



# Marginal and conditional approaches to longitudinal Rasch models

Maja Olsbjerg, Karl Christensen

## ► To cite this version:

Maja Olsbjerg, Karl Christensen. Marginal and conditional approaches to longitudinal Rasch models. Annales de l'ISUP, 2013, 57 (1-2), pp.109-126. hal-03611601

**HAL Id: hal-03611601**

**<https://hal.science/hal-03611601>**

Submitted on 17 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*Pub. Inst. Stat. Univ. Paris*  
 57, fasc. 1-2, 2013, 109-126

# Marginal and conditional approaches to longitudinal Rasch models

Maja Olsbjerg and Karl Bang Christensen

Department of Biostatistics, University of Copenhagen

**Abstract:** This paper describes different Rasch models for longitudinal data. An overview of marginal maximum likelihood (MML) estimation and conditional maximum likelihood (CML) estimation is presented. The assumptions implied in these models are discussed and two important aspects: invariance of item parameters over time and local independence across time that are identified. Several tests of these assumptions are proposed in the MML and the CML inference frame.

## 1. Introduction

Studying Patient Reported Outcomes (PRO's) often involve latent variables and Item Response Theory (IRT) models [1] are often used for modeling. Many applications involving PRO's are in longitudinal settings, where patients are followed over time. This calls for the use of longitudinal latent variable models yielding complex data structures where within-subject dependency should be taken into account in the statistical modeling. While solutions to address the issue of longitudinal data [2] have been proposed and evaluated, further investigation is warranted when latent variables are involved. Modeling of ordinal data with random effects has a long tradition in item response theory and some of the earliest descriptions of ordinal random effects regression models are found in the psychometric literature [3]. Latent variable models for multivariate longitudinal ordinal responses have been studied by many authors [4, 5]. The models include correlated error terms to accommodate longitudinal data and impose a structure on the latent variables. Another way of accounting for the dependence across time points arising from using the same question is by fitting item-specific random effects to account for correlations between time points. This is an extension of testlet theory [6] to longitudinal designs. Both of these approaches rely on distributional assumptions about the latent variables that are not easily tested. One way of avoiding this problem is to take advantage of Rasch models [7, 8] where conditional estimation can be carried out without problematic distributional assumptions. Ways of testing the assumptions of longitudinal IRT models and an investigation of the possibilities of using conditional rather than marginal approaches to longitudinal IRT will be discussed.

## 2. The Rasch model

The dichotomous Rasch model [7, 8] is the simplest IRT model. It describes the responses to manifest items  $X_1, \dots, X_I$  measuring a latent variable  $\theta$ . In the

M. Olsbjerg : [mola@sund.ku.dk](mailto:mola@sund.ku.dk) - University of copenhagen

K.B.Christensen : [kach@sund.ku.dk](mailto:kach@sund.ku.dk) - University of copenhagen

Rasch measurement model some important requirements should be met:

- (i)  $\theta \in \mathbb{R}$ .
- (ii)  $\theta \mapsto E(X_i|\theta)$  is increasing for all items  $i$ .
- (iii)  $Pr(X_1 = x_1, \dots, X_I = x_I | \theta) = \prod_{i=1}^I Pr(X_i = x_i | \theta)$  for all  $\theta \in \mathbb{R}$ .
- (iv)  $Pr(X_i = x_i | Y, \theta) = P(X_i = x_i | \theta)$  for all items  $i$  and all exogenous variables  $Y$ .

These are referred to as (i) unidimensionality, (ii) monotonicity, (iii) local independence and (iv) no differential item functioning (DIF). The response probability for item  $i$  for a given  $\theta$  is modeled as

$$Pr(X_i = x_i | \theta) = \frac{\exp(x_i(\theta - \beta_i))}{1 + \exp(\theta - \beta_i)} \quad (i = 1, \dots, I) \quad (2.1)$$

where  $\bar{\beta} = (\beta_i)_{i=1, \dots, I}$  is a vector of scalars, often called *item locations*, that describe the items and  $\theta$  is a scalar describing the person responding. It is clear that this way of modeling the probabilities automatically implies (i), (ii) and (iv).

For a sample of persons  $v = 1, \dots, N$  we assume that there is one value  $\theta_v$ , for each person and observe a response vector  $\bar{X}_v = (X_{vi})_{i=1, \dots, I}$ . By (2.1) and the assumption of local independence we obtain

$$Pr(\bar{X}_v = \bar{x}_v | \theta_v) = \prod_{i=1}^I Pr(X_{vi} = x_{vi} | \theta_v) \quad (2.2)$$

$$= \frac{\exp(\theta_v x_{v.} - \sum_i \beta_i x_{vi})}{\prod_{i=1}^I (1 + \exp(\theta_v - \beta_i))} \quad (2.3)$$

and see that  $x_{v.} = \sum_{i=1}^I x_{vi}$  is sufficient for  $\theta_v$ . The entire data matrix is denoted by  $\bar{\bar{X}} = (X_{vi})$ .

### 2.1. Identifiability

From (2.1) it is clear that the model is not identified since for any  $\kappa > 0$  the parameters  $\theta^* = \theta + \kappa$  and  $\beta_i^* = \beta_i + \kappa$  lead to the same probabilities. Restrictions are thus needed, e.g. by fixing item or person means. Typically the restriction  $\sum_{i=1}^I \beta_i = 0$  is used, but the choice of restriction does not change the model.

## 2.2. Item parameter estimation

Estimating item parameters jointly with the person parameters  $\theta_1, \dots, \theta_N$  using the likelihood function

$$\begin{aligned} L(\bar{\beta}, \theta_1, \dots, \theta_N) &= \prod_{v=1}^N Pr(\bar{X}_v = \bar{x}_v | \theta_v) \\ &= \frac{\exp(\sum_v \theta_v x_{v.} - \sum_i \beta_i x_{i.})}{\prod_{v=1}^N \prod_{i=1}^I (1 + \exp(\theta_v - \beta_i))} \end{aligned}$$

leads to inconsistent estimators [9] and for this reason marginal or conditional estimation is used.

### 2.2.1. Marginal maximum likelihood (MML) estimation

Assuming that the latent variable  $\theta$  is normally distributed with mean  $\mu$  and variance  $\sigma^2$ , the marginal probability of a response vector can be written as an integral  $Pr(\bar{X}_v = \cdot) = \int Pr(\bar{X}_v = \cdot | \theta) \varphi(\frac{\theta - \mu}{\sigma}) d\theta$ , where  $\varphi$  is the standard normal density. This yields the *marginal likelihood function*

$$L_M(\bar{\beta}, \mu, \sigma) = \prod_{v=1}^N \int Pr(\bar{X}_v = \bar{x}_v | \theta) \varphi\left(\frac{\theta - \mu}{\sigma}\right) d\theta. \quad (2.4)$$

Restrictions are also needed here: if item parameters are restricted  $\mu$  and  $\sigma$  can be estimated. Alternatively  $\mu$  can be set to zero so that  $\sigma$  and all item parameters can be estimated.

The likelihood equations derived from (2.4) can not be solved explicitly and estimation calls for numerical methods. More specifically, both integral approximation techniques and optimization algorithms are needed.

Several choices have been described extensively in the literature. Integral approximations of the Adaptive Quadrature type, with different choices of interpolating functions, are widely used [10], [11]. An example is the Gauss-Hermite Quadrature method which is the default in the NLMIXED procedure in SAS. Other options are Monte Carlo methods which comes in various different versions. In general these methods are particularly useful for higher dimensional integrals.

For optimization of the likelihood function a common choice is the Newton-Raphson algorithm which is also the default in NLMIXED.

If the distribution of  $\theta$  is correctly specified ML estimates of item parameters are consistent [13]. If the distribution is not normal then marginal ML estimates are not consistent. Justification of the distributional assumptions of the latent variable is a recurrent issue in random effects models and other distributions than the normal could be considered.



This model can be extended by imposing a linear structure

$$\begin{bmatrix} \theta_1 \\ \vdots \\ \theta_N \end{bmatrix} = \bar{Y}_\theta \bar{\eta}_\theta + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_N \end{bmatrix} \quad (2.5)$$

given by a design matrix  $\bar{Y}_\theta$  with one row for each person and a vector  $[\epsilon_1 \dots \epsilon_N]^T \sim MVN(0, \sigma^2 I_N)$  of residuals. This extension allows for the mean of the latent distribution to vary between different person groups. Models with this structure are often called *latent regression models* and have been discussed by many authors starting with [3].

### 2.2.2. Conditional maximum likelihood (CML) estimation

The sum  $R_v = X_v = \sum_{i=1}^I X_{vi}$  of item responses is sufficient for  $\theta_v$  the possible values being  $r = 0, 1, \dots, I$ . The distribution of the stochastic variable  $R_v$  is given by the probabilities

$$Pr(R_v = r | \theta_v) = \sum_{\bar{x} \in A_r} Pr(\bar{X}_v = \bar{x} | \theta_v)$$

computed by summation over the set  $A_r = \{\bar{x} : \sum_i x_i = r\}$  of all response vectors with sum  $r$ . Inserting (2.2) yields

$$Pr(R_v = r | \theta_v) = \frac{\gamma_r \exp(r\theta_v)}{\prod_{i=1}^I (1 + \exp(\theta_v - \beta_i))} \quad (2.6)$$

where

$$\gamma_r = \gamma_r(\beta_1, \dots, \beta_I) = \sum_{\bar{x} \in A_r} \exp(-\sum_{i=1}^I \beta_i x_i) \quad (2.7)$$

are symmetrical polynomials. Again summation is over all vectors  $\bar{x}$  with sum  $r$ . Due to the sufficiency item parameters can be estimated using the conditional distribution of the responses given the scores. The conditional probability of the vector of item responses given the sum  $r_v = x_v$  is

$$Pr(\bar{X}_v = \bar{x}_v | R_v = r_v) = \frac{\exp(-\sum_{i=1}^I \beta_i x_{vi})}{\gamma_{r_v}} \quad (2.8)$$

that does not depend on  $\theta_v$ . The *conditional likelihood function* is the product  $L_C(\bar{\beta}) = \prod_{v=1}^N Pr(\bar{X}_v = \bar{x}_v | X_v = x_v)$  that can be written

$$L_C(\bar{\beta}) = \frac{\exp(-\sum_{i=1}^I \beta_i x_{.i})}{\prod_{r=1}^I \gamma_r^{N_r}} \quad (2.9)$$

where  $x_{.i} = \sum_{v=1}^N x_{vi}$  are the item margins and  $N_r = \#\{v | x_v = r\}$  denotes the number of persons with score  $r = 1, \dots, I$ . Maximizing this likelihood yields

item parameter estimates which are conditionally consistent [14]. Restrictions are needed in order to ensure that the model is identified. Latent regression models imposing a structure like (2.5) have also been discussed in the conditional inference frame [15].

### 2.3. Model control

It is essential that the model fits the data in order for the results to be credible and a number of tests of the Rasch model have been proposed [16].

### 3. Longitudinal design

Now a longitudinal design is considered. Let  $X_{vit}$  denote the response given by person  $v$  to item  $i$  at time  $t$  and let the vector  $\bar{\theta}_v = (\theta_{vt})_{t=1,\dots,T}$  contain the values of the latent variable for person  $v$  across the time points  $t = 1, \dots, T$ . Let  $\beta_{it}$  denote the location of item  $i$  at time  $t$  and let  $\bar{\beta} = (\beta_{it})_{i=1,\dots,I; t=1,\dots,T}$ . The discussion of identifiability applies for the longitudinal case as well: if person parameters *and* item parameters are allowed to vary freely over time the model will not be identified and it will be impossible to distinguish changes in person locations from changes in item locations. This is intuitively clear as it is impossible to measure change with a measurement instrument that changes. The situation  $T = 2$  is by far the most frequently described in the literature.

### 4. Linear logistic test model

The *linear logistic test model* [17] is one of the earliest extensions of the dichotomous Rasch model to longitudinal designs. In the original formulation the model considers the same items observed for the same individuals at two different time points and imposes the model structure

$$Pr(X_{vit} = x_{vit} | \theta_v) = \frac{\exp(x_{vit}(\theta_v - \beta_{it}))}{1 + \exp(\theta_v - \beta_{it})} \quad (4.1)$$

thus specifying constant person locations and a change in item locations across time. If the item locations change by the same amount across items

$$\beta_{i2} = \beta_{i1} + \delta \quad \text{for all } i = 1, \dots, I \quad (4.2)$$

then this change can be interpreted as an average change in person locations. Conditional maximum likelihood estimation was originally proposed for the linear logistic test model, but marginal estimation is also feasible. It should be noted that the linear logistic test model does not model change at the individual person level, but rather at the item level. Specifying the linear structure (4.2) on the item parameters makes it possible to estimate the common change  $\delta$  that can be interpreted as change in the average person location. This has

since become standard in IRT. When each item is the result of a systematic combination of two or more factors like different item content, different situations, different time points or different modes of administration (e.g. telephone interview versus pen and paper questionnaire) a linear structure

$$\begin{bmatrix} \beta_{11} & \beta_{12} \\ \vdots & \vdots \\ \beta_{I1} & \beta_{I2} \end{bmatrix} = \bar{Y}_\beta \bar{\eta}_\beta \quad (4.3)$$

given by a design matrix  $\bar{Y}_\beta$  and a vector  $\bar{\eta}_\beta$  of parameters is imposed. This is called *facets* models [18]. Specifying a change in the average person location  $\beta_{i2} = \beta_{i1} + \delta$  can be done choosing

$$\bar{Y}_\beta = \begin{bmatrix} 1 & 0 & \dots & 0 & 1 \\ 0 & 1 & \dots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 1 \end{bmatrix} \quad \text{and} \quad \bar{\eta}_\beta = \begin{bmatrix} \beta_1 & \beta_1 \\ \vdots & \vdots \\ \beta_I & \beta_I \\ 0 & \delta \end{bmatrix}.$$

## 5. The multivariate Rasch model

For the longitudinal design a general model for the response  $X_{vit}$  of person  $v$  to item  $i$  at time  $t$  is given by the probabilities

$$Pr(X_{vit} = x_{vit} | \theta_{vt}) = \frac{\exp(x_{vit}(\theta_{vt} - \beta_{it}))}{1 + \exp(\theta_{vt} - \beta_{it})}. \quad (5.1)$$

As mentioned previously the model will not be identified if both person parameters and item parameters are allowed to vary freely over time and it will be impossible to distinguish changes in person locations from changes in item locations. The model (5.1) is thus too general and we will consider special cases. Andersen's longitudinal Rasch model [19] assumes time-invariant item parameters

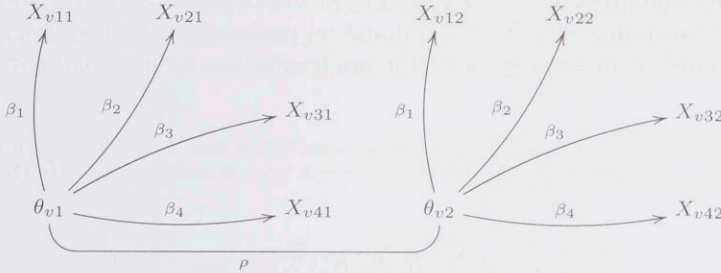
$$Pr(X_{vit} = x_{vit} | \bar{\theta}_v) = \frac{\exp(x_{vit}(\theta_{vt} - \beta_i))}{1 + \exp(\theta_{vt} - \beta_i)}. \quad (5.2)$$

Andersen used this model for individuals who respond to the same items at two time points, combining the values of the latent variable at the two time points in a two-dimensional latent density and estimating the correlation coefficient. This quantifies the change over time at the population level, but it is, however, important to notice that (5.2) is a model for change at individual level. The conditional independence assumptions and the parametric structure of the model is illustrated in Figure 1.

By imposing structure similar to (2.5) on the person parameters group level effects, rather than individual change can be estimated.



FIG 1. Graphical representation of the conditional independence assumptions and the parametric structure (time invariant item parameters) of the Andersen model (5.2).



### 5.1. Estimation

Conditional as well as marginal maximum likelihood estimation is feasible in the model (5.1). The general likelihood for an incomplete design can be written using indicator functions  $A_{vit}$  taking the value one if person  $v$  has responded to item  $i$  at time  $t$  and the value zero otherwise

$$\begin{aligned} L(\bar{\beta}, \bar{\theta}) &= \prod_{v=1}^N \prod_{t=1}^T \prod_{i=1}^I \left[ \frac{\exp(x_{vit}(\theta_{vt} - \beta_{it}))}{1 + \exp(\theta_{vt} - \beta_{it})} \right]^{A_{vit}} \\ &= \exp\left(\sum_{v=1}^N \sum_{t=1}^T R_{vt} \theta_{vt} - \sum_{i=1}^I \sum_{t=1}^T C_{it} \beta_{it}\right) K(\bar{\beta}, \bar{\theta})^{-1} \end{aligned} \quad (5.3)$$

where  $\bar{\beta}$  is the matrix of item parameters,  $\bar{\theta} = (\theta_{vt})_{v=1, \dots, N; t=1, \dots, T}$  is the matrix of person parameters,  $R_{vt} = \sum_{i=1}^I A_{vit} x_{vit}$  are person scores,  $C_{it} = \sum_{v=1}^N A_{vit} x_{vit}$  are item margins and

$$K(\bar{\beta}, \bar{\theta}) = \prod_{v=1}^N \prod_{t=1}^T \prod_{i=1}^I (1 + \exp(\theta_{vt} - \beta_{it}))^{A_{vit}}$$

a normalizing constant. If an item  $i$  is not administered to any persons at a time point  $t$  the corresponding parameter  $\beta_{it}$  can not be estimated. The model can be extended by imposing structure on the matrix

$$\bar{\theta} = \begin{bmatrix} \theta_{11} & \dots & \theta_{1T} \\ \vdots & & \vdots \\ \theta_{N1} & \dots & \theta_{NT} \end{bmatrix}.$$

Imposing structure on the matrix  $\bar{\beta}$  in a manner similar to (4.3) is also feasible.



### 5.1.1. Marginal maximum likelihood (MML) estimation

Assuming that the latent vector  $\bar{\theta} = (\theta_1, \dots, \theta_T)^T$  is multivariate normal with mean vector  $\bar{\mu}$  and covariance matrix  $\Sigma$ , the marginal probability of a response vector can be computed by integration using the *multivariate marginal likelihood function*

$$L_M(\bar{\beta}, \bar{\mu}, \Sigma) = \exp\left(-\sum_{i=1}^I \sum_{t=1}^T C_{it} \beta_{it}\right) \times \left[ \prod_{v=1}^N \int \exp\left(\sum_{t=1}^T R_{vt} \theta_t\right) K_v(\bar{\beta}, \bar{\theta})^{-1} \varphi_{\bar{\mu}, \Sigma}(\bar{\theta}) d\bar{\theta} \right] \quad (5.4)$$

where  $K_v(\bar{\beta}, \bar{\theta}) = \prod_{t=1}^T \prod_{i=1}^I (1 + \exp(\theta_t - \beta_{it}))^{A_{vit}}$  and  $\varphi_{\bar{\mu}, \Sigma}$  is the density of the multivariate normal distribution.

Again we need to make use of numerical methods for integral approximation and optimization of the likelihood [10], [11].

### 5.1.2. Conditional maximum likelihood (CML) estimation

From (5.3) it is clear that when  $A_{vit} = 1$  persons scores  $R_{vt} = x_{v,t}$  are sufficient for the person parameters and conditioning them out yields a multidimensional conditional likelihood function

$$L_C(\bar{\beta}) = \frac{\exp\left(-\sum_{t=1}^T \sum_{i=1}^I \beta_{it} x_{it}\right)}{\prod_{r=1}^I [\gamma_r^{(t)}]^{N_{rt}}} \quad (5.5)$$

$x_{it} = \sum_{v=1}^N x_{vit}$  are the item margins at time  $t$ ,  $\gamma_r^{(t)} = \gamma_r(\beta_{1t}, \dots, \beta_{It})$  and  $N_{rt} = \#\{v | x_{v,t} = r\}$  is the number of persons with score  $r$  at time  $t$ .

## 5.2. Identifiability

For each time point restrictions like those discussed in earlier sections are needed, but requiring invariance of item parameters across time points is also necessary. In the general formulation (5.1) the item set is not assumed to be constant across time as some items can be missing for all subjects at a given time point.

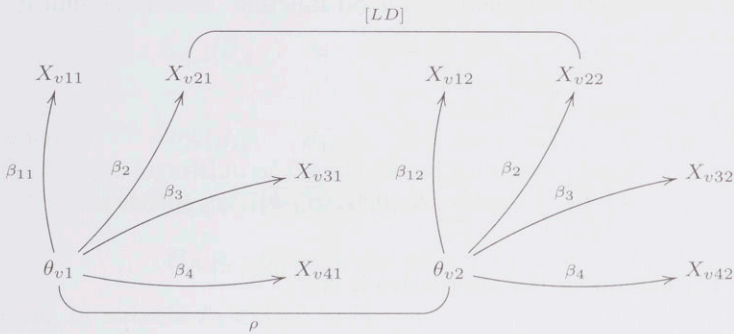
### 5.3. Model control

It is essential that the model fits the data in order for the results to be credible. Testing the fit of the Rasch model at each time point can be done using a wide range of tests proposed [16]. However, in the multivariate Rasch model additional tests of fit should be carried out in order to check the assumptions of invariant item parameters and local independence across time points as illustrated in Figure 1. Henceforth, assume that  $A_{vit} = 1$  for all persons  $v$ , items  $i$  and time points  $t$ .

## 6. Testing assumptions of the longitudinal Rasch model

In this section tests of two assumptions are discussed: invariance of the item parameter over time and local independence across time points. Violations of these assumptions are illustrated in Figure 2.

FIG 2. Violation of two model assumptions in (5.2): item parameters are not time invariant ( $\beta_{11} \neq \beta_{12}$ ) and there is local dependence across time points [LD].



### 6.1. Testing invariance of item parameters over time

As mentioned previously the model (5.3) has too many parameters in the sense that changes in person locations cannot be distinguished from changes in item locations. However, the assumptions of invariant item parameters and invariant person locations are not interchangeable. The response to a few items may have changed across time, but if responses to remaining items are essentially unchanged, the conclusion could be that changes are due to a change in the item locations rather than in the location of the persons. We consider the null hypothesis of time invariant item parameters

$$H_0 : \beta_{1t} = \beta_1, \dots, \beta_{It} = \beta_I \text{ for all } t = 1, \dots, T$$

and test it against a specified alternative

$$H_{A,a} : \beta_{it} = \beta_i \text{ for all } i \neq a \text{ and all } t = 1, \dots, T$$

in the MML and the CML inference frame. Let

$$\bar{\bar{\beta}}_0 = \begin{bmatrix} \beta_1 & \dots & \beta_1 \\ \vdots & & \vdots \\ \beta_I & \dots & \beta_I \end{bmatrix} \text{ and } \bar{\bar{\beta}}_{A,a} = \begin{bmatrix} \beta_1 & \dots & \beta_1 \\ \vdots & & \vdots \\ \beta_{a1} & \dots & \beta_{aT} \\ \vdots & & \vdots \\ \beta_I & \dots & \beta_I \end{bmatrix}.$$

For the CML inference frame tests against a general alternative are also considered.

### 6.1.1. Testing invariance of item parameters over time in the MML inference frame

In the MML inference frame the null hypothesis of time invariant item parameters can be tested against a specified alternative  $H_{A,a}$  where  $a$  represent one item for which invariance is violated. This can be done using likelihood ratio tests based on the multivariate marginal likelihood function. Under the null it becomes

$$L_0(\bar{\beta}_0, \bar{\mu}, \Sigma) = \exp\left(-\sum_{i=1}^I C_{i.}\beta_i\right) \times \left[\prod_{v=1}^N \int \exp\left(\sum_{t=1}^T R_{vt}\theta_t\right) K(\bar{\beta}_0, \bar{\theta})^{-1} \varphi_{\bar{\mu}, \Sigma}(\bar{\theta}) d\bar{\theta}\right]$$

where  $C_{i.} = \sum_{t=1}^T C_{it}$  and under the alternative it is

$$L_{A,a}(\bar{\beta}_{A,a}, \bar{\mu}, \Sigma) = \exp\left(-\sum_{i \neq a} C_{i.}\beta_i - \sum_{t=1}^T C_{at}\beta_{at}\right) \times \left[\prod_{v=1}^N \int \exp\left(\sum_{t=1}^T R_{vt}\theta_t\right) K(\bar{\beta}_{A,a}, \bar{\theta})^{-1} \varphi_{\bar{\mu}, \Sigma}(\bar{\theta}) d\bar{\theta}\right].$$

The above expressions looks quite complicated but are in fact just parametric multinomial distributions for which estimation is well-described. To realize this remember that the likelihood is simply a product of marginal probabilities of the observed  $I \times T$  response matrices. Hence, under the null it can be rewritten

$$L_0(\bar{\beta}_0, \bar{\mu}, \Sigma) = \prod_{v=1}^N Pr(\bar{X}_v = \bar{x}_v; \bar{\beta}_0) = \prod_{\bar{x} \in \{0,1\}^{I \times T}} Pr(\bar{X} = \bar{x}; \bar{\beta}_0)^{M_{\bar{x}}}$$

where the second product is taken over possible response matrices and  $M_{\bar{x}} = \#\{v : \bar{x}_v = \bar{x}\}$  is the number of persons with response matrix  $\bar{x}$ . This is proportional to the density of the vector  $(M_{\bar{x}})_{\bar{x} \in \{0,1\}^{I \times T}}$  which has a parametric multinomial distribution with probabilities depending on the parameters  $\bar{\beta}_0$ ,  $\bar{\mu}$  and  $\Sigma$ . Likewise the likelihood under the alternative has a parametric multinomial with parameters  $\bar{\beta}_{A,a}$ ,  $\bar{\mu}$  and  $\Sigma$ . See also [12].

Since these are nested models the logarithm of the likelihood ratio test statistic is  $\chi^2_{T-1}$ -distributed, see [12] section 6e.3. This can be extended to alternatives

$$H_{A,I_A} : \beta_{it} = \beta_i \text{ for all } i \notin I_A$$

given by subsets  $I_A \subset \{1, \dots, I\}$ .



### 6.1.2. Testing invariance of item parameters over time in the CML inference frame I

Under the null (5.5) is

$$L_0(\bar{\beta}_0) = \frac{\exp(-\sum_{i=1}^I \beta_i x_{i.})}{\prod_{r=1}^I [\gamma_r]^{N_r}}$$

where  $x_{i.} = \sum_{t=1}^T x_{it}$  and  $N_r = \sum_{t=1}^T N_{rt}$ , and under the alternative it is

$$L_{A,a}(\bar{\beta}_{A,a}) = \frac{\exp(-\sum_{i \neq a} \beta_i x_{i.} - \sum_{t=1}^T \beta_{at} x_{.at})}{\prod_{t=1}^T \prod_{r=1}^I [\gamma_r^{(t,a)}]^{N_{rt}}}$$

where  $\gamma_r^{(t,a)} = \gamma_r(\beta_1, \dots, \beta_{at}, \dots, \beta_I)$ , for  $t = 1, \dots, T$ . Since these are nested models the logarithm of the likelihood ratio test statistics are  $\chi_{T-1}^2$ -distributed [20]. This can be extended to alternatives

$$H_{A,I_A} : \beta_{it} = \beta_i \text{ for all } i \notin I_A \text{ and all } t = 1, \dots, T$$

given by subsets  $I_A \subset \{1, \dots, I\}$ .

### 6.1.3. Testing invariance of item parameters over time in the CML inference frame II: CML test based on crossing

A conditional likelihood ratio test of  $H_0 : \beta_{it} = \beta_i$  against the alternative  $H_{A,\{1, \dots, I\}}$  where  $\beta_{it} \neq \beta_{it'}$  for all  $i$  and  $t \neq t'$  is illustrated for  $T = 2$ . The test can be carried out by (i) randomly dividing respondents into two equally sized groups  $G = 1, 2$ , (ii) crossing over in such a way that for half the respondents we consider item responses from time 1 and for the other half item responses from time 2. In this way two data sets consisting of responses from two different time points are obtained (cf. Figure 3) and since there is only one observation per individual in each data set the potential problems of local dependence across time is overcome.

FIG 3. Illustration of the crossing utilized in the conditional test of item parameter invariance. The two data matrices in panel (ii) are used for the conditional likelihood ratio test.

(i)

$G$	$v$	$X_{11}$	$\dots$	$X_{I1}$
1	.	:		:
:	:	:		:
1	.	:		:
2	.	:		:
:	:	:		:
2	.	:		:

(ii)

$G$	$v$	$X_{11}$	$\dots$	$X_{I1}$	$X_{12}$	$\dots$	$X_{I2}$
1	.	:		:	*		*
:	:	:		:	*		*
1	.	:		:	*		*
2	.	:		:	*		*
:	:	:		:	*		*
2	.	:		:	*		*

Using conditional likelihood ratio tests to test for DIF with respect to  $G$  in each of the two new data sets is a way of testing whether the items work in the

same way at the two timepoints. In other words testing whether item parameters are time invariant.

Let the random variable  $G$  with  $P(G = 1) = P(G = 2) = \frac{1}{2}$  represent the allocation  $g(v) \in \{1, 2\}$  of persons  $v = 1, \dots, N$  into two groups. The likelihood under  $H_0$  and  $H_{A, \{1, \dots, I\}}$  is derived for the data set where responses for person  $v$  are those observed at time  $g(v)$ . Derivations for the other data set are identical. Thus, for  $v = 1, \dots, N$  and  $i = 1, \dots, I$  we consider response vectors  $\bar{X}_v^{(g)} = (X_{vig(v)})_{i=1, \dots, I}$  with response probabilities given by

$$P(X_{vig(v)} = x_{vig(v)} | \theta_v) = \frac{\exp(x_{vig(v)}(\theta_v - \beta_i))}{1 + \exp(\theta_v - \beta_i)}. \quad (6.1)$$

Conditioning on the scores  $R_v^{(g)} = \sum_{i=1}^I X_{vig(v)}$  we obtain the conditional probabilities that do not depend on  $\theta_v$  and the conditional likelihood under the null is

$$\begin{aligned} L_0(\bar{\beta}) &= \frac{\exp(-\sum_{v=1}^N \sum_{i=1}^I \beta_i x_{vig(v)})}{\prod_{v=1}^N \gamma_{r_v^{(g)}}} \\ &= \frac{\exp(-\sum_{i=1}^I \beta_i C_i^{(g)})}{\prod_{r=1}^I \gamma_r^{N_r^{(g)}}} \end{aligned} \quad (6.2)$$

where  $C_i^{(g)} = \sum_{v=1}^N x_{vig(v)}$  and  $N_r^{(g)} = \#\{v | \sum_{i=1}^I x_{vig(v)} = r\}$  are item margins and observed score distribution in the data set resulting from crossing. This likelihood is similar to (2.9). Under the alternative define scores  $r_v^{(1)} = \sum_{i=1}^I x_{vi1}$  for respondents in the first group, that is  $v \in g^{-1}(1)$  and scores  $r_v^{(2)} = \sum_{i=1}^I x_{vi2}$  for respondents in the second group  $v \in g^{-1}(2)$ . Furthermore define item margins  $C_i^{(1)} = \sum_{v \in g^{-1}(1)} x_{vi1}$  and  $C_i^{(2)} = \sum_{v \in g^{-1}(2)} x_{vi2}$ . The likelihood can be written

$$L_A(\bar{\beta}) = \frac{\exp(-\sum_{i=1}^I C_i^{(1)} \beta_{i1})}{\prod_{v \in g^{-1}(1)} \gamma_{r_v^{(1)}}} \frac{\exp(-\sum_{i=1}^I C_i^{(2)} \beta_{i2})}{\prod_{v \in g^{-1}(2)} \gamma_{r_v^{(2)}}}. \quad (6.3)$$

This is also a likelihood similar to (2.9) and since (6.3) and (6.2) are nested  $H_0$  can be tested against the alternative and the logarithm of the likelihood ratio test statistic is  $\chi_{I-1}^2$ -distributed [20].

This can be extended to the situation  $T > 2$  by dividing respondents into  $T$  groups using a map  $g : \{1, \dots, N\} \rightarrow \{1, \dots, T\}$ .

#### 6.1.4. Testing invariance of item parameters over time in the CML inference frame III: test based on estimated person locations

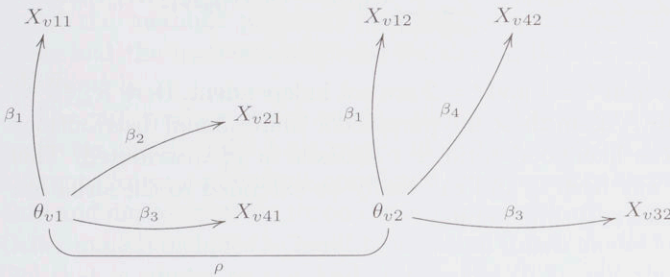
In the Rasch model the person locations can be estimated consistently using any item subset of items  $I_0 \subset \{1, \dots, I\}$  and this can be used as the basis for evaluating invariance of item parameters over time. Consider the CML estimates

$(\hat{\beta}_{i1})_{i=1,\dots,I}$  estimated using all data at  $t = 1$  and CML estimates  $(\hat{\beta}_{i2})_{i=1,\dots,I}$  estimated using all data at  $t = 2$ . Divide items into two disjoint subsets  $I_1$  and  $I_2$  and compare maximum likelihood estimates  $(\hat{\theta}_v)_{v=1,\dots,N}$  based on the subset of the data  $(X_{vi1})_{v=1,\dots,N;i \in I_1}$  and the subset of item parameters  $(\hat{\beta}_{i1})_{i \in I_1}$  to maximum likelihood estimates  $(\hat{\theta}_v^*)_{v=1,\dots,N}$  based on the subset of the data  $(X_{vi1})_{v=1,\dots,N;i \in I_2}$  and the subset of item parameters  $(\hat{\beta}_{i2})_{i \in I_2}$ . If item parameters are invariant then no systematic differences should be apparent. A way of evaluating would be to inspect a scatter plot of  $(\hat{\theta}_v, \hat{\theta}_v^*)_{v=1,\dots,N}$ . More than one random division of items into disjoint subsets can be considered.

## 6.2. Modeling items that are not invariant over time

If an item parameter  $\beta_{i_0}$  is not invariant over time this can be modeled in an incomplete design where items  $1, \dots, I$  with item parameters  $\beta_1, \dots, \beta_I$  are used at time  $t = 1$  and at times  $t = 2, \dots, T$  item  $i_0$  is replaced by a new item with item parameter  $\beta_{i_0 t}$ . This is illustrated in Figure 4 where item 2 is not invariant and is at time 2 regarded as a new item, item 4. A construction like this is sometimes referred to as *virtual items*.

FIG 4. Graphical representation of the extension to incomplete designs of the Andersen model (5.2).



## 6.3. Testing local dependence across time points

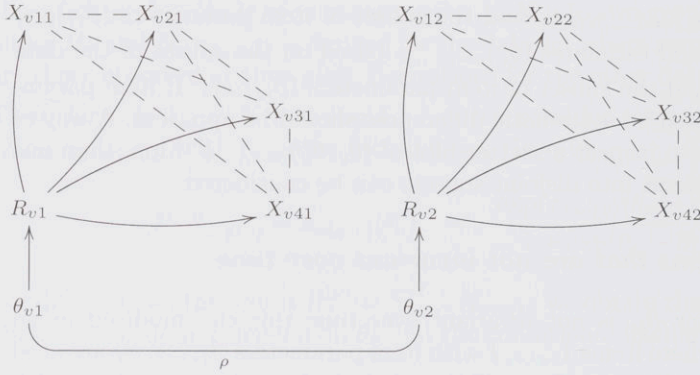
The assumption of local independence across time points in the model (5.2) may not be valid e.g. when respondents remember their previous response to an item. Tests of this assumption can be formulated in the MML and the CML inference frame. We consider the hypothesis of local independence of items across time points

$$H_0 : Pr(X_{vi1} = x_{vi1}, \dots, X_{viT} = x_{viT} | \bar{\theta}_v) = \prod_{t=1}^T Pr(X_{vit} = x_{vit} | \theta_{vt})$$

and test it against a specified alternative where, for an item  $i_0$ ,  $X_{vi_01}, \dots, X_{vi_0T}$  are not conditionally independent.



FIG 5. Andersen model (5.2) as a Graphical Rasch model: scores  $R_{v1}$  and  $R_{v2}$  separate items at time one from items at time two. Dashed lines indicate that items are not conditionally independent given their sum.



### 6.3.1. Testing local dependence across time points in the MML inference frame

The likelihood (5.3) can be tested against the extensions

$$L_{A,i_0}(\bar{\beta}, \bar{\theta}, \xi_{12}) = \exp\left(\sum_{v=1}^N \sum_{t=1}^T R_{vt} \theta_{vt} - \sum_{i \neq i_0} \sum_{t=1}^T C_{it} \beta_{it} - C_{i_0 1} C_{i_0 2} \xi_{12}\right) K_{i_0}(\bar{\beta}, \bar{\theta})^{-1}$$

where responses to item  $i_0$  at  $t = 1$  and  $t = 2$  are not independent. Here  $K_{i_0}(\bar{\beta}, \bar{\theta})$  is a normalizing constant. Again these are parametric multinomial distributions and the logarithm of the likelihood ratio test statistic is  $\chi_1^2$  distributed. This test can be performed for any item  $i_0$  and can easily be extended to the situation with several items and several time points.

### 6.3.2. Testing local dependence across time points in the CML inference frame I: Graphical Rasch model tests

Graphical Rasch models [21] are latent structure models where measurement models with properties that are similar to those of ordinary Rasch models are embedded in multivariate structural frameworks consisting of chain graph models [22]. Since items at each time point fit the Rasch model the raw scores  $R_{v1}$  and  $R_{v2}$  are sufficient for  $\theta_{v1}$  and  $\theta_{v2}$ , respectively. Adding the scores to the graphical representation yields a graphical Rasch model (Figure 5).

A consequence is that  $X_{vi1} \perp X_{vi2} | R_{v1}$  and  $X_{vi1} \perp X_{vi2} | R_{v2}$  for all items and these assumptions can be tested using the Mantel-Haenszel test for each item pair stratifying by either  $R_{v1}$  or  $R_{v2}$ .

#### 6.4. Overcoming local dependence across time points

If the data fits a Rasch model at each time local dependence across time points can be overcome by choosing disjoint sets of items for each time point, e.g. choosing  $X_{v11}, X_{v21}, \dots$  and  $X_{v22}, X_{v42}, \dots$  since this yields a model where results are not confounded by LD. This model can be used for estimating  $(\theta_{v1})_{v=1, \dots, N}$  and  $(\theta_{v2})_{v=1, \dots, N}$  on the same scale.

### 7. Discussion

Latent variable models for multivariate longitudinal ordinal responses often rely on distributional assumptions about the latent variables that are not easily tested. One way of avoiding this problem is to use Rasch models [7, 8] where conditional estimation can be carried out without problematic distributional assumptions. Assumptions about invariance of the measurement instrument over time are also made and this paper discussed how to test the assumptions of longitudinal IRT models and investigated the possibilities of using conditional rather than marginal approaches. A recent application in clinical trials [23] stresses that 'Inferences made using IRT are only valid when the model fits the data', and tests a multidimensional model for dichotomized items using Lagrange multiplier tests [24]. However the specific types of misfit discussed in this paper are not addressed.

All the methods proposed here can be generalized to ordinal polytomous items and the methods proposed for the MML inference frame also apply for 2PL models.

The multivariate Rasch models discussed here is a special case of random coefficients multinomial logit model [25]. This model is formulated in an MML inference frame and include correlated error terms to accommodate longitudinal data and impose a structure on the latent variables, e.g. an autoregressive model. Other models include the multidimensional Rasch model for learning and change [26] that is similar to the Andersen model (5.2), but the person locations are parametrized in terms of the initial level and changes.

The methods proposed for the MML inference frame can be incorporated in SAS using proc NLMIXED [27, 28] and conditional approaches can also be implemented in SAS [29]. The R package eRm [30] for fitting Rasch models using CML handles special cases of longitudinal data like the linear logistic test model.

### Acknowledgements

The authors would like to thank Svend Kreiner for valuable comments.

### References

- [1] HAMBLETON, R. K. and VAN DER LINDEN, W. J. (1997). *Handbook of Modern Item Response Theory*. Springer-Verlag, New York.



- [2] LIANG, K.-Y. and ZEGER, S. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13–22.
- [3] ANDERSEN, E. B. and MADSEN, M. (1977). Estimating the parameters of the latent population distribution. *Psychometrika*, 42, 357–74.
- [4] LIU and HEDEKER. (2006). A mixed-effects regression model for longitudinal multivariate ordinal data. *Biometrics*, 62, 261–228.
- [5] CAGNONE, MOUSTAKI and VASDEKIS. (2009). Latent variable models for multivariate longitudinal ordinal responses. *British Journal of Mathematical and Statistical Psychology*, 62, 401–415.
- [6] WANG and WILSON, M. (2005). The Rasch testlet model. *Applied Psychological Measurement*, 29, 126–149.
- [7] RASCH, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danish National Institute for Educational Research, Copenhagen.
- [8] FISCHER, G. H. and MOLENAAR, I. W. (1995). *Rasch models: Foundations, recent developments, and applications*. Springer-Verlag, New York.
- [9] NEYMAN, J. and SCOTT, E. L. (1948). Consistent estimates based on partially consistent observations. *Econometrika*, 16, 1–32.
- [10] PINHEIRO, J. C. and BATES, D. M. (1995). *Approximations to the log-likelihood function in the nonlinear mixed-effects model*. Journal of Computational and Graphical Statistics, 4, 12–35.
- [11] PINHEIRO, J. C. and CHAO, E. C. (2006). *Efficient laplacian and adaptive gaussian quadrature algorithms for multilevel generalized linear mixed models*. Journal of Computational and Graphical Statistics, 15(1), 58–81.
- [12] RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*. 2nd ed. New York: J. Wiley and Sons.
- [13] ZWINDERMAN, A. H. and VAN DEN WOLLENBERG, A. L. (1990). Robustness of Marginal Maximum Likelihood Estimation in the Rasch Model. *Applied Psychological Measurement*, 14, 73–81.
- [14] ANDERSEN, E. B. (1973) Conditional inference for multiple-choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, 26, 31–44.
- [15] CHRISTENSEN KB, BJORNER JB, KREINER S and PETERSEN JH. (2004). Latent Regression in log linear Rasch models. *Communications in Statistics A*, 33, 1295–1313.
- [16] GLAS, C. A. W. and VERHELST, N. D. (1995). Tests of Fit for Polytomous Rasch Models. In: Fischer & Molenaar (editors) *Rasch Models - Foundations, Recent Developments, and Applications*. Berlin: Springer-Verlag, 325–352.
- [17] FISCHER, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika*, 48, 3–26.
- [18] LINACRE, J. M. (1989). *Many-faceted Rasch Measurement*. Chicago: MESA.
- [19] ANDERSEN, E. B. (1985). Estimating latent correlations between repeated testings. *Psychometrika*, 50, 3–16.
- [20] KELDERMAN, H. (1984). Loglinear Rasch model tests. *Psychometrika*, 49, 223–245.
- [21] KREINER, S. and CHRISTENSEN, K. B. (2002). Graphical Rasch models.



- In: Mesbah, M., Cole, B. F., Lee, M. T., (editors) *Statistical Methods for Quality of Life Studies: Design, Measurements and Analysis*. Boston: Kluwer Academic Publishers, 187-203.
- [22] LAURITZEN, S. (1996). *Graphical models*. Oxford: Clarendon Press
  - [23] GLAS, C. A. W., GEERLINGS, H., VAN DE LAAR, M. A. F. J. and TAAL, E. (2009). Analysis of longitudinal randomized clinical trials using item response models. *Contemporary Clinical Trials*, 30, 158-170.
  - [24] GLAS, C. A. W. (1999). Modification indices for the 2-pl and the nominal response model. *Psychometrika*, 64, 273-294.
  - [25] ADAMS, WILSON and WANG. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*, 21, 1-23.
  - [26] EMBRETSON, S. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika*, 56, 495-516.
  - [27] CHRISTENSEN, K. B. (2006). Fitting Polytomous Rasch Models in SAS. *Journal of Applied Measurement* 7 407-417.
  - [28] HARDOUIN, J.-B. and MESBAH, M. (2007). The SAS Macro-Program %AnaQol to Estimate the Parameters of Item Responses Theory Models. *Communications in Statistics Simulation and Computation* 36 437-453.
  - [29] CHRISTENSEN, K. B. and BJORNER, J. B. (2003). *SAS macros for Rasch based latent variable modelling*. Research report 13, Department of Biostatistics, University of Copenhagen.
  - [30] MAIR, P. and HATZINGER, R. (2007). Extended Rasch Modeling: The eRm Package for the Application of IRT Models in R *Journal of Statistical Software*, 20, 9.