Biometrical Journal **45** (2003) 7, 837–850

Incorporating Inter-item Correlations in Item Response Data Analysis

Xiaoming Sheng[1], Atanu Biswas[2], and K. C. Carrière[3]*

[1] DFPM-Health Research Center, University of Utah School of Medicine, Salt Lake City, UT 84108 USA
[2] Applied Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata, India 700 018
[3] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1

Abstract

This paper concerns with the analysis of item response data, which are usually measured on a rating scale and are therefore ordinal. These study items tended to be highly inter-correlated. Rasch models, which convert ordinal categorical scales into linear measurements, are widely used in ordinal data analysis. In this paper, we improve the current methodology in order to incorporate inter-item correlations. We have advocated the latent variable approach for this purpose, in combination with generalized estimating equations to estimate the Rasch model parameters. The data on a study of families of lung cancer patients demonstrate the utility of our methods.

Key words: Polychoric correlation; Latent variable; Rasch model; Generalized estimating equations; Relative efficiency.

1. Introduction

Consider the following situation, where I multiple-choice questions are given to N subjects. For question i $(i = 1, \dots, I)$ and subject j $(j = 1, \dots, N)$, we obtain a response X_{ij}, which takes a value k $(k = 0, \dots, m_i)$; there are $m_i + 1$ possible values (categories) for question i. Rasch models (Rasch, 1960), which convert ordinal categorical scales into linear measurements, are widely used in ordinal data analysis. The dichotomous Rasch model was designed to measure separate latent traits for each response category. Over the past four decades, several generalizations of the model have been developed. In the threshold approach, the dichotomous Rasch model is assumed to hold with a certain probability of passing the thresholds between two neighboring response categories (Andrich, 1978). The least restrictive of these generalizations is given by the partial credit model (Wright and Masters, 1982), often referred to as the polytomous Rasch model because of its generality. The probability that subject j ($j = 1, \ldots, N$) responds to category $k+1$ ($k = 0, \ldots, m_i$) for item i ($i = 1, \ldots, I$), i.e. $P(X_{ij} = k) = \pi_{ijk}$, is defined as

$$\pi_{ijk} = \frac{\exp \sum_{k'=0}^{k} (\beta_j - \delta_{ik'})}{\sum_{k=0}^{m_i} \exp \sum_{k'=0}^{k} (\beta_j - \delta_{ik'})} = \frac{\exp (k\beta_j - \sigma_{ik})}{\sum_{k=0}^{m_i} \exp (k\beta_j - \sigma_{ik})}, \tag{1}$$

where β_j is the ability parameter for subject j, δ_{ik} is the difficulty parameter of item i, and σ_{ik} is the cumulative threshold parameter defined as $\sigma_{ik} = \sum_{k'=0}^{k} \delta_{ik'}$. As an example, β_j can be seen as student's ability to solve the questions in a certain math test, while δ_{ik} describes how difficult these questions are.

We note that these questionnaire items are inter-correlated. For example, in a study using items that measure symptom distress on a scale, many of the items are closely correlated. If one complains about frequency of chest pains, he/she will most likely also complain about their intensity. However, in the current Rasch method, the item difficulty parameters are assumed conditionally independent given subject ability parameters. As shown in the literature (for example, LIANG and ZEGER, 1986), ignoring the presence of significant correlations can lead to serious bias in the study conclusions and a loss of efficiency.

In this paper, we develop a method of accounting for inter-item correlation (OLSSON, 1979), known as the polychoric correlation, in item response data. We formulate the polychoric correlation coefficient using the concept of latent variables, variables that are usually continuous, although unobservable, and that are the source of the apparent complexity of the data (MILLER et al. 1962). We use the generalized estimating equations approach (GEE, LIANG and ZEGER, 1986) to obtain consistent estimates of the parameters of the Rasch model.

2. Correlation for Ordinal Random Variables and the Rasch Models

When considering dependency among ordinal categorical data, OLSSON (1979) showed that the Pearson correlation leads to biased estimates. Instead, one should use the polychoric correlation, i.e. the correlation among the underlying latent random variables. This is different from the traditional latent trait model approaches (for example, ANDERSON, 1977, and FISCHER, 1983), where the latent trait is related to the person ability parameter $\boldsymbol{\beta}$, and the correlation is that for $\boldsymbol{\beta}$ as well.

POON and LEE (1987) developed the most general model for estimating the polychoric correlation. The full maximum likelihood estimators of the polychoric correlation coefficient and threshold parameters were obtained via the Fletcher-

Powell algorithm. A computationally more efficient approach, called "the partition maximum likelihood method," was also proposed by POON and LEE (1987). POON, LEE, and BENTLER (1990) used a pseudo maximum likelihood approach, which is computationally more efficient than the full maximum likelihood approach. RONNING and KUKUK (1996) compared the efficiency of the estimates from a joint likelihood against those of a conditional likelihood for measuring the association of two ordinal variables. The maximum likelihood estimates of the correlation and threshold parameters are consistent in both approaches; however, estimates from the conditional model are less efficient.

We now show that polytomous Rasch model (1) can also be derived from the concept of a latent variable approach, if the cut-off threshold parameters are chosen appropriately. Corresponding to the ordinal data X_{ij} and $m_i = m$ for all i, we assume the existence of a continuous random variable W_{ij} such that $X_{ij} = k$ if and only if $W_{ij} \in (c_{ij,k-1}, c_{ij,k}]$ where $k = 0, \ldots, m$ with $c_{ij,-1} = -\infty$ and $c_{ij,m} = \infty$. For $k = 0, \ldots, m-1$, let

$$c_{ij,k} = \Psi^{-1}\left(\frac{\sum_{k'=0}^{k} \exp\,(k'\beta_j - \sigma_{ik'})}{\sum_{k'=0}^{m} \exp\,(k'\beta_j - \sigma_{ik'})}\right) \tag{2}$$

where $\Psi^{-1}(\alpha)$ is the upper $100(1-\alpha)\%$ point of an assumed distribution. Immediately, we get

$$\pi_{ijk} = \Psi(c_{ij,k}) - \Psi(c_{ij,k-1}) = \frac{\exp\,(k\beta_j - \sigma_{ik})}{\sum_{k'=0}^{m} \exp\,(k'\beta_j - \sigma_{ik'})}\,.$$

This is the Rasch model (1) from a latent variables point of view. If we assume a normal distribution for W_{ij}, we take $c_{ij,k} = \Phi^{-1}(.)$ in (2), where $\Phi^{-1}(\alpha)$ is the upper $100(1-\alpha)\%$ point of the standard normal distribution.

We can extend this idea to a multi-item case, where there can be more than one item. Suppose $\boldsymbol{W_j} = (W_{1j}, \ldots, W_{Ij})'$ is the latent response vector corresponding to the jth individual. We assume that $\boldsymbol{W}_j$ follows a multivariate distribution with a joint c.d.f. $\boldsymbol{\Psi}_j$ such that each W_{ij} has the marginal c.d.f. Ψ_{ij}. Again, take $c_{ij,k}$ to be

$$c_{ij,k} = \Psi_{ij}^{-1}\left(\frac{\sum_{k'=0}^{k} \exp\,(k'\beta_j - \sigma_{ik'})}{\sum_{k'=0}^{m} \exp\,(k'\beta_j - \sigma_{ik'})}\right).$$

We see that the Rasch model (1) holds for each component X_{ij}. Moreover they now have dependency according to the specification of $\boldsymbol{\Psi}_j$.

Intuitively, we treat the correlation between any two items as if it is same for all individuals. Thus one immediate choice of the joint distribution is Φ, the I-dimensional multivariate normal distribution with mean $\mathbf{0}$ and a variance matrix with compound symmetry structure and an equal correlation ρ among items: $\mathbf{\Sigma} = (1-\rho)\boldsymbol{I} + \rho\mathbf{1}\mathbf{1}'$ where $\boldsymbol{I}$ is the I-dimensional identity matrix and $\mathbf{1}$ is a vector of 1's. Then, the joint probability that the jth individual has the response k_i on the ith item ($i = 1, \ldots, I$), is

$$\pi_{(1,\ldots,I),j}(k_1,\ldots,k_I) = \sum_{i_1,\ldots,i_I=(0,1)} (-1)^{i_1+\ldots+i_I} \times \Phi(c_{1j,k_1-i_1},\ldots,c_{Ij,k_I-i_I})\,.$$

For any two items ($i \neq i'$) within a multi-item framework, we have

$$\pi_{(i,i'),j}(k,k') = \Phi(c_{ij,k}, c_{i'j,k'}) - \Phi(c_{ij,k}, c_{i'j,k'-1}) - \Phi(c_{ij,k-1}, c_{i'j,k'}) + \Phi(c_{ij,k-1}, c_{i'j,k'-1})\,. \tag{3}$$

3. Estimation and Inference Method

The parameters to be estimated are $\boldsymbol{\beta}$ (consisting $Im-1$ distinct values of β_j's), $\boldsymbol{\delta}$ (Im item difficulty parameters), and ρ. Denote $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\delta}')'$ and $\boldsymbol{\theta}_\rho = (\boldsymbol{\theta}', \rho)'$.

3.1 *Estimation of the polychoric correlation coefficient*

Given any two correlated items $\boldsymbol{x}_i$ and $\boldsymbol{x}_{i'}$, $(i,i') \in \{1,\ldots,I\}$, the full likelihood for ρ given $\boldsymbol{\theta}$ is:

$$\mathcal{L}_\rho(\boldsymbol{\theta}) = \prod_j \prod_{k,k'} \left(\pi_{(i,i'),j}(k,k')\right)^{I(j,k,k')}$$

where $I(j,k,k')$ is an indicator function of whether the jth subject responds to categories k and k' to the respective two items, and $\pi_{(i,i'),j}(k,k')$ is given as in (3). Note that the likelihood function is not shown as a direct function of ρ, but through the expression of $\pi_{(i,i'),j}(k,k')$. Therefore, the log-likelihood can be obtained as

$$l_\rho(\boldsymbol{\theta}) = \log \mathcal{L}_\rho(\boldsymbol{\theta}) = \sum_j \sum_{k,k'} I(j,k,k') \log\left(\pi_{(i,i'),j}(k,k')\right),$$

and the estimating equation for ρ is given by

$$\frac{\partial l_\rho(\boldsymbol{\theta})}{\partial\rho} = \sum_j \sum_{k,k'} \frac{I(j,k,k')}{\pi_{(i,i'),j}(k,k')} \times \pi'_{(i,i'),j}(k,k') = 0\,, \tag{4}$$

where $\pi'_{(i,i'),j}(k,\ k') = [\phi(c_{ij,k},\ c_{i'j,k'}) - \phi(c_{ij,k},\ c_{i'j,k'-1}) - \phi(c_{ij,k-1},\ c_{i'j,k'}) + \phi(c_{ij,k-1},\ c_{i'j,k'-1})]$ and $\phi(.)$ is the density function of a bivariate standard normal distribution with correlation coefficient ρ.

We can plug in an initial estimator of $\boldsymbol{\theta}$ into (4) to get an estimate of ρ through an iterative Newton-Raphson method. The second-order derivative of the log-likelihood function with respect to ρ is:

$$\frac{\partial^2 l_\rho(\boldsymbol{\theta})}{\partial\rho^2} = \sum_j \sum_{k,k'} I(j,k,k') \times \left[\frac{\pi''_{(i,i'),j}(k,k')}{\pi_{(i,i'),j}(k,k')} - \left(\frac{\pi'_{(i,i'),j}(k,k')}{\pi_{(i,i'),j}(k,k')}\right)^2\right]$$

where $\pi''_{(i,i'),j}(k,\ k') = [\phi'(c_{ij,k},\ c_{i'j,k'}) - \phi'(c_{ij,k},\ c_{i'j,k'-1}) - \phi'(c_{ij,k-1},\ c_{i'j,k'}) + \phi'(c_{ij,k-1},\ c_{i'j,k'-1})]$ and $\phi'(x,y) = (\rho/(1-\rho^2) + (x-\rho y)\,(y-\rho x)/(1-\rho^2)^2) \times \phi(x,y)$.

Therefore, the iterative equation gives

$$\hat{\rho}^{(t)} = \hat{\rho}^{(t-1)} - \left[\frac{\partial^2 l_\rho(\boldsymbol{\theta})}{\partial\rho^2}\right]^{-1} \left[\frac{\partial l_\rho(\boldsymbol{\theta})}{\partial\rho}\right]\Bigg|_{\hat{\rho}^{(t-1)}}, \qquad t = 1,\ldots,T \tag{5}$$

at some T.

This is how the polychoric correlation coefficient $\rho_{i,i'}$ between any two items i, i' is estimated. Based on the estimators of $\rho_{i,i'}$ for all $(I(I-1)/2)$ item pairs, the common ρ for all I items can be estimated by taking the average of these $\hat{\rho}_{i,i'}$ as follows:

$$\hat{\rho} = \frac{2\sum_{i<i'} \hat{\rho}_{i,i'}}{I(I-1)}\,.$$

It follows from the general results of the maximum likelihood estimates that if the initial estimate of $\boldsymbol{\theta}$ is consistent given the true value of ρ, then $\hat{\rho}_{i,i'}$, and therefore $\hat{\rho}$ is consistent and asymptotically normal with mean ρ and variance estimated by

$$\widehat{\text{Var}}\,(\hat{\rho}) = \frac{-4\left(\frac{\partial^2 l_\rho(\boldsymbol{\theta})}{\partial\rho^2}\right)^{-1}\Bigg|_{\hat{\rho}^{(T)}}}{(I(I-1))^2}\,.$$

3.2 *Estimation of the Rasch parameters*

Since the correlation among the items is modelled implicitly through the polychoric correlation, it is difficult to write the likelihood function in the usual way. We use the GEE approach to obtain consistent estimates of the parameters of the Rasch model.

First, let us define some mathematical characteristics of the response variable. For $(i,i') \in \{1,\ldots,I\}$, define $\mu_{ij} = E(X_{ij}) = \sum_{k=0}^{m} k\pi_{ijk}$, $\sigma_{(i,i),j} = \text{Var}\,(X_{ij}) = \sum_{k=0}^{m} k^2\pi_{ijk} - \mu_{ij}^2$, and

$$\sigma_{(i,i'),j} = \text{Cov}\,(X_{ij}, X_{i'j}) = \sum_{k,k'=0}^{m} kk'\pi_{(i,i'),j}(k,k') - \mu_{ij}\mu_{i'j}\,.$$

The estimating equation for $\boldsymbol{\theta}$ is given as

$$\boldsymbol{U}(\boldsymbol{\theta}) = \sum_{j=1}^{N} \boldsymbol{D}_j' \boldsymbol{\Omega}_j^{-1} (\boldsymbol{x}_j - \boldsymbol{\mu}_j) = \boldsymbol{0}\,, \tag{6}$$

where $\boldsymbol{x}_j = (X_{1j}, \dots, X_{Ij})'$, $\boldsymbol{\mu}_j = E(\boldsymbol{x}_j)$, $\boldsymbol{\Omega}_j = (\sigma_{(i,i'),j})_{(I \times I)}$ is the variance-covariance matrix of $\boldsymbol{x}_j$, and $\boldsymbol{D}_j = \partial \boldsymbol{\mu}_j / \partial \boldsymbol{\theta}$ is an $I \times p$ matrix.

Let us re-write the estimating equation (6) in the following way:

$$\boldsymbol{U}(\boldsymbol{\theta}, \hat{\rho}(\boldsymbol{\theta})) = \sum \boldsymbol{U}_j(\boldsymbol{\theta}, \hat{\rho}(\boldsymbol{\theta})) = \boldsymbol{0}\,, \tag{7}$$

where $\hat{\rho}(\boldsymbol{\theta})$ is a consistent estimator given $\boldsymbol{\theta}$. As shown in ANDERSON (1973), when the test length is fixed, the estimates of the "structural" parameters $\boldsymbol{\delta}$ (NEYMAN SCOTT, 1948) are not necessarily consistent. According to HABERMAN (1977), a sufficient condition for consistency of the estimates is that the number of items also gets large fast enough so that $(\log N)/I \to 0$. Similarly, we obtain the following asymptotic properties of $\hat{\boldsymbol{\theta}}$, solutions to (7).

Theorem 1: *Under mild regularity conditions, when* $N(\log N)^2/I^2 \to 0$, *and that* $\hat{\rho}$ *is consistent for* ρ *given true parameter* $\boldsymbol{\theta}$, $N^{\frac{1}{2}}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ *is asymptotically normally distributed with covariance matrix given by*

$$\text{Var}\,(N^{\frac{1}{2}}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})) = \lim_{N \to \infty} \left(\frac{\sum_j \boldsymbol{D}_j' \boldsymbol{\Omega}_j^{-1} \boldsymbol{D}_j}{N} \right)^{-1}$$

as $N \to \infty$. □

We omit the proof, as it follows along the similar lines as those in LIANG and ZEGER (1986).

The solution to (6) may also be obtained iteratively:

$$\hat{\boldsymbol{\theta}}^{(t)} = \hat{\boldsymbol{\theta}}^{(t-1)} + \left[\sum_{j=1}^{N} \boldsymbol{D}_j' \boldsymbol{\Omega}_j^{-1} \boldsymbol{D}_j \right]^{-1} \left[\sum_{j=1}^{N} \boldsymbol{D}_j' \boldsymbol{\Omega}_j^{-1} (\boldsymbol{x}_j - \boldsymbol{\mu}_j) \right] \Bigg|_{\hat{\boldsymbol{\theta}}^{(t-1)}}\,,$$

$$t = 1, \dots, T \tag{8}$$

at some T.

3.3 *Two-step iteration method*

To estimate parameters, we start with some initial estimates obtainable using the Spearman's correlation coefficient for ordered categorical data for ρ and the estimates from the traditional Rasch model for $\boldsymbol{\theta}$. Given these initial values, we use (5) and (8) alternatively to get the estimates of $\boldsymbol{\theta}$ and ρ. That is, at each step of the iteration, we solve for (5) to obtain an estimator of ρ, and then given the updated value of ρ, solve for (8) to obtain an estimator of $\boldsymbol{\theta}$. We repeat this procedure until we reach convergence. Since, given any initial consistent estimators, the estimates obtained from (5) or (8) are consistent after the first step (LEHMANN and CASELLA, 1998), the final estimates from this two-step iteration method after the last step T are also consistent and asymptotically normal (OLSSON, 1979). This method also allows us to compute a consistent estimate of its asymptotic variance matrix. The fact that in each of the two steps a portion of the parameters $\boldsymbol{\theta}_\rho$ is replaced by its consistent estimator implies some loss of efficiency. However, this loss is minor and negligible (OLSSON, 1979).

Therefore, the estimate $\hat{\rho}_{i,i'}^{(T)}$ obtained from (5) after the last iteration at step T is asymptotically normal with mean ρ and variance $\operatorname{Var}(\hat{\rho}_{i,i'})$, which can be estimated by

$$\widehat{\operatorname{Var}}(\hat{\rho}_{i,i'}) = -\left(\frac{\partial^2 l_\rho(\boldsymbol{\theta})}{\partial \rho^2}\right)^{-1}\Bigg|_{\hat{\rho}_{i,i'}^{(T)}}$$

where the right-hand side is evaluated at $\hat{\rho}_{i,i'}^{(T)}$ after the last iteration step T. Also $\hat{\boldsymbol{\theta}}^{(T)}$ obtained from (8) after the last iteration at step T is asymptotically normal with mean $\boldsymbol{\theta}$ and variance matrix $\operatorname{Var}(\hat{\boldsymbol{\theta}})$, which can be estimated by

$$\widehat{\operatorname{Var}}(\hat{\boldsymbol{\theta}}) = \left(\sum_{j=1}^{N} \boldsymbol{D}_j' \boldsymbol{\Omega}_j^{-1} \boldsymbol{D}_j\right)^{-1}\Bigg|_{\hat{\boldsymbol{\theta}}^{(T)}}$$

when the right hand side is evaluated at $\hat{\boldsymbol{\theta}}^{(T)}$ after the last iteration step T.

4. Efficiency Considerations

We now show the performance of our strategy through a computer simulation. We replicated the simulation 1000 times to obtain the empirical distribution of the estimates. We generated the latent variable for $N = 100$ subjects for $I = 4$ items from a multivariate normal distribution. The correlation parameter ρ is chosen to be 0, 0.3, and 0.7 respectively. Then the ordinal response data are obtained according to the latent traits as generated, with a given distinct set of parameters. For the dichotomous case, we used

$$\boldsymbol{\beta}' = (0, 0.25, 0.5, 0.75, 1)$$

for the ability of persons parameters and

$$\boldsymbol{\delta} = \begin{pmatrix} 0 & -0.3 \\ 0 & -0.1 \\ 0 & 0.1 \\ 0 & 0.3 \end{pmatrix}$$

Table 1

Summary of estimates of ρ in the dichotomous ($I = 4$, $m = 1$, $N = 100$) and polytomous models ($I = 4$, $m = 2$, $N = 100$)

	Dichotomous			Polytomous		
ρ	$\bar{\rho}$	$\widehat{\mathrm{Var}}\,(\hat{\rho})$	$\widehat{\mathrm{MSE}}\,(\hat{\rho})$	$\bar{\rho}$	$\widehat{\mathrm{Var}}\,(\hat{\rho})$	$\widehat{\mathrm{MSE}}\,(\hat{\rho})$
0.0	0.0289	0.0096	0.0104	0.0203	0.0397	0.0402
0.3	0.2516	0.0178	0.0201	0.3242	0.0740	0.0746
0.7	0.7511	0.0301	0.0327	0.7309	0.1185	0.1195

Note: The entries show the average, variance, and MSE of the estimated $\hat{\rho}$ in 1000 simulations.

for the difficulty of items, while for the polytomous case, they are

$$\boldsymbol{\beta}' = (0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2)$$

and

$$\boldsymbol{\delta} = \begin{pmatrix} 0 & -0.2962 & 0.2185 \\ 0 & -0.4270 & 0.4717 \\ 0 & -0.3885 & 0.6263 \\ 0 & -0.3852 & 0.1803 \end{pmatrix}$$

with $\mathbf{1}'\boldsymbol{\delta}\mathbf{1} = 0$ for identifiability (see WRIGHT and MASTERS, 1982).

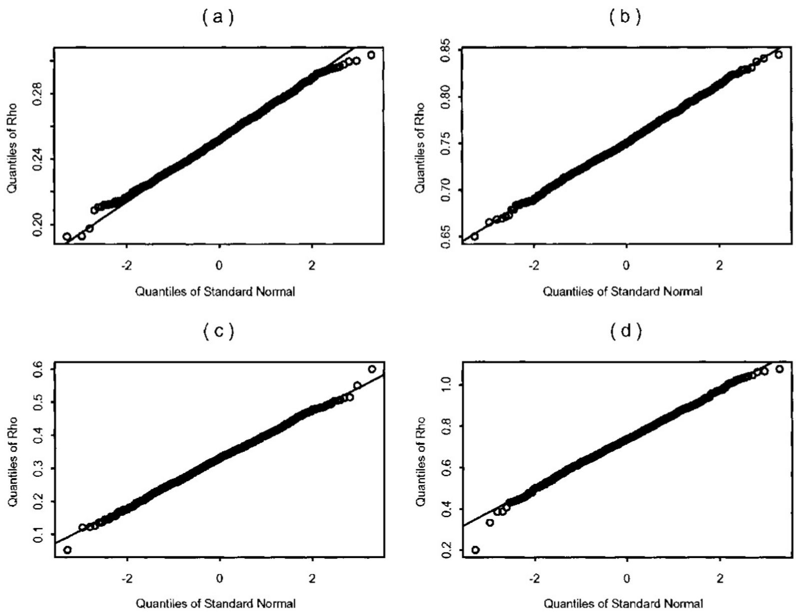

Fig. 1. Q-Q plots of $\hat{\rho}$ vs. standard normal quantiles
Note: Shown are standard normal quantiles against estimates of ρ for (a) $\rho = 0.3$, (b) $\rho = 0.7$ for the dichotomous model ($I = 4$, $m = 1$, $N = 100$), (c) $\rho = 0.3$, and (d) $\rho = 0.7$ for the polytomous model ($I = 4$, $m = 2$, $N = 100$)

Figure 1 gives the Q-Q plots of $\hat{\rho}$ when the true polychoric correlation coefficient is 0.3, and 0.7 respectively. We can see that the estimates closely follow a normal distribution in both cases. Table 1 reports the performance of our estimates of ρ in simulation. Both bias and variance increase with increasing values of ρ. The bias in estimation is significantly reduced in the polytomous model, but the variance is fourfold as compared to that in the dichotomous model.

Table 2 shows the relative efficiencies of the estimates for $\boldsymbol{\beta}$ with inter-item correlation over those assuming independence between items, based on both empirical and asymptotic distributions. When the true correlation is moderate, for example $\rho = 0.3$, there is little improvement in efficiency between the estimates obtained by recognizing the correlation among the items and those obtained under the independence assumption. The relative efficiencies are very close to 1. However, when the correlation is high, for example, $\rho = 0.7$, estimates based on recognizing the correlation gain a significant efficiency as compared to the traditional Rasch model, which assumes independence. The relative efficiencies could be as low as 0.73, and at most 0.84. The result of the estimation of $\boldsymbol{\delta}$ also gives a similar conclusion. Therefore, for the sake of brevity we have omitted the presentation of results on $\boldsymbol{\delta}$.

Table 2
Relative efficiencies of the estimators for $\boldsymbol{\beta}$ in the dichotomous ($I = 4$, $m = 1$, $N = 100$) and polytomous ($I = 4$, $m = 2$, $N = 100$) models

β	$\rho = 0$	$\rho = 0.3$		$\rho = 0.7$	
		Empirical	Asymptotic	Empirical	Asymptotic
Dichotomous					
0.25	1.000	.9763	.9814	.8250	.8403
0.50	1.000	.9847	.9896	.7502	.7338
0.75	1.000	.9755	.9783	.7958	.8058
overall	1.000	.9794	.9839	.7994	.7841
Polytomous					
0.25	1.000	.9459	.9245	.7455	.7393
0.50	1.000	.9635	.9566	.7196	.7234
0.75	1.000	.9545	.9630	.8259	.8404
1.00	1.000	.9293	.9302	.7403	.7582
1.25	1.000	.9777	.9694	.7993	.8037
1.50	1.000	.9329	.9445	.7587	.7459
1.75	1.000	.9798	.9632	.8149	.7904
overall	1.000	.9687	.9639	.7769	.7848

Note: The entries are relative efficiencies of the estimators for $\boldsymbol{\beta}$, based on both empirical and asymptotic distributions, and adjusted by the inter-item correlation compared to models assuming the independence of items. The row labelled 'overall' corresponds to those of the trace of the variance matrix. Ratios of less than 1 indicate improvement produced by taking inter-item correlation into consideration. The first half are for the dichotomous model and the second half for the polytomous model.

Table 3
Significant coefficients of reduced regression models of family care satisfaction scores in the lung cancer study

	Average Score	Traditional Rasch	Rasch With Correlation
intercept	1.402(0.312)	1.303(0.175)	1.397(0.154)
Family member:			
Age			
51–65	**−0.346(0.106)	***−0.210(0.058)	*−0.149(0.058)
65+	−0.026(0.098)	−0.041(0.061)	−0.048(0.059)
Ethnicity			
European	N/S	**0.112(0.048)	*0.101(0.047)
FAD	**−0.346(0.101)	***−0.254(0.051)	***−0.268(0.052)
DIS	−0.114(0.074)	***−0.140(0.020)	***−0.121(0.019)
DIS2	***0.033(0.009)	***0.035(0.004)	***0.033(0.004)
Patient:			
Age			
51–65	N/S	**0.273(0.105)	N/S
65+	N/S	0.070(0.103)	N/S
Education			
HighSch	N/S	−0.106(0.078)	−0.045(0.078)
College	N/S	*0.149(0.065)	*0.166(0.066)
SDS	**0.211(0.071)	*0.109(0.042)	*0.099(0.042)

Note: Shown are the estimates (standard errors) of coefficients with $p < 0.1$ from regression models using the simple average scoring method, the traditional Rasch method, and the simultaneous Rasch analysis with polychoric correlation; the level of significance is as indicated (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$; N/S: non-significant).

5. Numerical Example: Family Health of Lung Cancer Patients

In this section, we illustrate the methods described in the previous sections by using data from a study of lung cancer patients (KRISTJANSON et al., 1997). This study was undertaken to examine family care characteristics (family care expectations, perceptions, care satisfaction) and family health status across the illness trajectory. The objective was to understand the level of "family care satisfaction" in association with other family health and care measures.

The study involved 117 patients. Information on each patient and his/her family members was collected at the time of entry to a health care facility. Other family health and care variables were measured at various times and on as many as twelve occasions by means of questionnaire.

Demographic information was collected and following family health variables assessment tools were used: SDS (symptom distress scale with 13 items; 1 – normal, 2 – occasional distress, 3 – frequent distress, 4 – usual distress, 5 – constant distress), FAMCAR (family care satisfaction with 20 items; 1 = very satisfied, 2 = satisfied, 3 = undecided, 4 = dissatisfied, 5 = very dissatisfied), SOS (symptom of stress scale with 94 items; 0 = never, 1 = infrequently, 2 = sometimes, 3 = often, 4 = very frequently), and FAD (family assessment device with 12 items; 1 = strongly agree, 2 = agree, 3 = disagree, 4 = strongly disagree). The family care variables include: FEXP (family expectations scale with 16 items, ranging from 0 = not at all important to me to 10 = very important to me) and FPER (family perceptions scale with 21 items; 1 = strongly agree, 2 = agree, 3 = uncertain, 4 = disagree, 5 = strongly disagree).

For each of the above family health and care variables measured by a rating scale, we used the two-step iteration method (5) and (8) for the Rasch model to obtain its polychoric correlation and an "ability of persons" parameter for each family member adjusted for the polychoric correlation among items.

The polychoric correlations are quite high to be ignorable, with over .85 among the family expectation items, nearly .5 among the family assessment device items, and about .3 for other items.

The "ability of persons" parameter can alternatively be obtained via the popular average scoring approach taking a simple summary score for each family member on a number of given items and the traditional Rasch model approach using the CON procedure (Wright and Masters, 1982), which assumes no dependency among items or parameters. Our goal here is to compare the three approaches: (1) the popular average scoring approach; (2) the traditional Rasch method; (3) the proposed Rasch method incorporating polychoric correlation, an improvement over Sheng and Carrière (2002). Traditionally, these scores are then subjected to regression analyses to build a model, describing relationship between a score of a dependent variable and scores for a set of regressors.

We now demonstrate the varying degree of conclusions possible from different scoring methods in a regression model. The dependent variable is the family care satisfaction level (FAMCAR), predictors are the demographic variables and other family health and care variables, all estimated from each of the three approaches described above. A new variable is created, named "discrepancy" (DIS) that describes the discrepancy between the family expectations scale and the family perceptions scale. Here, we note that the explanatory variables are estimated with error from the Rasch model and that it may be appropriate to consider a measurement error model instead. However, for the purpose of comparisons, we focus on the performance of these scores in the traditional multiple regression model in this paper.

Table 3 displays the ordinary least squares estimates of all significant coefficients ($p < 0.05$) in the final model. The following variables are common in all three models: the family member's age, the family assessment device level, the symptom distress scale, the discrepancy score, and the squared discrepancy score. When the family members are between 51 and 65 years old and have a poor family environment as measured by FAD, they are less satisfied with health care. When the patients develop less symptom distress, the family care satisfaction level tends to be higher.

In contrast, the use of Rasch scores revealed a few more covariates as significant, in addition to the above. The traditional Rasch method finds the patient's age and education level and the family member's ethnic background to be significantly related to the family care satisfaction level. The family care satisfaction level is positively associated with patients who are middle-aged, from a European family background, and have an education level beyond high school. On the other hand, the Rasch method incorporating the polychoric correlation finds the patient's education and the family member's ethnic background to have significant relationships with the family care satisfaction level. The family care satisfaction level is positively associated with patients who had an education beyond high school and a European family background.

In summary, the directions and the levels of association between family health and care variables remained consistent among different scoring methods. However, the significance levels and the key variables identified were not the same, as discussed above. We conclude that the family's ethnic background matters in determining family care satisfaction levels. Europeans tend to be happier with family health care, and patients who have had some college education seem to manage better, and these factors lead to generally positive attitudes among family members about patient care. These additional significant factors provide important information about managing better patient and family care. The analysis was done using an Splus program. The program is available to readers upon request.

6. Concluding Remarks

We have proposed a latent variable approach to the Rasch model to incorporate inter-item correlations. Currently, there are methods available to achieve this purpose, such as the multivariate probit model. However, such models do not address the properties of fundamental measurement. Using the generalized estimating equations method, we developed an estimating method for the Rasch model parameters under item-to-item correlations. The simulation study has shown the relative efficiency of the proposed estimation method. Generally, the efficiency gain of the estimates increased, as the level of polychoric correlation becomes high.

As expected, choice of an estimation method makes a difference in how we interpret the data. In the example we illustrated to describe the relationships between family care satisfaction and other family health and care measures, we found that the significance levels and the key variables identified depended on the approaches chosen.

Overall, using the Rasch methods produced more precise results than the simple average scoring method. Clearly, the current approaches have some limitations. They do not fulfill the properties of the fundamental measurement (the average scoring method) or they assume conditional independence of difficulty parameters given the ability parameters (the traditional Rasch method). The method based on

GEE, incorporating the polychoric correlation rectifies all these limitations. An alternate approach may be to first formulate the linear latent model and estimate person's parameters. Although it would not appear to lead to any more simplified procedure, it is worth pursuing as a future work.

One limitation of the proposed estimation method is that it is computationally quite intensive. It may take a long time for the iteration procedure to converge, and sometimes it even diverges. Better numerical methods are needed to increase the efficiency of estimates. Further, a more general correlation pattern between the items might be desirable. Sufficiency and conditional likelihood inference in this approach are not included in this paper. They might be done in a similar way as that in ANDERSEN (1977).

Acknowledgement

This work was funded in part by a grant from the Natural Sciences and Engineering Research Council of Canada and the Alberta Heritage Foundation for Medical Research (AHFMR). The third author is a Senior Scholar with the AHFMR.

References

ANDRICH, D., 1978: Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, **2**, 581–594.

ANDERSON, E. B., 1973: *Conditional Inference and Models for Measuring*. Copenhagen: Mentalhygiejnisk Forlag.

ANDERSEN, E. B., 1977: Sufficient Statistics and latent trait models. *Psychometrika*, **42**, 69–81.

CRAMÉR, H., 1946: *Mathematical Methods of Statistics*. Princeton, NJ: Princeton University Press.

FISCHER, G. H., 1983: Logistic Latent Trait models with linear Constraints. *Psychometrika*, **48**, 3–26.

HABERMAN, S. J., 1977: Maximum Likelihood Estimates in Exponential Response Models. *Annals of Statistics*, **5**, 815–841.

KRISTJANSON, L., LEIS, A., KOOP, P., CARRIÈRE, K. C., and MUELLER, B., 1997: Family members' care expectations, care perceptions and satisfaction with advanced cancer care: a pilot study. *Journal of Palliative Care*, **13**, 5–13.

LEHMANN, E. L., and CASELLA, G., 1998: *Theory of Point Estimation*. New York, NY: Springer-Verlag.

LIANG, K. Y. and ZEGER, S. L., 1986: Longitudinal data analysis using generalized linear models. *Biometrika*, **73**, 13–22.

MILLER, C. R., SABAGH, G., and DINGMAN, H. F., 1962: Latent class analysis and differential mortality, *Journal of the American Statistical Association*, **57**, 430–438.

NEYMAN, J. and SCOTT, E. L., 1948: Consistent Estimates Based on Partially Consistent Observations. *Econometrica*, **16**, 1–32.

OLSSON, U., 1979: Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, **44**, 443–460.

POON, W. Y. and LEE, S. Y., 1987: Maximum likelihood estimation of multivariate polyserial and polychoric correlation coefficient. *Psychometrika*, **52**, 409–430.

POON, W. Y., LEE, S. Y., and BENTLER, P. M., 1990: Pseudo maximum likelihood estimation of multivariate polychoric and polyserial correlations. In *Proceedings of the Annual Meeting of the American Statistical Association*, 291–296.

Rasch, G., 1960: *Probabilistic models for some intelligence and attainment tests*. [Danish Institute of Educational Research 1960, University of Chicago Press 1980, MESA Press 1993] Chicago, IL: MESA Press.

Ronning, G. and Kukuk, M., 1996: Efficient estimation of ordered probit models. *Journal of American Statistical Association,* **91**, 1120–1129.

Sheng, X. and Carrière, K. C., 2002: An Improved CML Estimation Procedure for the Rasch Model with Item Response Data. *Statistics in Medicine*, **21**, 407–416.

Wright, B. D. and Masters, G. N., 1982: *Rating Scale Analysis: Rasch Measurement*. Chicago, IL: MESA Press.