



%lrasch_mml: A SAS Macro for Marginal Maximum Likelihood Estimation in Longitudinal Polytomous Rasch Models

Maja Olsbjerg
University of Copenhagen

Karl Bang Christensen
University of Copenhagen

Abstract

Item response theory models are often applied when a number of items are used to measure a unidimensional latent variable. Originally proposed and used within educational research, they are also used when focus is on physical functioning or psychological well-being. Modern applications often need more general models, typically models for multidimensional latent variables or longitudinal models for repeated measurements. This paper describes a SAS macro that fits two-dimensional polytomous Rasch models using a specification of the model that is sufficiently flexible to accommodate longitudinal Rasch models. The macro estimates item parameters using marginal maximum likelihood estimation. A graphical presentation of item characteristic curves is included.

Keywords: polytomous Rasch model, longitudinal Rasch model, marginal maximum likelihood (MML) estimation, item parameter drift, response dependence, SAS macro.

1. Introduction

Item response theory (IRT) models were developed to describe probabilistic relationships between correct responses to a set of test items and continuous latent traits (Linden and Hambleton 1997). IRT models were originally developed and used in educational testing, where the models describe how the probability of a correct answer to an item in a test depends on ability, but they are applicable whenever location of persons and items on an underlying latent scale is of interest. Traditional applications in education often use dichotomous (correct/incorrect) item scoring, but polytomous items are common in other applications. The use of IRT models in new research fields increases the need for implementation in standard statistical software like SAS (SAS Institute Inc. 2013) or SPSS (IBM Corporation 2015). Estimation in IRT models using SAS has been the topic of several research papers (Rijmen,

Tuerlinckx, De Boeck, and Kuppens 2003; Smits, De Boeck, and Verhelst 2003; Nandakumar and Hotchkiss 2012). In particular, implementation of polytomous Rasch models in SAS has been discussed (Sheu, Chen, Su, and Wang 2005; Christensen 2006), and SAS macros that used these ideas are available (Christensen and Bjorner 2003; Hardouin and Mesbah 2007; Christensen and Olsbjerg 2013; Christensen 2013).

Many applications require more general models. Typically when multidimensional latent variables are considered or when repeated measurements are used. Longitudinal Rasch models were studied by Pastor and Beretvas (2006), who illustrated how these models can be seen as hierarchical generalized linear models and implemented in the software program **HLM** (Raudenbush, Bryk, Cheong, and Congdon 2004). **HLM** uses penalized quasi-likelihood for estimation, but as noted by Pastor and Beretvas, various estimation procedures and software programs for these kinds of models exist. An example is the random weights linear logistic test model (Rijmen and De Boeck 2002), which is a special case of the multidimensional random coefficients multinomial logit model (Adams, Wilson, and Wang 1997a) implemented in the computer program **ConQuest** (Wu, Adams, Wilson, and Haldane 2007). For these more general models implementation in standard software is also useful. This paper describes the SAS macro `%lrasch_mml` available from

http://biostat.ku.dk/~kach/index.html#lrasch_mml.

The macro fits polytomous longitudinal Rasch models using marginal maximum likelihood (MML; Bock and Aitkin 1981; Thissen 1982; Zwiderman and Wollenberg 1990). It estimates item parameters and the parameters of a two-dimensional latent distribution and plots item characteristic curves. It is sufficiently flexible to model item parameter drift and response dependence across time points. Recently IRT models have been used increasingly in health status measurement and evaluation of patient reported outcomes (PROs) like physical functioning and psychological well-being (Reeve *et al.* 2007). E.g., the simplest IRT model, the Rasch (1960) model (Fischer and Molenaar 1995; Christensen, Kreiner, and Mesbah 2013), is increasingly used for validation of measurement instruments (Tennant and Conaghan 2007) and has been shown to be superior to classical approaches (Blanchin *et al.* 2011). The macro is illustrated using a data set from a longitudinal study of health related quality of life (HRQoL) in non-melanoma skin cancer patients.

It is of considerable importance that model assumptions are checked before the results of a statistical analysis are reported. This is especially important for IRT models where the model requirements are a mathematical formulation of measurement requirements. Thus, existing methodology should be used to make sure that the Rasch model fits at each time point. This can be done using the fit statistics implemented in proprietary software packages like **RUMM** (Andrich, Sheridan, and Luo 2010) and **WINSTEPS** (Linacre 2011) or in free stand-alone software like **DIGRAM** (Kreiner 2003). In SAS the macros `%AnaQo1` (Hardouin and Mesbah 2007), `GLIMMIX_Rasch` (Chen, Li, and Kromrey 2013) and `%rasch_mml` (Christensen and Olsbjerg 2013) that report infit and outfit test statistics can be used.

Beyond the requirement of fit of the Rasch model at each time point special requirements arise from the use of the Rasch model for longitudinal data. Two of these requirements, item parameter drift and local dependence across time points, can be tested in a likelihood frame work using the SAS macro `%lrasch_mml`.

1.1. Item parameter drift

To measure trends over time is important in many areas, also for latent variables (Goldstein 1983), but it is important that the item parameters do not change over time (Miller and Fitzpatrick 2009; Chan, Drasgow, and Sawin 1999; Chan *et al.* 1999; Wells, Subkoviak, and Serlin 2002). The macro `%lrasch_mml` can be used to test this assumption since models where one or more items do not stay constant over time can be fitted. Different methods for detection of item parameter drift exist (Donoghue and Isham 1998; DeMars 2004; Galdin and Laurencelle 2010). The macro `%lrasch_mml` makes it possible to test the assumption of item parameter invariance using likelihood ratio tests.

1.2. Local dependence across time points

Local dependence has been extensively studied in the psychometric literature, especially within the class of Rasch models (Kelderman 1984; Kreiner and Christensen 2004, 2007; Marais and Andrich 2008a,b). However in the context of longitudinal studies references are sparse. In a study of response dependence and measurement of change, Marais (2009) found that when the assumption of local independence across time is violated it can lead to incorrect conclusions. A few tests of this assumption have been discussed (Olsbjerg and Christensen 2013). Andrich and Kreiner (2010) proposed a way of quantifying local dependence for two dichotomous items. Their method is based on splitting an item into two new ones according to the responses to another item. This method has been generalized to polytomous items and to other IRT models and can be used to overcome local response dependence across time points (Olsbjerg and Christensen 2014). The macro `%lrasch_mml` makes it possible to include splitted items and to test the assumption of local independence across time points using likelihood ratio tests.

2. The unidimensional polytomous Rasch model

IRT models can be seen as a mathematical formalization of the following measurement requirements: (i) items measure only one latent variable, (ii) expected item scores increase with the underlying latent variable, (iii) items are sufficiently different to avoid redundancy, and (iv) items function in the same way across sub-populations. Let θ denote the latent variable and $\bar{X} = (X_i)_{i \in I}$ a vector of item responses. The requirements can be written as

$$(i) \quad \theta \in \mathbb{R},$$

$$(ii) \quad \forall i \in I : \theta \mapsto E(X_i|\theta) \text{ is increasing,}$$

$$(iii) \quad P(\bar{X} = \bar{x}|\theta) = \prod_{i \in I} P(X_i = x_i|\theta) \text{ for all } \theta,$$

$$(iv) \quad P(X_i = x_i|Y, \theta) = P(X_i = x_i|\theta) \text{ for all items } i \in I \text{ and all variables } Y.$$

The requirements (i)–(iv) are referred to as unidimensionality, monotonicity, local independence and absence of differential item functioning (DIF), respectively. Fit of observed data to an IRT model thus implies that these requirements are met and evaluation of model fit is crucial. Assume that item $i \in I$ has $m_i + 1$ response categories represented by the numbers

$0, \dots, m_i$ and let the stochastic variable X_i with realization x_i denote the response. For items $i \in I$ the *polytomous Rasch model* is given by probabilities

$$P(X_i = x_i | \theta) = \frac{\exp(x_i \theta + \eta_{ix_i})}{K_i}, \quad (1)$$

where $K_i = K_i(\theta) = K_i(\theta, \bar{\eta}_i) = \sum_{l=0}^{m_i} \exp(l\theta + \eta_{il})$ with $\bar{\eta}_i = (\eta_{ih})_{h=1, \dots, m_i}$ and $\eta_{i0} = 0$. An alternative parametrization can be obtained by replacing the vector of item parameters $\bar{\eta}_i$ with the thresholds

$$\bar{\beta}_i = (\beta_{ik})_{k=1, \dots, m_i}, \quad (2)$$

defined by $\beta_{ik} = -(\eta_{ik} - \eta_{ik-1})$. These can be interpreted as the locations on the latent continuum where the probabilities of choosing adjacent categories intersect. Under the assumption (iii) of local independence the distribution of the vector $\bar{X} = (X_i)_{i \in I}$ with realization $\bar{x} = (x_i)_{i \in I}$ is given by the probabilities

$$P(\bar{X} = \bar{x} | \theta) = \exp(r\theta) \frac{\exp(\sum_{i \in I} \eta_{ix_i})}{K(\theta)}, \quad (3)$$

where $K(\theta) = \prod_{i \in I} K(\theta, \bar{\eta}_i)$ and $r = \sum_{i=1}^I x_i$. It is clear from Equation 3 that the sum of item responses $R = \sum_{i=1}^I X_i$ is sufficient for θ . The contribution to the *joint log likelihood* for a person v with item responses $(x_i)_{i \in I_v}$, $I_v \subset I$ is

$$l((\bar{\eta}_i)_{i \in I_v}, \theta_v) = r\theta_v + \sum_{i \in I_v} \eta_{ix_i} - \log K_{I_v}(\theta_v), \quad (4)$$

with $r = \sum_{i \in I_v} x_i$ and $K_{I_v}(\theta_v) = \prod_{i \in I_v} K_i(\theta_v)$. Restrictions are needed to ensure that the model in Equation 4 is identified since for all $(\theta, \bar{\eta}_i)$

$$P(X_i = x_i | \theta, \bar{\eta}_i) = P(X_i = x_i | \theta^*, \bar{\eta}_i^*)$$

for $\theta^* = \theta - \kappa$ and $\bar{\eta}_i^* = (\eta_{ih} + \kappa h)_{h=1, \dots, m_i}$. We can make sure that the model is identified by imposing a linear restriction on the β 's or on the θ 's.

Jointly estimating all parameters yields inconsistent estimates, since the number of parameters increases with the sample size (Neyman and Scott 1948). This can be overcome by assuming that θ is drawn from a normal distribution and maximizing the marginal log likelihood (Bock and Aitkin 1981; Thissen 1982; Zwinderman and Wollenberg 1990). The contribution for person v with item responses $(x_i)_{i \in I_v}$, $I_v \subset I$ to the marginal likelihood is

$$l_M((\bar{\eta}_i)_{i \in I_v}) = \sum_{i \in I_v} \eta_{ix_i} + \log \int_{\mathbb{R}} \frac{\exp(r\theta)}{K_{I_v}(\theta)} \varphi(\theta) d\theta. \quad (5)$$

This model can be fitted with PROC NLMIXED in SAS using an adaptive Gaussian quadrature. If the mean of the normal distribution is fixed at zero all item parameters can be identified, and if a linear restriction is imposed on the item parameters the mean *and* variance of the normal distribution can be identified. Thus depending on the linear restriction imposed to ensure identifiability two different ways of parameterizing the model exist:

- (i) $\sum_{i \in I} \sum_{h=1}^{m_i} \eta_{ih} = 0$ (or equivalently $\sum_{i \in I} \sum_{h=1}^{m_i} \beta_{ih} = 0$), $\theta \sim N(\mu, \sigma^2)$;
- (ii) $(\beta_{ih})_{i \in I; h=1, \dots, m_i}$ unrestricted, $\theta \sim N(0, \sigma^2)$

The polytomous Rasch model was proposed by Andersen (1977). Masters (1982) called this model the *partial credit model* deriving probabilities in Equation 1 from the requirement that the conditional probabilities $P(X_i = k | X_i \in \{k-1, k\}; \theta)$ fit a dichotomous Rasch model.

3. The two-dimensional polytomous Rasch model

Assume that the set of items can be split into two new, i.e.,

$$I = I_1 \cup I_2, \quad (6)$$

with items in I_1 and I_2 measuring latent variables θ_1 and θ_2 respectively. Two situations are common: (i) θ_1 and θ_2 are distinct, but correlated latent variables, (ii) θ_1 and θ_2 represent repeated measurements of the same latent variable. The latter is the focus in this paper. If the distribution of the items is as specified by the polytomous Rasch model and the assumption (iii) of local independence holds the vector \bar{X} with realization \bar{x} is distributed as follows

$$\begin{aligned} P(\bar{X} = \bar{x} | \theta) &= \frac{\exp(\sum_{i \in I_1} x_i \theta_1 + \sum_{i \in I_1} \eta_{ix_i} + \sum_{i \in I_2} x_i \theta_2 + \sum_{i \in I_2} \eta_{ix_i})}{K_1(\theta_1) K_2(\theta_2)} \\ &= \frac{\exp(r_1 \theta_1 + r_2 \theta_2 + \sum_{i \in I} \eta_{ix_i})}{K_1(\theta_1) K_2(\theta_2)} \end{aligned} \quad (7)$$

where

$$\begin{aligned} r_1 &= \sum_{i \in I_1} x_i, & r_2 &= \sum_{i \in I_2} x_i, \\ K_1(\theta_1) &= \prod_{i \in I_1} K_i(\theta_1, \bar{\eta}_i), & K_2(\theta_2) &= \prod_{i \in I_2} K_i(\theta_2, \bar{\eta}_i). \end{aligned}$$

Again Neyman's factorization theorem shows that $R_1 = \sum_{i \in I_1} X_i$ and $R_2 = \sum_{i \in I_2} X_i$ are sufficient for θ_1 and θ_2 , respectively.

The contribution to the *joint log likelihood* for a person v with item responses $(x_i)_{i \in I_v}$, $I_v \subset I$ is given by

$$l((\bar{\eta}_i)_{i \in I_v}, \theta_1, \theta_2) = r_1 \theta_1 + r_2 \theta_2 + \sum_{i \in I_v} \eta_{ix_i} - \log K_1(\theta_1) - \log K_2(\theta_2) \quad (8)$$

with $K_1(\theta_1) = K_{I_v \cap I_1}(\theta_1)$ and $K_2(\theta_2) = K_{I_v \cap I_2}(\theta_2)$. Again, jointly estimating all parameters does not provide consistent estimates and restrictions are needed in order to ensure that the model is identified. This can be done by placing a restriction on either the item parameters or the person parameters.

3.1. Estimation of item parameters

Assuming that (θ_1, θ_2) is drawn from a two-dimensional normal distribution

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \sim N \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix} \right) \quad (9)$$

yields the *two-dimensional marginal log likelihood function* to which the contribution of a person v with item responses $(x_i)_{i \in I_v}$, $I_v \subset I$, is

$$l_M((\bar{\eta}_i)_{i \in I_v}) = \sum_{i \in I_v} \eta_{ix_i} + \log \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\exp(r_1\theta_1 + r_2\theta_2)}{K_1(\theta_1)K_2(\theta_2)} \varphi(\theta_1, \theta_2) d\theta_1 d\theta_2. \quad (10)$$

Two-dimensional Rasch models of this kind were originally discussed by Andersen (1985) and Embretson (1991), who both considered longitudinal measurement. Both models fit within the general framework for multidimensional Rasch models described by Adams and colleagues (Adams *et al.* 1997a; Adams, Wilson, and Wu 1997b).

3.2. Parameter restrictions

Restrictions are needed in order to ensure that the model is identified. Typically this is done by assuming that $\mu_1 = \mu_2 = 0$. For the special case of longitudinal data further restrictions are often imposed and other parametrizations are of interest. Consider the situation where the same set of items is used at two time points and let the response to item i at time t be denoted by X_{it} . In the longitudinal Rasch model it is a natural restriction that the item parameters do not change over time, i.e., that for all items i

$$\bar{\eta}_{i,2} = \bar{\eta}_{i,1}. \quad (11)$$

When this restriction is imposed there is no need for the assumption that $\mu_1 = \mu_2 = 0$. In fact, for longitudinal data a model where the mean changes over time, but the item parameters are invariant as in Equation 11, is often preferable. The original formulations of the longitudinal Rasch model for dichotomous items (Andersen 1985; Embretson 1991) used this restriction, but a more general model where only a subset of the items are restricted to be equal would also be identified. Depending on the linear restriction imposed to ensure identifiability different ways of parameterizing the model exist. The macro %lrasch_mml uses

- (i) $(\bar{\eta}_{i,1})_{i \in I; h=1, \dots, m_i}$ unrestricted,
- (ii) $\bar{\eta}_{i,2} = \bar{\eta}_{i,1}$ for all $i \in I$,
- (iii) $\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ \mu \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \right)$.

The assumption that item parameters do not change over time can be relaxed. This is outlined in the following Section 3.3.

3.3. Relaxing assumptions

The longitudinal Rasch model formulated for the situation where θ_1 and θ_2 are (correlated) values of the same latent variable at two distinct time points and the two sets of items I_1 and I_2 are identical. This can be relaxed in two ways: by allowing item parameter drift (as outlined in Section 1.1) and by allowing local dependence across time points (as outlined in Section 1.2).

Item parameter drift can be allowed by specifying a model where the item sets I_1 and I_2 are overlapping, but not identical. To illustrate how this could be done, note that allowing an item $i \in I$ to change over time can be specified using $I_1 = (I \setminus \{i\}) \cup \{i'_1\}$ and $I_2 = (I \setminus \{i\}) \cup \{i'_2\}$ where the item is modeled as two separate items, each one only administered at one of the time points.

Two similar items are highly correlated, and the correlation could be even higher than what the underlying latent variable accounts for. For this reason the requirement (iii) of local independence is formulated as a requirement of non-redundancy. In the longitudinal Rasch model this requirement is imposed at each time point, but an additional requirement of local dependence across time points is also imposed: Responses given to the same item at the two points should be independent given (θ_1, θ_2) . In the unidimensional Rasch model local dependence can be modeled using log linear Rasch models (Kelderman 1984; Kreiner and Christensen 2004), but also by “splitting” one item with respect to the observed values of another (Andrich and Kreiner 2010). In our implementation of the longitudinal Rasch model the latter of these approaches is used. A detailed account of how splitting an item at the second time point based on observed responses at the first time point is provided by Olsbjerg and Christensen (2014). Briefly, a model where local dependence across time points is modeled for an item $i \in I$ can be formulated by defining $m_i + 1$ time 2 items

$$i_2(m) = \begin{cases} m, & \text{if } x_{1i} = m; \\ ., & \text{otherwise.} \end{cases} \quad (m = 0, 1, \dots, m_i) \quad (12)$$

where $.$ denotes a missing value. The extended model uses $I_1 = I$ and

$$I_2 = (I \setminus \{i\}) \cup \{i_2(m) : m = 0, 1, \dots, m_i\}.$$

Since missing values are easily included in the implemented MML framework the parameters of this model are readily estimated.

4. Implementation in SAS

The SAS macro %lrasch_mml uses PROC NL MIXED to estimate item parameters, and the parameters of the two-dimensional normal distribution. PROC NL MIXED fits nonlinear mixed models (Rijmen *et al.* 2003; Smits *et al.* 2003) and is very flexible because the conditional distribution given the random effects can be specified to be a general distribution using SAS programming statements. The NL MIXED procedure maximizes an approximation to the likelihood integrated over the random effects. Different integral approximations are available, the principal one being adaptive Gaussian quadrature.

The macro has two required input statements:

- DATA specifying the input data set;
- ITEM_NAMES identifying the names and response formats of the items.

The macro also has four optional input statements:

- ANCHOR specifying the items that should be restricted to have the same item parameters at the two time points (default value ALL, specifying that all items have the same parameters, cf. Equation 11);

- **SPLIT** specifying the items that should be split for local dependence, cf. [Olsbjerg and Christensen \(2014\)](#) (default value `NONE`, indicating that no items should be split);
- **ICC** specifying whether or not the macro should plot item characteristic curves (default value `NO`). Note that these curves should rightly be referred to as categories probability curves.
- **OUT** specifying the prefix of output data sets generated by the macro (default value `LRASCH`).

The assumption that item parameters do not change over time can thus be relaxed. This makes it possible to test for item parameter drift, and to specify models where item parameter drift is present. The SAS macro `%lrasch_mml` creates three output data sets:

- The data set `OUT_logl` contains the maximum value of the conditional log likelihood function and the Akaike information criterion (AIC; [Akaike 1974](#)), the Bayesian information criterion (BIC; [Schwarz 1978](#)) and the sample size corrected version of the AICC ([Burnham and Anderson 2002](#)).
- The data set `OUT_thresholds` contains the item parameter estimates (using the PCM parametrization), the standard errors of the item parameter estimates, and 95% confidence intervals.
- The data set `OUT_poppar` contains population parameter estimates, i.e., the change in mean μ , the standard deviations σ_1 and σ_2 and the latent correlation ρ .

5. Example: Longitudinal HRQoL data

Patient-reported outcomes (PROs) are used to capture patients' perception of a disease and its impact on daily living. SCQoL is a HRQoL questionnaire developed for non-melanoma skin cancer (NMSC; [Vinding, Christensen, Esmann, Olesen, and Jemec 2013](#)) that consists of 9 items scored on a standard 4-point Likert scale. The data used in this example were collected from patients with NMSC undergoing surgery at the Department of Plastic Surgery at Roskilde Hospital.

The analyses presented here can be performed using the SAS macro `%lrasch_mml.sas` and the supplementary files available along with this manuscript (or from the homepage http://biostat.ku.dk/~kach/index.html#lrasch_mml):

- the example data `scqol.sas7bdat` containing the responses to the SCQoL questionnaire,
- the sample code `v67c02.sas` described in the following.

Participants responded to the skin cancer quality of life (SCQoL) questionnaire before the operation and 3 months after. A total of 101 patients responded at baseline, 14 patients did not respond at follow-up. The marginal distribution of item responses is computed using the statements

Time point	Item	0	1	2	3
Baseline	SC01	62 (61.4%)	30 (29.7%)	9 (8.9%)	0
	SC02	23 (22.8%)	38 (37.6%)	24 (23.8%)	16 (15.8%)
	SC03	53 (52.5%)	30 (29.7%)	15 (14.9%)	3 (3.0%)
	SC04	34 (33.7%)	35 (34.7%)	24 (23.8%)	8 (7.9%)
	SC05	92 (91.1%)	4 (4.0%)	2 (2.0%)	3 (3.0%)
	SC06	67 (66.3%)	27 (26.7%)	6 (5.9%)	1 (1.0%)
	SC07	32 (31.7%)	51 (50.5%)	13 (12.9%)	5 (5.0%)
	SC08	72 (71.3%)	19 (18.8%)	7 (6.9%)	3 (3.0%)
	SC09	35 (34.7%)	36 (35.6%)	21 (20.8%)	9 (8.9%)
Follow-up	SC01	45 (51.7%)	24 (27.6%)	14 (16.1%)	4 (4.6%)
	SC02	29 (33.3%)	33 (37.9%)	18 (20.7%)	7 (8.1%)
	SC03	65 (74.7%)	18 (20.7%)	4 (4.6%)	0
	SC04	44 (50.6%)	34 (39.1%)	6 (6.9%)	3 (3.5%)
	SC05	72 (82.8%)	8 (9.2%)	5 (5.8%)	2 (2.3%)
	SC06	70 (80.5%)	15 (17.2%)	2 (2.3%)	0
	SC07	28 (32.2%)	43 (49.4%)	12 (13.8%)	4 (4.6%)
	SC08	66 (75.9%)	17 (19.5%)	3 (3.5%)	1 (1.2%)
	SC09	19 (21.8%)	28 (32.2%)	29 (33.3%)	11 (12.6%)

Table 1: The marginal distribution of item responses in the SCQoL data.

```
proc freq data=data.scqol;
  table b_sc01-b_sc09 fu_SC01-fu_SC09 / nocum;
run;
```

and is shown in Table 1.

We note that the maximum response '3' is not observed for the item SC01 at baseline and for the items SC03 and SC06 at follow-up. In order to fit the longitudinal Rasch model to these data we have to specify the structure of the input data. We specify variable names and response formats by defining the ITEM_NAMES data set, writing

```
data inames;
  input item_no item_names1 $ item_names2 $ item_text $ max1 max2;
  datalines;
  1 b_SC01 fu_SC01 SC01 2 3
  2 b_SC02 fu_SC02 SC02 3 3
  3 b_SC03 fu_SC03 SC03 3 2
  4 b_SC04 fu_SC04 SC04 3 3
  5 b_SC05 fu_SC05 SC05 3 3
  6 b_SC06 fu_SC06 SC06 3 2
  7 b_SC07 fu_SC07 SC07 3 3
  8 b_SC08 fu_SC08 SC08 3 3
  9 b_SC09 fu_SC09 SC09 3 3
  ;
run;
```

the ITEM_NAMES data set contains information about the items, specifically the variable

`item_names1` contains item names of time 1 items, the variable `item_names2` contains item names of time 2 items, the variable `item_text` contains text for plots, and the variables `max1` and `max2` contain the maximum score for time 1 and time 2 items, respectively. Thus the code specifies that the item SC01 at baseline and the items SC03 and SC06 at follow-up are to be considered to have only three response options. In this situation it does not make sense to require invariance of items 1, 3 and 6.

We include the SAS macro (provided in the supplements) using the statement

```
%include 'lrasch_mml.sas';
```

or alternatively including it directly from the web via

```
FILENAME lrasch URL 'http://biostat.ku.dk/~kach/macro/lrasch_mml.sas';
%include lrasch;
```

We fit the longitudinal Rasch model using the statement

```
%lrasch_mml(DATA=data.scqol, ITEM_NAMES=inames,
  ANCHOR=b_SC02 b_SC04 b_SC05 b_SC07 b_SC08 b_SC09,
  SPLIT=NONE, OUT=SCQOL, ICC=YES);
```

This statement specifies that the parameters of the items SC02, SC04, SC05, SC07, SC08, and SC09 do not change over time. The log likelihood value $-2 \log L = 2948.2$ is saved in the data set SCQOL_logl and printed in the output window.

scqol: LRASCH MML estimation loglikelihood

Descr	Value
-2 Log Likelihood	2948.2
AIC (smaller is better)	3022.2
AICC (smaller is better)	3023.9
BIC (smaller is better)	3118.9

Estimates of the item parameters are saved in the data set SCQOL_thresholds and are also printed in the output window.

scqol: LRASCH MML estimated thresholds

Label	Estimate	Standard Error	Lower	Upper
b_SC01 1	0.84	0.28	0.28	1.40
b_SC01 2	2.27	0.43	1.41	3.13
b_SC02 1	-0.88	0.25	-1.38	-0.38
b_SC02 2	0.82	0.25	0.32	1.33
b_SC02 3	1.58	0.32	0.95	2.21
b_SC03 1	0.49	0.29	-0.08	1.06

b_SC03 2	1.52	0.37	0.79	2.26
b_SC03 3	3.12	0.68	1.77	4.48
b_SC04 1	-0.16	0.23	-0.62	0.31
b_SC04 2	1.40	0.28	0.85	1.95
b_SC04 3	2.23	0.40	1.43	3.03
b_SC05 1	2.95	0.34	2.27	3.63
b_SC05 2	1.68	0.52	0.66	2.70
b_SC05 3	2.06	0.64	0.79	3.32
b_SC06 1	1.09	0.29	0.52	1.66
b_SC06 2	2.62	0.50	1.63	3.62
b_SC06 3	3.61	1.13	1.37	5.85
b_SC07 1	-0.82	0.24	-1.29	-0.36
b_SC07 2	1.87	0.28	1.32	2.43
b_SC07 3	2.28	0.44	1.41	3.15
b_SC08 1	1.53	0.25	1.04	2.02
b_SC08 2	2.29	0.40	1.49	3.09
b_SC08 3	2.55	0.64	1.28	3.82
b_SC09 1	-0.74	0.25	-1.24	-0.24
b_SC09 2	0.54	0.25	0.05	1.04
b_SC09 3	1.90	0.32	1.26	2.53
fu_SC01 1	0.45	0.32	-0.19	1.09
fu_SC01 2	1.07	0.39	0.29	1.84
fu_SC01 3	2.35	0.62	1.12	3.57
fu_SC02 1	-0.88	0.25	-1.38	-0.38
fu_SC02 2	0.82	0.25	0.32	1.33
fu_SC02 3	1.58	0.32	0.95	2.21
fu_SC03 1	1.44	0.33	0.78	2.09
fu_SC03 2	2.43	0.60	1.24	3.63
fu_SC04 1	-0.16	0.23	-0.62	0.31
fu_SC04 2	1.40	0.28	0.85	1.95
fu_SC04 3	2.23	0.40	1.43	3.03
fu_SC05 1	2.95	0.34	2.27	3.63
fu_SC05 2	1.68	0.52	0.66	2.70
fu_SC05 3	2.06	0.64	0.79	3.32
fu_SC06 1	1.77	0.35	1.08	2.45
fu_SC06 2	3.04	0.79	1.46	4.61
fu_SC07 1	-0.82	0.24	-1.29	-0.36
fu_SC07 2	1.87	0.28	1.32	2.43
fu_SC07 3	2.28	0.44	1.41	3.15
fu_SC08 1	1.53	0.25	1.04	2.02
fu_SC08 2	2.29	0.40	1.49	3.09
fu_SC08 3	2.55	0.64	1.28	3.82
fu_SC09 1	-0.74	0.25	-1.24	-0.24
fu_SC09 2	0.54	0.25	0.05	1.04
fu_SC09 3	1.90	0.32	1.26	2.53

The item parameters are shown in Table 2, note that the data set `SCQOL_thresholds` also con-

Item	$\hat{\beta}$	<i>s.e.</i> ($\hat{\beta}$)	Item	$\hat{\beta}$	<i>s.e.</i> ($\hat{\beta}$)
b_SC01 1	0.84	0.28	fu_SC01 1	0.45	0.32
b_SC01 2	2.27	0.43	fu_SC01 2	1.07	0.39
			fu_SC01 3	2.35	0.62
b_SC02 1	-0.88	0.25	fu_SC02 1	-0.88	0.25
b_SC02 2	0.82	0.25	fu_SC02 2	0.82	0.25
b_SC02 3	1.58	0.32	fu_SC02 3	1.58	0.32
b_SC03 1	0.49	0.29	fu_SC03 1	1.44	0.33
b_SC03 2	1.52	0.37	fu_SC03 2	2.43	0.60
b_SC03 3	3.12	0.68			
b_SC04 1	-0.16	0.23	fu_SC04 1	-0.16	0.23
b_SC04 2	1.40	0.28	fu_SC04 2	1.40	0.28
b_SC04 3	2.23	0.40	fu_SC04 3	2.23	0.40
b_SC05 1	2.95	0.34	fu_SC05 1	2.95	0.34
b_SC05 2	1.68	0.52	fu_SC05 2	1.68	0.52
b_SC05 3	2.06	0.64	fu_SC05 3	2.06	0.64
b_SC06 1	1.09	0.29	fu_SC06 1	1.77	0.35
b_SC06 2	2.62	0.50	fu_SC06 2	3.04	0.79
b_SC06 3	3.61	1.13			
b_SC07 1	-0.82	0.24	fu_SC07 1	-0.82	0.24
b_SC07 2	1.87	0.28	fu_SC07 2	1.87	0.28
b_SC07 3	2.28	0.44	fu_SC07 3	2.28	0.44
b_SC08 1	1.53	0.25	fu_SC08 1	1.53	0.25
b_SC08 2	2.29	0.40	fu_SC08 2	2.29	0.40
b_SC08 3	2.55	0.64	fu_SC08 3	2.55	0.64
b_SC09 1	-0.74	0.25	fu_SC09 1	-0.74	0.25
b_SC09 2	0.54	0.25	fu_SC09 2	0.54	0.25
b_SC09 3	1.90	0.32	fu_SC09 3	1.90	0.32

Table 2: The estimated item parameters from the SCQoL data.

tains 95% confidence intervals for the item parameter estimates. The data set SCQOL_POPPAR contains estimates of the average change in latent variable ($\mu = -0.047$, 95% CI: -0.348 to 0.253), estimated variances, and estimated latent correlation ($\rho = 0.653$, 95% CI: 0.475 to 0.831). This data set is also printed in the output window.

scqol: LRASCH estimated change in latent mean (mu), variances, and latent correlation (rho)

Parameter	Estimate	Standard Error	Lower	Upper
mu	-0.047	0.152	-0.348	0.253
rho	0.653	0.090	0.475	0.831
sigma1	1.319	0.138	1.045	1.594
sigma2	1.141	0.136	0.871	1.412

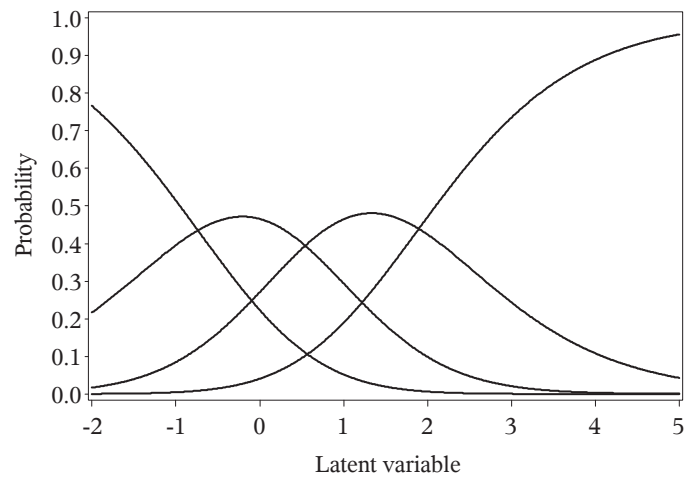
b_SC09 and fu_SC09

Figure 1: Item characteristic curves (ICCs) for an anchored item (plotted with the option `ICC = YES`).

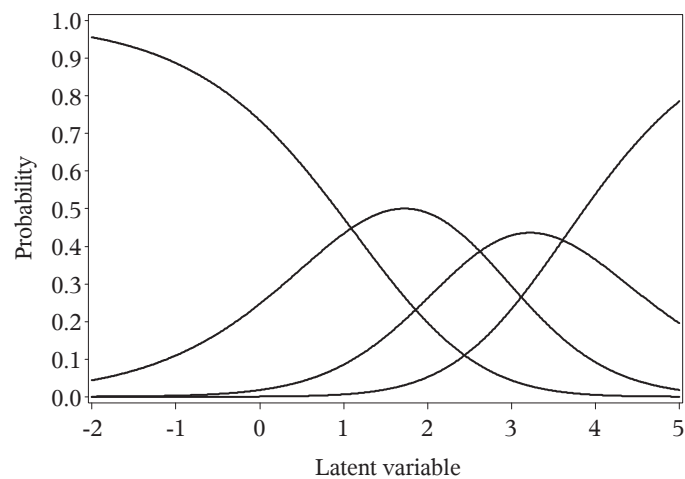
b_SC06

Figure 2: Item characteristic curves (ICCs) for an unanchored item (plotted with the option `ICC = YES`).

The ICCs for selected items are shown in Figures 1 and 2.

5.1. Item parameter drift

To test that the parameters do not change over time we specify, e.g.,

```
%lrasch_mml(DATA=data.scqol, ITEM_NAMES=inames,
  ANCHOR=b_SC04 b_SC05 b_SC07 b_SC08 b_SC09,
  SPLIT=NONE, OUT=AN2);
```

where the item SC02 is no longer anchored. The log likelihood value is $-2 \log L = 2941.8$.

AN2: LRASCH MML estimation loglikelihood

Descr	Value
-2 Log Likelihood	2941.9
AIC (smaller is better)	3021.9
AICC (smaller is better)	3023.8
BIC (smaller is better)	3126.5

Compared to the previously fitted model, where $\beta = (-0.88, 0.82, 1.58)$ at both time points this model has $\beta = (-1.25, 0.73, 1.44)$ at time 1 and $\beta = (-0.41, 1.06, 1.98)$ at time 2.

AN2: LRASCH MML estimated thresholds

Label	Estimate	Standard Error	Lower	Upper
b_SC01 1	0.84	0.29	0.28	1.41
b_SC01 2	2.31	0.44	1.44	3.18
b_SC02 1	-1.25	0.33	-1.91	-0.58
b_SC02 2	0.73	0.32	0.10	1.35
b_SC02 3	1.44	0.38	0.69	2.20
b_SC03 1	0.49	0.29	-0.09	1.06
b_SC03 2	1.55	0.37	0.82	2.29
b_SC03 3	3.17	0.68	1.82	4.53
b_SC04 1	-0.09	0.24	-0.56	0.38
b_SC04 2	1.47	0.28	0.92	2.03
b_SC04 3	2.30	0.41	1.49	3.10
b_SC05 1	3.02	0.34	2.34	3.70
b_SC05 2	1.75	0.52	0.72	2.78
b_SC05 3	2.12	0.64	0.86	3.39
b_SC06 1	1.10	0.29	0.53	1.68
b_SC06 2	2.66	0.50	1.67	3.66
b_SC06 3	3.67	1.13	1.42	5.91
b_SC07 1	-0.76	0.24	-1.23	-0.28
b_SC07 2	1.94	0.28	1.38	2.51
b_SC07 3	2.35	0.44	1.47	3.22
b_SC08 1	1.60	0.25	1.10	2.10

b_SC08 2	2.36	0.41	1.56	3.17
b_SC08 3	2.62	0.64	1.34	3.89
b_SC09 1	-0.68	0.26	-1.18	-0.17
b_SC09 2	0.61	0.25	0.11	1.12
b_SC09 3	1.97	0.32	1.32	2.61
fu_SC01 1	0.59	0.33	-0.06	1.24
fu_SC01 2	1.19	0.40	0.41	1.98
fu_SC01 3	2.46	0.62	1.23	3.69
fu_SC02 1	-0.41	0.34	-1.08	0.26
fu_SC02 2	1.06	0.36	0.35	1.77
fu_SC02 3	1.98	0.50	0.99	2.98
fu_SC03 1	1.57	0.34	0.90	2.24
fu_SC03 2	2.55	0.60	1.35	3.74
fu_SC04 1	-0.09	0.24	-0.56	0.38
fu_SC04 2	1.47	0.28	0.92	2.03
fu_SC04 3	2.30	0.41	1.49	3.10
fu_SC05 1	3.02	0.34	2.34	3.70
fu_SC05 2	1.75	0.52	0.72	2.78
fu_SC05 3	2.12	0.64	0.86	3.39
fu_SC06 1	1.90	0.35	1.20	2.59
fu_SC06 2	3.15	0.80	1.57	4.73
fu_SC07 1	-0.76	0.24	-1.23	-0.28
fu_SC07 2	1.94	0.28	1.38	2.51
fu_SC07 3	2.35	0.44	1.47	3.22
fu_SC08 1	1.60	0.25	1.10	2.10
fu_SC08 2	2.36	0.41	1.56	3.17
fu_SC08 3	2.62	0.64	1.34	3.89
fu_SC09 1	-0.68	0.26	-1.18	-0.17
fu_SC09 2	0.61	0.25	0.11	1.12
fu_SC09 3	1.97	0.32	1.32	2.61

Thus this model has three additional parameters. This yields a likelihood ratio test statistic of $-2 \log L = 2948.2 - 2941.9 = 6.3$, which on three degrees of freedom is not significant. We note that the population parameters in AN2_POPPAR do not differ much from those estimated under the previous model.

AN2: LRASCH estimated change in latent mean (μ), variances, and latent correlation (ρ)

Parameter	Estimate	Standard Error	Lower	Upper
μ	0.090	0.162	-0.232	0.412
ρ	0.653	0.090	0.476	0.831
σ_1	1.353	0.143	1.069	1.637
σ_2	1.124	0.136	0.855	1.393

This likelihood ratio test can be computed using the SAS code

```

data lrt;
  merge  SCQOL_logl(rename=(value=logl_0))
        AN2_logl(rename=(value=logl));
  where Descr='-2 Log Likelihood';
  lrt=logl_0-logl;
  df=3;
  p=1-probchi(lrt,df);
run;
proc print noobs;
run;

```

Descr	logl_0	logl	lrt	df	p
-2 Log Likelihood	2948.2	2941.9	6.32293	3	0.096914

The insignificant result means that for the item SC02 there is no evidence of item parameter drift. Similar tests should be carried out for the other items in the scale.

5.2. Local dependence across time points

The macro %lrasch_mml makes it possible to split items and to test the assumption local independence across time points we thus specify

```

%lrasch_mml(DATA=data.scqol, ITEM_NAMES=inames,
  ANCHOR=b_SC02 b_SC04 b_SC05 b_SC07 b_SC08 b_SC09,
  SPLIT=b_SC02, OUT=SP2);

```

where the item SC02 is split. The log likelihood value is $-2 \log L = 2933.1$.

SP2: LRASCH MML estimation loglikelihood

Descr	Value
-2 Log Likelihood	2933.1
AIC (smaller is better)	3031.1
AICC (smaller is better)	3034.1
BIC (smaller is better)	3159.2

As described above splitting means that four versions of SC02 are included at time 2 depending on the observed value of SC02 at time 1. These consist of: an item with $\beta = (0.52, 1.47, 0.51)$ for respondents with $b_SC02 = 0$, an item with $\beta = (-0.70, 1.37, 2.64)$ for respondents with $b_SC02 = 1$, an item with $\beta = (-0.13, 0.42, 1.93)$ for respondents with $b_SC02 = 2$, and an item with $\beta = (-1.96, 1.21, 2.02)$ for respondents with $b_SC02 = 3$.

SP2: LRASCH MML estimated thresholds

Label	Estimate	Standard Error	Lower	Upper
-------	----------	----------------	-------	-------

b_SC01 1	0.84	0.28	0.27	1.40
b_SC01 2	2.31	0.44	1.44	3.17
b_SC02 1	-1.24	0.33	-1.91	-0.58
b_SC02 2	0.72	0.32	0.10	1.35
b_SC02 3	1.44	0.38	0.69	2.19
b_SC03 1	0.49	0.29	-0.09	1.06
b_SC03 2	1.55	0.37	0.81	2.29
b_SC03 3	3.17	0.68	1.81	4.53
b_SC04 1	-0.09	0.24	-0.56	0.38
b_SC04 2	1.47	0.28	0.91	2.02
b_SC04 3	2.29	0.41	1.48	3.10
b_SC05 1	3.02	0.34	2.33	3.70
b_SC05 2	1.74	0.52	0.72	2.77
b_SC05 3	2.12	0.64	0.85	3.38
b_SC06 1	1.10	0.29	0.52	1.67
b_SC06 2	2.66	0.50	1.66	3.66
b_SC06 3	3.67	1.13	1.42	5.92
b_SC07 1	-0.76	0.24	-1.23	-0.28
b_SC07 2	1.94	0.28	1.37	2.50
b_SC07 3	2.34	0.44	1.47	3.22
b_SC08 1	1.59	0.25	1.09	2.09
b_SC08 2	2.36	0.41	1.55	3.16
b_SC08 3	2.61	0.64	1.34	3.89
b_SC09 1	-0.67	0.25	-1.18	-0.17
b_SC09 2	0.61	0.25	0.11	1.11
b_SC09 3	1.96	0.32	1.32	2.60
fu_SC02 1 (b_SC02 = 0)	0.52	0.66	-0.79	1.82
fu_SC02 2 (b_SC02 = 0)	1.47	1.16	-0.84	3.78
fu_SC02 3 (b_SC02 = 0)	0.51	1.46	-2.39	3.40
fu_SC02 1 (b_SC02 = 1)	-0.70	0.48	-1.66	0.25
fu_SC02 2 (b_SC02 = 1)	1.37	0.53	0.31	2.43
fu_SC02 3 (b_SC02 = 1)	2.64	1.11	0.43	4.85
fu_SC02 1 (b_SC02 = 2)	-0.13	0.67	-1.46	1.21
fu_SC02 2 (b_SC02 = 2)	0.42	0.62	-0.80	1.64
fu_SC02 3 (b_SC02 = 2)	1.93	0.75	0.45	3.42
fu_SC02 1 (b_SC02 = 3)	-1.96	1.14	-4.22	0.30
fu_SC02 2 (b_SC02 = 3)	1.21	0.70	-0.19	2.60
fu_SC02 3 (b_SC02 = 3)	2.02	0.94	0.15	3.88
fu_SC01 1	0.59	0.33	-0.06	1.24
fu_SC01 2	1.19	0.40	0.40	1.97
fu_SC01 3	2.45	0.62	1.23	3.68
fu_SC03 1	1.57	0.34	0.90	2.23
fu_SC03 2	2.54	0.60	1.34	3.74
fu_SC04 1	-0.09	0.24	-0.56	0.38
fu_SC04 2	1.47	0.28	0.91	2.02
fu_SC04 3	2.29	0.41	1.48	3.10

fu_SC05 1	3.02	0.34	2.33	3.70
fu_SC05 2	1.74	0.52	0.72	2.77
fu_SC05 3	2.12	0.64	0.85	3.38
fu_SC06 1	1.89	0.35	1.20	2.59
fu_SC06 2	3.14	0.80	1.56	4.72
fu_SC07 1	-0.76	0.24	-1.23	-0.28
fu_SC07 2	1.94	0.28	1.37	2.50
fu_SC07 3	2.34	0.44	1.47	3.22
fu_SC08 1	1.59	0.25	1.09	2.09
fu_SC08 2	2.36	0.41	1.55	3.16
fu_SC08 3	2.61	0.64	1.34	3.89
fu_SC09 1	-0.67	0.25	-1.18	-0.17
fu_SC09 2	0.61	0.25	0.11	1.11
fu_SC09 3	1.96	0.32	1.32	2.60

The population parameters are of the same magnitude as in the two previous models.

SP2: LRASCH estimated change in latent mean (μ), variances, and latent correlation (ρ)

Parameter	Estimate	Standard		
		Error	Lower	Upper
μ	0.092	0.164	-0.234	0.418
ρ	0.624	0.096	0.433	0.814
σ_1	1.347	0.142	1.064	1.629
σ_2	1.107	0.135	0.839	1.374

Compared to the model where β at time 2 is equal to β at time 1 this model has 12 additional parameters. This yields a likelihood ratio test statistic of $-2 \log L = 2948.2 - 2933.1 = 15.1$, which on 12 degrees of freedom is not significant. This likelihood ratio test can be computed using the SAS code

```
data lrt;
  merge SCQOL_logl(rename=(value=logl_0))
        SP2_logl(rename=(value=logl));
  where Descr='-2 Log Likelihood';
  lrt=logl_0-logl;
  df=12;
  p=1-probchi(lrt,df);
run;
proc print;
run;
```

Descr	logl_0	logl	lrt	df	p
-2 Log Likelihood	2948.2	2933.1	15.0844	12	0.23685

The insignificant result means that for the item `SC02` there is no evidence of local dependence across time points. The observed correlation is explained solely by the latent variables θ_1 and θ_2 . Similar tests should be carried out for the other items in the scale. This can be done in steps: by first testing local dependence across time points for each item against the simple longitudinal Rasch model, and then evaluating the combined evidence. The flexibility of the proposed macro then makes it possible to add local dependence across time for a single item yielding an extended model. Local dependence across time points for the remaining items can then be tested against this extended model.

6. Discussion

The proprietary software packages **RUMM** (Andrich *et al.* 2010) and **WINSTEPS** (Linacre 2011) for fitting Rasch models are widely used. All of these fit unidimensional models only, even though many applications deal with multidimensional or longitudinal data. The two-dimensional Rasch model as originally discussed by Andersen (1985) and Embretson (1991) was formulated for longitudinal data. To obtain consistent item parameter estimates MML estimation (Bock and Aitkin 1981; Thissen 1982; Zwinderman and Wollenberg 1990) is used. This approach to item parameter estimation assumes that the latent variables are sampled from a population and introduces an assumption about the distribution of the latent variable.

A weakness of the proposed macro is that fit indices, like INFIT or OUTFIT, are not implemented. However, since these fit test statistics are implemented in the SAS macros for unidimensional Rasch models **AnaQol** (Hardouin and Mesbah 2007), **GLIMMIX_Rasch** (Chen *et al.* 2013), and `%rasch_mml` (Christensen and Olsbjerg 2013) testing the fit of the Rasch model at each time point is feasible. While tests of item fit in the longitudinal model are not implemented, some of the features of the longitudinal model structure can be tested using SAS macros for unidimensional Rasch models. Indeed, because local dependence across time points can be modeled using item splitting (as implemented using the `SPLIT` option), a SAS macro for unidimensional Rasch models that can accommodate missing data can be used to test the fit of locally dependent items in a longitudinal Rasch model. Another weakness of the proposed macro is that the implemented graphical display only shows curves of the expected categories probabilities, but that no comparison with observed data is possible. Again SAS macros for unidimensional Rasch models like `%AnaQol` (Hardouin and Mesbah 2007), `%rasch_mml` (Christensen and Olsbjerg 2013), and `%rasch_cml` (Christensen 2013) can be used for this.

The general implementation allows the user to specify models where the item parameters do change over time. The macro can be used to test the assumption of item parameter invariance using likelihood ratio tests, thus adding to existing methods for detection of item parameter drift (Donoghue and Isham 1998; DeMars 2004; Galdin and Laurencelle 2010). The macro also makes it possible to study local dependence across time points, by splitting of the item at follow-up into new items according to the responses given at baseline (Olsbjerg and Christensen 2014). The macro `%lrasch_mml` makes it possible to include splitted items and to test the assumption of local independence across time points using likelihood ratio tests, thus adding to existing tests of this assumption (Olsbjerg and Christensen 2013).

References

- Adams R, Wilson M, Wang WC (1997a). “The Multidimensional Random Coefficients Multinomial Logit Model.” *Applied Psychological Measurement*, **21**(1), 1–23. doi:10.1177/0146621697211001.
- Adams R, Wilson M, Wu M (1997b). “Multilevel Item Response Models: An Approach to Errors in Variables Regression.” *Journal of Educational and Behavioral Statistics*, **22**(1), 47–76. doi:10.3102/10769986022001047.
- Akaike H (1974). “A New Look at the Statistical Model Identification.” *IEEE Transactions on Automatic Control*, **19**(6), 716–723. doi:10.1109/tac.1974.1100705.
- Andersen E (1977). “Sufficient Statistics and Latent Trait Models.” *Psychometrika*, **42**(1), 69–81. doi:10.1007/bf02293746.
- Andersen E (1985). “Estimating Latent Correlations between Repeated Testings.” *Psychometrika*, **50**(1), 3–16. doi:10.1007/bf02294143.
- Andrich D, Kreiner S (2010). “Quantifying Response Dependence Between Two Dichotomous Items Using the Rasch Model.” *Applied Psychological Measurement*, **34**(3), 181–192. doi:10.1177/0146621609360202.
- Andrich D, Sheridan B, Luo G (2010). **RUMM2030**. RUMM Laboratory, Perth, Australia.
- Blanchin M, Hardouin J, Neel T, Kubis G, Blanchard C, é E, Sébille V (2011). “Comparison of CTT and Rasch-Based Approaches for the Analysis of Longitudinal Patient Reported Outcomes.” *Statistics in Medicine*, **30**(8), 825–838. doi:10.1002/sim.4153.
- Bock R, Aitkin M (1981). “Marginal Maximum Likelihood Estimation of Item Parameters: An Application of an EM Algorithm.” *Psychometrika*, **46**(4), 443–459. doi:10.1007/bf02293801.
- Burnham K, Anderson D (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 2nd edition. Springer-Verlag, New York.
- Chan K, Drasgow F, Sawin L (1999). “What Is the Shelf Life of a Test? The Effect of Time on the Psychometrics of a Cognitive Ability Test Battery.” *Journal of Educational Measurement*, **84**(4), 610–619. doi:10.1037/0021-9010.84.4.610.
- Chen Y, Li I, Kromrey J (2013). “GLIMMIX_Rasch: A SAS Macro for Fitting the Dichotomous Rasch Model.” In *SouthEast SAS Users Group 2013 Proceedings*. SAS Institute Inc., Cary, NC. URL <http://analytics.ncsu.edu/sesug/2013/P0-10.pdf>.
- Christensen K (2006). “Fitting Polytomous Rasch Models in SAS.” *Journal of Applied Measurement*, **7**(4), 407–417.
- Christensen K (2013). “Conditional Maximum Likelihood Estimation in Polytomous Rasch Models Using SAS.” *ISRN Computational Mathematics*, **Article ID 617475**, 8 pages. doi:10.1155/2013/617475.

- Christensen K, Bjorner J (2003). “SAS Macros for Rasch Based Latent Variable Modelling.” *Technical report*, Department of Biostatistics, University of Copenhagen. Research report 03/13.
- Christensen K, Kreiner S, Mesbah M (2013). *Rasch Models in Health*. John Wiley & Sons, Hoboken, NJ. doi:10.1002/9781118574454.
- Christensen K, Olsbjerg M (2013). “Marginal Maximum Likelihood Estimation in Polytomous Rasch Models Using SAS.” *Publications de l’Institut de statistique de l’Université de Paris*, 57(1–2), 69–84.
- DeMars C (2004). “Detection of Item Parameter Drift over Multiple Test Administrations.” *Applied Measurement in Education*, 17(3), 265–300. doi:10.1207/s15324818ame1703_3.
- Donoghue J, Isham S (1998). “A Comparison of Procedures to Detect Item Parameter Drift.” *Applied Psychological Measurement*, 22(1), 33–51. doi:10.1177/01466216980221002.
- Embretson S (1991). “A Multidimensional Latent Trait Model for Measuring Learning and Change.” *Psychometrika*, 56(3), 495–516. doi:10.1007/bf02294487.
- Fischer G, Molenaar I (1995). *Rasch Models: Foundations, Recent Developments, and Applications*. Springer-Verlag, New York.
- Galdin M, Laurencelle L (2010). “Assessing Parameter Invariance in Item Response Theory’s Logistic Two Item Parameter Model: A Monte Carlo Investigation.” *Tutorials in Quantitative Methods for Psychology*, 6(2), 39–51.
- Goldstein H (1983). “Measuring Changes in Educational Attainment over Time: Problems and Possibilities.” *Journal of Educational Measurement*, 20(4), 369–377. doi:10.1111/j.1745-3984.1983.tb00214.x.
- Hardouin J, Mesbah M (2007). “The SAS Macro-Program %AnaQo1 to Estimate the Parameters of Item Responses Theory Models.” *Communications in Statistics – Simulation and Computation*, 36(2), 437–453. doi:10.1080/03610910601158351.
- IBM Corporation (2015). *IBM SPSS Statistics 23*. IBM Corporation, Armonk, NY. URL <http://www.ibm.com/software/analytics/spss/>.
- Kelderman H (1984). “Loglinear Rasch Model Tests.” *Psychometrika*, 49(2), 223–245. doi:10.1007/bf02294174.
- Kreiner S (2003). “Introduction to **DIGRAM**.” *Research Report 10*, Department of Statistics, University of Copenhagen.
- Kreiner S, Christensen K (2004). “Analysis of Local Dependence and Multidimensionality in Graphical Loglinear Rasch Models.” *Communications in Statistics – Theory and Methods*, 33(6), 1239–1276. doi:10.1081/sta-120030148.
- Kreiner S, Christensen K (2007). “Validity and Objectivity in Health-Related Scales: Analysis by Graphical Loglinear Rasch Models.” In M Davier, C Carstensen (eds.), *Multivariate and Mixture Distribution Rasch Models: Extensions and Applications*. Springer-Verlag, New York.

- Linacre J (2011). **WINSTEPS Rasch Measurement Computer Program**. Beaverton, Oregon. URL <http://www.winsteps.com/>.
- Linden W, Hambleton R (1997). *Handbook of Modern Item Response Theory*. Springer-Verlag, New York. doi:10.1007/978-1-4757-2691-6.
- Marais I (2009). “Response Dependence and the Measurement of Change.” *Journal of Applied Measurement*, **10**(1), 17–29.
- Marais I, Andrich D (2008a). “Effects of Varying Magnitude and Patterns of Local Dependence in the Unidimensional Rasch Model.” *Journal of Applied Measurement*, **9**(2), 105–124.
- Marais I, Andrich D (2008b). “Formalising Dimension and Response Violations of Local Independence in the Unidimensional Rasch Model.” *Journal of Applied Measurement*, **9**(3), 200–215.
- Masters G (1982). “A Rasch Model for Partial Credit Scoring.” *Psychometrika*, **47**(2), 149–174. doi:10.1007/bf02296272.
- Miller G, Fitzpatrick S (2009). “Expected Equating Error Resulting From Incorrect Handling of Item Parameter Drift among the Common Items.” *Educational and Psychological Measurement*, **69**(3), 357–368. doi:10.1177/0013164408322033.
- Nandakumar R, Hotchkiss L (2012). “PROC NL MIXED: For Estimating Parameters of IRT Models.” *Applied Psychological Measurement*, **36**(1), 60–63. doi:10.1177/0146621611420419.
- Neyman J, Scott E (1948). “Consistent Estimates Based on Partially Consistent Observations.” *Econometrica*, **16**(1), 1–32. doi:10.2307/1914288.
- Olsbjerg M, Christensen K (2013). “Marginal and Conditional Approach to Longitudinal Rasch Models.” *Publications de l’Institut de statistique de l’Université de Paris*, **57**(1–2), 109–126.
- Olsbjerg M, Christensen K (2014). “Local Dependence in Longitudinal IRT Models.” *Research Report 2*, Department of Statistics, University of Copenhagen.
- Pastor D, Beretvas S (2006). “Longitudinal Rasch Modeling in the Context of Psychotherapy Outcomes Assessment.” *Applied Psychological Measurement*, **30**(2), 100–120. doi:10.1177/0146621605279761.
- Rasch G (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Danish National Institute for Educational Research, Copenhagen.
- Raudenbush S, Bryk A, Cheong Y, Congdon R (2004). **HLM 6 for Windows**. Scientific Software International, Inc., Skokie, IL.
- Reeve B, Hays R, JB B, Cook, F K, Crane P, Teresi J, Thissen D, Revicki D, Weiss D, Hambleton R, Liu H, Gershon R, Reise S, Lai J, Cella D, Group P (2007). “Psychometric Evaluation and Calibration of Health-Related Quality of Life Item Banks. Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS).” *Medical Care*, **45**(5). doi:10.1097/01.mlr.0000250483.85507.04. Suppl. 1.

- Rijmen F, De Boeck P (2002). “The Random Weights Linear Logistic Test Model.” *Applied Psychological Measurement*, **26**(3), 271–285. doi:10.1177/0146621602026003003.
- Rijmen F, Tuerlinckx F, De Boeck P, Kuppens P (2003). “A Non-Linear Mixed Model Framework for Item Response Theory.” *Psychological Methods*, **8**(2), 185–205. doi:10.1037/1082-989x.8.2.185.
- SAS Institute Inc (2013). *The SAS System, Version 9.4*. SAS Institute Inc., Cary, NC. URL <http://www.sas.com/>.
- Schwarz G (1978). “Estimating the Dimension of a Model.” *The Annals of Statistics*, **6**(2), 461–464. doi:10.1214/aos/1176344136.
- Sheu CF, Chen CT, Su YH, Wang WC (2005). “Using SAS PROC NL MIXED to Fit Item Response Theory Models.” *Behavior Research Methods*, **37**(2), 202–218. doi:10.3758/bf03192688.
- Smits D, De Boeck P, Verhelst ND (2003). “Estimation of the **MIRID**: A Program and a SAS-Based Approach.” *Behavior Research Methods, Instruments and Computers*, **35**(4), 537–549. doi:10.3758/bf03195533.
- Tennant A, Conaghan P (2007). “The Rasch Measurement Model in Rheumatology: What Is It and Why Use It? When Should It Be Applied, and What Should One Look for in a Rasch Paper?” *Arthritis Care & Research*, **57**(8), 1358–1362. doi:10.1002/art.23108.
- Thissen D (1982). “Marginal Maximum Likelihood Estimation for the One-Parameter Logistic Model.” *Psychometrika*, **47**(2), 175–186. doi:10.1007/bf02296273.
- Vinding G, Christensen K, Esmann S, Olesen A, Jemec G (2013). “Quality of Life in Non-Melanoma Skin Cancer—The Skin Cancer Quality of Life (SCQoL) Questionnaire.” *Dermatologic Surgery*, **39**(12), 1784–1793. doi:10.1111/dsu.12353.
- Wells C, Subkoviak M, Serlin R (2002). “The Effect of Item Parameter Drift on Examinee Ability Estimates.” *Applied Psychological Measurement*, **26**, 77–87. doi:10.1177/0146621602261005.
- Wu M, Adams R, Wilson M, Haldane S (2007). *ACER ConQuest Version 2: Generalised Item Response Modelling Software*. Australian Council for Educational Research, Camberwell.
- Zwinderman A, Wollenberg A (1990). “Robustness of Marginal Maximum Likelihood Estimation in the Rasch Model.” *Applied Psychological Measurement*, **14**(1), 73–81. doi:10.1177/014662169001400107.

Affiliation:

Karl Bang Christensen
Department of Biostatistics
University of Copenhagen
Øster Farimagsgade 5, entr. B
P.O. Box 2099 DK-1014 Copenhagen K, Denmark
E-mail: kach@sund.ku.dk
URL: <http://biostat.ku.dk/~kach/>