Assuming the Poisson model instead of the ZIP model for data with excessive numbers of zeros can result in inconsistent parameter estimates ([13]).

The ZIP model became more commonly used after the publication of Lambert [7]. The ZIP model combines a distribution degenerating at zero with a baseline Poisson model. Lambert assumed that a population consisted of two groups. The first group occurs with probability ω and produces only zeros, while the second group occurs with $1 - \omega$ probability and leads to a Poisson-distributed count with a mean of λ. The purpose of the present study is to develop zero-inflated Rasch (ZI-Rasch) model, a combination of the Rasch model and the ZIP model, and to investigate the relative appropriateness of the ZI-Rasch model for data with excessive numbers of zeros.

2. ZI-Rasch Model and Parameter Estimation

In this study, we propose a new Rasch-family model termed the zero-inflated Rasch (ZI-Rasch) model to handle zero-inflated data. We suppose that a score that includes all of the categories of y_{ijk} of the i-th test taker ($i = 1, \cdots, N$) on the j-th item ($j = 1, \cdots, J$) composed of k categories ($k = 1, \cdots, J_k$) is a random variable distributed as the ZIP (τ) model having two distinct underlying states. The first state occurs with probability ω_{jk} and produces only zeros, whereas the other state occurs with probability $1 - \omega_{jk}$ and leads to a Poisson model with a mean of λ_{ijk}. In general, the zeros from the first state are known as true zeros or structural zeros (a zero score from test takers with low ability levels) while those from the Poisson

model are known as false zeros or sampling zeros (a zero score from test takers of medium or high ability levels). In addition, τ is a shape-scalar parameter related to the probability producing only zeros, $\omega_{jk} = (1 + \lambda_{ijk}{}^\tau)^{-1}$ ([7]). It is natural that the probability ω_{jk} under true zeros is in inverse proportion to the mean λ_{ijk} under the Poisson model. Although the vectors of the covariates may or may not be the same sets of variables, we use the ZIP (τ) model for two reasons in this study. First, all covariates such as items are the same. Second, it is possible to halve the number of parameters needed for the ZIP model, which may accelerate the computations considerably. From the model, the probability that test taker i has the score, a count of completing k categories, y_{ijk}, on item j is given by

$$(2.1) \qquad P(Y_{ijk} = y_{ijk}) = \begin{cases} \omega_{jk} + (1 - \omega_{jk})e^{-\lambda} & y_{ijk} = 0 \\ \\ (1 - \omega_{jk})\frac{e^{-\lambda}\lambda_{ijk}{}^y}{y_{ijk}!} & y_{ijk} = 1, 2, \cdots, 0 \leq \omega_{jk} \leq 1 \end{cases}$$

The link functions are $\log(\frac{\omega_{jk}}{1-\omega_{jk}}) = -\tau \mathbf{X}'\boldsymbol{\beta}$ and $\log(\lambda_{ijk}) = \mathbf{X}'\boldsymbol{\beta}$. The parameter ω_{jk} has a logit-link with covariates denoted as $\mathbf{X}$, and the mean parameter λ_{ijk} is log-linked to the same covariates $\mathbf{X}$, where $\mathbf{X}$ is an item-indicator dummy variable to indicate the item number and $\boldsymbol{\beta}$ is the associated regression parameter. It is also clear that this model is reduced to the Poisson model when $\omega_{jk} = 0$.

Following Newton's theory of gravitation ($F = MA$) from the laws of physics, we specify a simple multiplicative or additive structure to relate the test taker (T) to the category for the item (D) ([3], [12]). The relationship of the accelerations to the masses and the force, after logarithmic transformation, is $A_1 = F_{11} + M_1$ and $A_2 = F_{11} + M_2$. Then, by subtraction, $A_1 - A_2 = M_1 - M_2$, where the force F_{11} has been deleted. Hence, through a comparison of accelerations independently in which the force is contained, the masses can be compared.

Beginning with the multiplicative structure, we assume that

$$(2.2) \qquad\qquad\qquad \lambda_{ijk} = \frac{T_i}{D_{jk}}, \; T_i \geq 0, \; D_{jk} \geq 0.$$

Taking the logarithms on both sides of (2.2) produces the additive structure

$$(2.3) \qquad\qquad\qquad \lambda_{ijk} = \log T_i - \log D_{jk} = \theta_i - \delta_{jk}$$

where $\theta_i = \log T_i$ and $\delta_{jk} = \log D_{jk}$. Regarding achievement testing, the location of the i-th test taker, T_i or θ_i, expresses the ability, and the location of the k-th category in the j-th item, D_{jk} or δ_{jk}, expresses the mean category difficulty.

In this study, we consider the probability function ϕ, the probability of completing the k-th category, in a simple dichotomous model by substituting the item difficulty δ_j for the mean category difficulty δ_{jk}. On a dichotomously scored test, k takes a value of 1 if item j is answered correctly, and 0 otherwise. The probability of test taker i succeeding rather than failing on item j takes the following form

$$(2.4) \qquad \phi_{ij} = \frac{P(Y_{ij}=1)}{P(Y_{ij}=0)+P(Y_{ij}=1)} = \frac{\frac{\lambda_{ij}}{e^\lambda}}{\frac{1}{e^\lambda}+\frac{\lambda_{ij}}{e^\lambda}} = \frac{\lambda_{ij}}{1+\lambda_{ij}}.$$

Of course, when only two responses are possible,

$$(2.5) \qquad P(Y_{ij}=0)+P(Y_{ij}=1) = 1.$$

At this point, we complete the ZI-Rasch model incorporating more zeros than those permitted under the Poisson assumption. By substituting (2.4) for (2.1), this can be formulated as

$$(2.6) \qquad P(Y_{ij}=y_{ij}) = \begin{cases} \omega_j + (1-\omega_j)\frac{1}{e^{\theta_i-\delta_j}} & y_{ij}=0 \\[2mm] (1-\omega_j)\frac{e^{\theta_i-\delta_j}}{1+e^{\theta_i-\delta_j}} & y_{ij}=1, 0 \le \omega_{jk} \le 1 \end{cases}$$

where δ_j is the item j difficulty parameter and θ_i is the ability of test taker i. One link function is $\log(\frac{\omega_j}{1-\omega_j}) = -\tau \mathbf{X}'\boldsymbol{\delta}$, where τ is a shape-scalar parameter, $\mathbf{X}$ is the $J \times J$ item indicator matrix, and $\boldsymbol{\delta}$ denotes the $J \times 1$ item difficulty parameters. In addition, the other link function is $\log(\frac{\lambda_{ij}}{1-\lambda_{ij}}) = \boldsymbol{\theta} - \mathbf{X}'\boldsymbol{\delta} = \theta_i \mathbf{1} - \mathbf{X}'\boldsymbol{\delta}$, where $\boldsymbol{\theta}$ is the vector of the test takers' abilities and $\mathbf{1}$ is the $J \times 1$ vector having all J elements equal to one. In this model, the first state occurs with the probability of ω and produces only zeros, whereas the other state occurs with the probability of $(1-\omega)$ and proceeds to a Rasch model. It is also clear that this model reduces to the Rasch model when $\omega = 0$.

Parameter estimation for the family of Rasch measurement models and ZIP models typically involves maximum likelihood. In this study, we deal with the marginal maximum likelihood estimation (MMLE) of this model. The MMLE requires the maximization of the likelihood which requires integrating the joint probability function of the responses with respect to the test taker distributions. The log-likelihood

function $(LL(\omega, \delta, \theta))$ is

$$(2.7)\quad LL = \sum_{i=1}^{I}\sum_{j=1}^{J}[I_{(y_{ij}=0)}\log\{\frac{1}{1+e^{-\tau X_{ij}\delta_j}} + \frac{e^{-\tau X_{ij}\delta_j}}{1+e^{-\tau X_{ij}\delta_j}}(\frac{1}{1+e^{(\theta_i-X_{ij}\delta_j)}})\}]$$

$$+\ \sum_{i=1}^{I}\sum_{j=1}^{J}[I_{(y_{ij}=1)}\log\{\frac{e^{-\tau X_{ij}\delta_j}}{1+e^{-\tau X_{ij}\delta_j}}(\frac{e^{(\theta_i-X_{ij}\delta_j)}}{1+e^{(\theta_i-X_{ij}\delta_j)}})\}]$$

and the marginal likelihood is

$$(2.8)\quad \prod_{i=1}^{I}[\int_{\theta}\prod_{j=1}^{J}\{(\frac{e^{-\tau X_{ij}\delta_j}}{1+e^{-\tau X_{ij}\delta_j}})(\frac{e^{(\theta_i-X_{ij}\delta_j)}}{1+e^{(\theta_i-X_{ij}\delta_j)}})\}^{y_{ij}}$$

$$\times\{\frac{1}{e^{-\tau X_{ij}\delta_j}}(1+\frac{e^{(\theta_i-X_{ij}\delta_j)}}{1+e^{-\tau X_{ij}\delta_j}})\}^{1-y_{ij}}f(\theta)d\theta]$$

where θ denotes the random effects sampled from a normal distribution, denoted by the probability density function $f(\theta)$. With an expectation-maximization (EM) algorithm, it would be much simpler to maximize the log likelihood of (2.7) if the values of the test taker parameters τ, δ and θ were observed. However, they are not observed. Thus, an iterative EM algorithm is used, where the unobserved values of τ, δ, and θ are replaced by their conditional expectations and temporary estimates of (τ, δ, θ), for instance $(\tau_0, \delta_0, \theta_0)$, in the E-step. The M-step maximizes (2.8) and provides new estimates of (τ, δ, θ), in this case $(\tau_1, \delta_1, \theta_1)$, for the next E-step ([6]).

3. Methods

3.1. Simulation Procedures This study used the factors listed in the Table 1, types of which are also used in parameter recovery studies. Although these parameters were more precisely estimated for long tests and many test takers, this study used a few (500), a moderate number (1000), and a sufficient number of test takers (2000). In addition, when the ZI-Rasch model is used, zero-inflated data are needed. To obtain these data, this study assumed a very high level of difficulty relative to the ability level.

Table 1. *Factors and Conditions in the Simulation Study*

Factors	Conditions
Number of Test Takers	500, 1000, 2000
Number of Items	30, 50
Difficulty of Items	N(5, 1), N(5.5, 1), N(6, 1)
Estimation Model	Rasch Model, ZI-Rasch Model

Each condition in this study was replicated 4 times following a previous study in [10]. The relatively small number of replications may be due to the computing time in fitting the simulated dataset. Because the number of replications was small, all standard errors of estimates were reported.

3.2. Data In this study, we need zero-inflated data, which can be obtained by administering very difficult tests. We considered the beta and normal distributions as underlying distributions for the ability and item difficulty parameters, respectively. By manipulating the difference between the ability and item difficulty parameter distributions, we tried to obtain various data sets with varying proportions of zeros. For each condition, test takers' dichotomous responses were generated using the following procedures.

Step1. Test takers' ability parameters (θ_B) were sampled randomly from a beta distribution with alpha 10 and beta 2. The beta distribution is supported in the range from 0 to 1, and it is possible to alter the range. Then, the test takers' ability parameters (θ_Z) have a range of (-3, 3), as a standard normal distribution, when the function $\theta_Z = \theta_B \times (3 - (-3)) - 3$, is used. Hence, the mean and the standard deviation of the θ_Z distributions are approximately 2.06 and 0.62, respectively. When the beta distribution is used, researchers can control the shape and scale parameters and thus obtain various distributions, such as a strictly increasing, a strictly decreasing, a normal-shaped or a U-shaped distribution.

Step 2. Item difficulty parameters were assumed to follow a normal distribution with means of 5.0, 5.5, and 6.0 and a standard deviation of 1, as we had to consider large differences such as three standard deviations from the mean of the test takers' ability in order to obtain zero-inflated data.

Step 3. Using the given test taker ability and item difficulty parameters, the probability of a correct response of each test taker to each item was calculated using the Rasch model. Although we have much interest in zero-inflated data, the Rasch model was used to generate data responses rather than the ZI-Rasch model in order to evaluate the appropriateness of the ZI-Rasch model relative to the Rasch model from a more conservative perspective.

Step 4. A random number was generated from a uniform (0, 1) distribution. If the random number was less than or equal to the probability of a correct response, an item score of one was recorded. Otherwise, an item score of zero was recorded.

3.3. Evaluation Criteria This study determined how well the estimated model

fitted the observed data, where the fit was the degree of the difference between observed and predicted values based on log-likelihood approaches such as the likelihood ratio statistic ($-2LL$), Akaike's information criterion (AIC), the Consistent Akaike Information Criterion (AICC), and Schwarz's Bayesian information criterion (BIC). These indices penalize for model complexity and permit comparisons of non-nested models; however, they can only be used descriptively. In addition, these indices indicate better fits when they are small. To evaluate the extent to which the ZI-Rasch model could recover the item parameters precisely, concurrent parameter recovery studies were conducted. The precision of parameter estimation was evaluated in terms of the root mean square error (RMSE) and the correlations (r) between the estimated parameters and the actual value of the parameters. The proposed method provided better item recovery when these values produced small RMSE and high r results for the estimated and simulated parameters. Before computing these indices, the estimated parameters were transformed to the same scale as the true parameter using standardized values. The RMSE was the square root of the average squared difference between the estimated and true values of an item parameter across J items and the test taker parameter across I test takers. The RMSE was determined as follows:

$$(3.1) \qquad RMSE = \sqrt{\frac{\sum_{r=1}^{R} \sum_{j=1}^{J} (\hat{\delta}_j - \delta_j)^2}{J \times R}}$$

Here, $\hat{\delta}_j$ is the estimated value of the given item parameter from item j, δ_j denotes the true value of the given item parameter from item j, and R represents the number of iterations.

3.4. Analyses Using all of the procedures, simulated test data were generated using WinGen3 in [5], a computer program designed to generate item responses based on various item response models using simulated test taker ability and item difficulty parameters in [16]. To obtain data-model fit indices and parameter estimates, we used the PROC NLMIXED procedure in SAS software [14].

4. Results

4.1. Data-Model Fit The data-model fit was evaluated by computing the $-2LL$, AIC, AICC and BIC values. As shown in Table 2, all data-model fit indices of the ZI-Rasch model were smaller than were those of the Rasch model.

Table 2: *Model-Data Fit Indices between the Rasch Model and the ZI-Rasch Model*

Item Difficulty	No. of Items	No. of Test Takers	−2LL		AIC		AICC		BIC	
			Rasch	ZI-Rasch	Rasch	ZI-Rasch	Rasch	ZI-Rasch	Rasch	ZI-Rasch
N(5.0, 1)	30	500	7439	7339	7499	7401	7499	7401	7626	7531
		1000	14602	14451	14662	14513	14662	14513	14810	14665
		2000	29243	28939	29303	29001	29303	29001	29471	29175
	50	500	12510	12353	12610	12455	12610	12455	12821	12670
		1000	23609	23367	23709	23469	23709	23469	23954	23719
		2000	48419	47914	48519	48016	48519	48016	48799	48302
N(5.5, 1)	30	500	6733	6656	6793	6718	6793	6718	6919	6849
		1000	13039	12898	13099	12960	13099	12960	13246	13112
		2000	25939	25654	25999	25716	25999	25716	26167	25889
	50	500	9810	9701	9910	9803	9911	9803	10121	10018
		1000	19127	18898	19227	19000	19227	19000	19472	19250
		2000	38647	38247	38747	38349	38747	38349	39027	38635
N(6.0, 1)	30	500	4137	4121	4197	4183	4197	4183	4323	4314
		1000	8206	8143	8266	8205	8266	8205	8413	8358
		2000	16259	16125	16319	16187	16319	16187	16487	16360
	50	500	6318	6261	6418	6363	6418	6363	6628	6578
		1000	12321	12226	12421	12328	12421	12328	12666	12578
		2000	25379	25147	25479	25249	25479	25249	25759	25535

Note: N(μ, σ) indicates that the difficulty parameters were selected from a normal distribution with a mean of μ and a standard deviation of σ.

For example, the $-2LL$ values from the Rasch model with difficulty levels of N(5, 1) to N(6.0, 1) with 30 items and 500 test takers were 7439, 6733, and 4137. On the other hand, the values of -2LL from the ZI-Rasch model with the difficulty level ranging from N(5, 1) to N(6.0, 1) were 7339, 6656 and 4121.

Depending on the difficulty levels of N(5.0, 1), N(5.5, 1), and N(6.0, 1) with 30 and 50 items, the corresponding proportions of zeros were approximately 0.13 and 0.19, 0.42 and 0.08, and 0.14 and 0.30 for the sets. It was natural for the proportion of zeros to increase when the mean value of the normal distribution also increased, as the mean values denote the item difficulty levels. In addition, high difficulty levels, many items, and many test takers were generally associated with small fit values. Based on overall data-model fit indices, regardless of all simulation conditions, the ZI-Rasch model explained the data better than the Rasch model. Regarding all of the data-model fit indices, smaller was better. These results therefore confirm that the ZI-Rasch model is superior to the Rasch model in terms of the data-model indices, even when the response data came from the Rasch model.

4.2. Parameter Recovery To evaluate how well the item difficulty parameters recovered under varying simulation conditions, comparisons between the Rasch and the ZI-Rasch models were carried out using RMSE and r as the evaluation criteria. Table 3 summarizes the RMSE of the item difficulty parameter estimates.

Table 3: *The RMSE between the Simulated and Estimated Item Difficulty Parameters and Estimated Item Difficulty Parameters*

Item Difficulty	No. of Test Takers	No. of Item=30		No. of Item=50	
		Rasch	ZI-Rasch	Rasch	ZI-Rasch
	500	0.165	0.165	0.170	0.170
N(5.0, 1)	1000	0.132	0.135	0.139	0.139
	2000	0.087	0.091	0.106	0.105
	500	0.246	0.244	0.267	0.267
N(5.5, 1)	1000	0.208	0.206	0.180	0.176
	2000	0.167	0.174	0.153	0.150
	500	0.370	0.371	0.286	0.284
N(6.0, 1)	1000	0.245	0.245	0.212	0.211
	2000	0.188	0.187	0.198	0.192

Note: $N(\mu, \sigma)$ indicates that the difficulty parameters were selected from a normal distribution with a mean of μ and a standard deviation of σ.

In Table 3, the RMSE values are generally lower in the ZI-Rasch model than they are in the Rasch model. In contrast, RMSE values are generally lower in the Rasch

Table 4: *Correlations between Simulated and Estimated Item Difficulty Parameters*

Item Difficulty	No. of Test Takers	No. of Item=30		No. of Item=50	
		Rasch	ZI-Rasch	Rasch	ZI-Rasch
	500	0.981	0.981	0.975	0.975
N(5.0, 1)	1000	0.988	0.987	0.983	0.983
	2000	0.995	0.994	0.990	0.990
	500	0.968	0.968	0.965	0.965
N(5.5, 1)	1000	0.977	0.977	0.984	0.985
	2000	0.985	0.984	0.988	0.989
	500	0.939	0.940	0.940	0.941
N(6.0, 1)	1000	0.973	0.973	0.967	0.967
	2000	0.984	0.985	0.971	0.973

Note: $N(\mu, \sigma)$ indicates that the difficulty parameters were selected from a normal distribution with a mean of μ and a standard deviation of σ.

model than they are in the ZI-Rasch model with 30 items, except at a difficulty level of N(5.5, 1) with 500 and 1000 test takers and N(6.0, 1) with 2000 test takers. However, there was little difference in most cases. In addition, the RMSE values increased for both models when the difficulty level increased from N(5.0, 1) to N(6.0, 1) and decreased when the number of test takers increased from 500 to 2000. The number of items did not appear to affect the item difficulty parameter estimation for both the Rasch and the ZI-Rasch models. Moreover, large proportions of zeros in the data were associated with large RMSE values because the difficulty parameters of the ZI-Rasch model were generated from using the item distributions selected from the Rasch model. Considering small RMSE values, the ZI-Rasch model would likely yield more accurate estimates compared to the Rasch model when there are many test takers with zero-inflated data.

Table 4 above summarizes the average correlations between the true simulated and estimated item parameters. The correlation values are generally higher in the ZI-Rasch model than they are in the Rasch model. In contrast, the correlation values are usually higher in the Rasch model than in the ZI-Rasch model with 30 items, except at a difficulty level of N(6.0, 1) with 500 and 2000 test takers. However, the differences are insignificant in most cases. In addition, the correlation values showed a tendency to decrease in both models when the difficulty level increased from N(5.0, 1) to N(6.0, 1), and to increase when the number of test takers increased from 500 to 2000.

5. Conclusion

The goal of this study was to develop a ZI-Rasch model based on the Rasch and

ZIP models as an alternative when zero-inflated data exists. Hence, this study compared the Rasch model to the ZI-Rasch model using data-model fit indices and parameter estimates as evaluation criteria. Specifically, this study investigated effects of the number of items, the number of test takers, and the difficulty level of the items between these models. In the data-model fit indices, regardless of the simulation condition, the ZI-Rasch model produced better fitting values compared to the Rasch model, even when response data were generated from the Rasch model. In addition, the ZI-Rasch model generally produced item difficulty parameter estimates similar to those of the Rasch model.

The ZI-Rasch model tended to produce better fitting statistics compared to the Rasch model. The results of $-2LL$, AIC, AICC and BIC analyses showed that the ZI-Rasch model yielded smaller values than the Rasch model. This implies that the ZI-Rasch model is a viable alternative to the popular Rasch model with zero-inflated data.

The ZI-Rasch model tended to produce difficulty estimates similar to those of the Rasch model, even after the response data were generated from the Rasch model. In the item difficulty parameter estimations, the number of test takers and the difficulty level of items had an influence on the RMSE and the correlations in the Rasch and the ZI-Rasch models. Both the ZI-Rasch model and the Rasch model tended to produce better item difficulty estimates as the number of test takers increased and the difficulty level of the items decreased. In conclusion, the ZI-Rasch model developed in this study appears to be a better model for data with excessive numbers of zeros than popular Rasch model.

References

1. American Invitational Mathematics Examination: In Wikipedia, The Free Encyclopedia. Retrieved from http://en.wikipedia.org/w/index.php?title=American Invitational Mathematics Examination& oldid = 418555225, 2011.

2. T. Andreesscu & R. Gelca: *Mathematical Olympiad Challenges.* Birkhäuser Boston. Retrieved from http://www.amc.maa.org, 2010.

3. D. Andrich: *Rasch models for measuremen.* Newbury Park, CA: SAGE, 1988.

4. D. Fletcher, D.I. Mackenzie & E. Villouta: Modelling skewed data with many zeros: A simple approach combining ordinary and logistic regression. *Environ. Ecol. Stat.* **12** (2005), 45-54.

5. K.T. Han & R.K. Hambleton: User's manual for WinGen: Windows software that generates IRT model parameters and item responses. *Center for Educational Assessment Research Report* **642** (2007), 516-524.

6. M.S. Johnson: Marginal maximum likelihood estimation of item response models in R. *J. stat. soft.* **20** (2007), 1-25.

7. D. Lambert: Zero Inflated Poisson regression with an application to defects in manufacturing. *Technometrics.* **34** (1992), 1-14.

8. J.D. Lewsey & W.M. Thomson: The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socioeconomic status. *Community. Dent. Oral.* **32** (2004), 183-189.

9. W. Mansell & P. Curtis: *Top private school dumps too easy GCSEs.* The GuardianE, 2009.

10. T.G. Martin & B.A. Wintle, J.R. Rhodes, P.M. Kuhnert, S.A. Field, J. Low-Choy, A.J. Tyre & H.P. Possingham: Zero tolerance ecology improving ecological inference by modelling the source of zero observations. *Ecol. Lett.* **8** (2005), 1235-1246.

11. H. Ogasawara: Rasch's multiplicative Poisson model with covariates. *Psychometrika.* **61** (1996), 73-92.

12. G. Rasch: *Probabilistic models for some intelligence and attainment tests.* Copenhagen: Denmark Paedogogische Institut, 1960.

13. M. Ridout, C.G.B. Demetrio & J. Hind: *Models for count data with many zeros.* International Biometric Conference, Cape Town, 1998.

14. SAS Institute: *SAS/STAT user'guide(Version 8)*, 2000.

15. A.H. Welsh, R.B. Cunningham, C.F. Donnelly & D.B. Lindenmayer: Modeling the abundance of rare species: Statistical models for counts with extra zeros. *J. Ecol. Mode.* (1996), 297-308.

16. S. Yang: *A comparison of unidimensional and multidimensional Rasch models using parameter estimates and fit indices when assumption of unidimensionality is violated.* Unpublished doctoral dissertation, Ohio State University, 2007.

[a]Peabody College, Vanderbilt University, Tennessee 37203, USA
Email address: sungyeun.kim@vanderbilt.edu

[b]Department of Education, Yonsei University, Seoul 120-749, Republic of Korea
Email address: guemin@yonsei.ac.kr