Item response theory models for measuring level and change in latent variables

KB Christensen
Dep. of Biostatistics, Univ. of Copenhagen
http://biostat.ku.dk/~kach

Item response theory (IRT) model

Latent variable θ measured indirectly through item responses.

```
              Karl                     Marco
------------------------------------------------------------->
2+2                  18+23                   log(40)-3
```

Statistical models place items and persons on the same scale

(i) evaluation of the measurement instrument

(ii) modeling of the structure of the latent variables.

IRT models: Rasch [1PL], Birnbaum [2PL]

Probabilities for item i

$$P(X_i = 1|\theta, \beta_i) = \frac{\text{exp}(\theta - \beta_i)}{1 + \text{exp}(\theta - \beta_i)} \qquad [1PL]$$

(depends on person location θ and the location of item i).

$$P(X_i = 1|\theta, \beta_i) = \frac{\text{exp}(\alpha_i(\theta - \beta_i))}{1 + \text{exp}(\alpha_i(\theta - \beta_i))} \qquad [2PL]$$

depends on the person location θ and the location and the scale of item i.

Item and person locations

An IRT model gives us estimates $\widehat{\beta}_i$ of the item locations (the item difficulties) and of each persons location $\widehat{\theta}$.

```
                     x      x
                     x      x      x
                   x x      x      xx
                   x x      x      xx   xx
        x          x x      x      xx   xx  x
--------------------------------------------------------------->
2+2                  18+23                 log(40)-3
```

Item information functions ...

... the accuracy with which we can estimate person locations.

Any item provides some information about this, but the amount of information depends on targeting.

Example - Rasch [1PL] model

$$I_i(\theta) = P_i(X_i = 1|\theta)(1 - P_i(X_i = 1|\theta))$$

More item $\sim$ more information. Test information function

$$I(\theta) = \sum_i I_i(\theta)$$

Precision

We use IRT models because we want to estimate the person location θ, or to *measure* θ, but how much measurement error is there.

Variance of the estimate $\widehat{\theta}$

$$V(\widehat{\theta}) = \frac{1}{I(\widehat{\theta})}$$

the standard error of measurement

$$SEM(\widehat{\theta}) = \sqrt{\frac{1}{I(\widehat{\theta})}}$$

is expressed in the same units as the measurement itself.

confidence interval around the estimate.

Stages in IRT analysis

(i) estimate item parameters
(ii) estimate person locations $\theta_1, \ldots, \theta_N$
(iii) evaluate model fit
(iv) evaluate targeting

Joint estimation of α's, $\beta's$, and θ's yield inconsistent estimates

Neyman, Scott. Econometrika, 1948, 16:1–32.

(i) estimate item parameters

IRT approach: assume $\theta \sim N(0, \sigma^2)$ together with local independence this yields the likelihood function

$$L(\alpha, \beta, \sigma^2) = \int \prod_{i=1}^{k} \frac{\exp(x_i \alpha_i(\theta - \beta_i))}{1 + \exp(\alpha_i(\theta - \beta_i))} \varphi(\theta) d\theta$$

maximized numerically. Estimates and asymptotic s.e.

If the distributional assumption is OK estimates are unbiased.

Rasch [1PL]: Param. est. without reference to dist. of θ

(ii) estimate person locations $\theta_1, \ldots, \theta_N$:

This is done using estimates of $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k)$. For a person with responses $(x_1, \ldots, x_k)$

$$L(\theta) = \prod_{i=1}^{k} \frac{\exp(x_i \widehat{\alpha}_i(\theta - \widehat{\beta}_i))}{1 + \exp(\widehat{\alpha}_i(\theta - \widehat{\beta}_i))}$$

easy to maximize. 'Asymptotic' s.e., biased estimates. Weighted likelihood

$$L(\theta) = \prod_{i=1}^{k} \frac{\exp(x_i \widehat{\alpha}_i(\theta - \widehat{\beta}_i))}{1 + \exp(\widehat{\alpha}_i(\theta - \widehat{\beta}_i))} g(\theta)$$

has smaller bias.

Warm. Psychometrika, 1989, 54:427–450.

Rasch Model

Items $X_1, \ldots, X_I$ measuring latent variable θ

$$P(X_i = 1|\theta) = \frac{\exp(\theta - \beta_i)}{1 + \exp(\theta - \beta_i)} = \frac{\xi\delta_i}{1 + \xi\delta_i}. \tag{1}$$

Resp. vector $\overline{X} = (X_1, \ldots, X_I)$, $\overline{x} = (x_1, \ldots, x_I)$

$$P(\overline{X} = \overline{x}|\theta) = \frac{\xi^r \prod_{i=1}^I \delta_i^{x_i}}{\prod_{i=1}^I (1 + \xi\delta_i)} \tag{2}$$

Score $R = \sum_{i=1}^I X_i$, $r = \sum_{i=1}^I x_i$.

$$P(R = r|\theta) = \frac{\xi^t \gamma(r)}{\prod_{i=1}^I (1 + \xi\delta_i)} = \frac{\xi^r \gamma(r)}{\sum_{l=0}^I \xi^l \gamma(l)} \tag{3}$$

where

$$\gamma(r) = \sum_{\overline{x}:\sum_{i=1}^I x_i = r} \prod \delta_i^{x'_i}. \tag{4}$$

Conditional likelihood

based on

$$P(\overline{X} = \overline{x} | R = r, \theta) = \frac{\prod_{i=1}^{I} \delta_i^{x_i}}{\gamma(r)} \tag{5}$$

Consistent item parameter estimates

Contribution to log likelihood

$$l_C(\beta) = \log P(\overline{X} = \overline{x} | R = r, \theta) = \sum_{i=1}^{I} x_i \log(\delta_i) - \log \gamma(r) \tag{6}$$

Implem. in R (eRm) Journal of Statistical Software, 20(9), 1-20
and SAS (`%rasch_cml.sas`)biostat.ku.dk/~kach/index.html♯cml

Extended likelihood

based on

$$P(\overline{X} = \overline{x}) = P(\overline{X} = \overline{x} | R = r) \underbrace{P(R = r | \theta)}_{\kappa(r)} \tag{7}$$

where $P(R = \quad \cdot \; | \theta)$ are used as unrestricted parameters $\kappa(r)$, estimated using

$$\widehat{\kappa(r)} = \frac{n(r)}{n} \tag{8}$$

where $n(r)$ is the number of persons with score r.

Contribution to log likelihood

$$l_E(\beta, \kappa) = \log P(\overline{X} = \overline{x}) = l_C(\beta) + \log \kappa(r) \tag{9}$$

Extended likelihood, ctd.

Many parameters $(\beta_i)_{i=1,\ldots,I}$ and $(\kappa(t))_{t=0,1,\ldots,I}$.
Score value r' is not obs.: $\widehat{\kappa(r')} = 0$ on boundary of param. space.

Simulate distribution of test statistics using two-step procedure:

(i) sample R_0 from the empirical dist. $(\widehat{\kappa}(r))_{r=0,1,\ldots,I}$
(ii) sample $\overline{X}$ from $P(\overline{X} = \cdot \,|R = R_0)$ *

Can generate, say, 1000 data sets and generate distribution of test statistic under the null.

*recursively: $P(X_1 = \cdot \,|R = R_0)$, $P(R_2 = \cdot \,| \sum_{i=1}^{I} X_i = R_0 - X_1)$, ...)

Item fit

Compare item responses X_i to expected values.

$$\theta \mapsto E(X_i|\theta) = P(X_i = 1|\theta) \tag{10}$$

Problem: θ is unobs.

Use $\widehat{\theta}$ or R

$$\widehat{\theta} \mapsto E(X_i|\widehat{\theta}) \tag{11}$$

$$t \mapsto E(X_i = h|R = r) \tag{12}$$

Item fit, ctd.

Unified approach: plot empirical curves

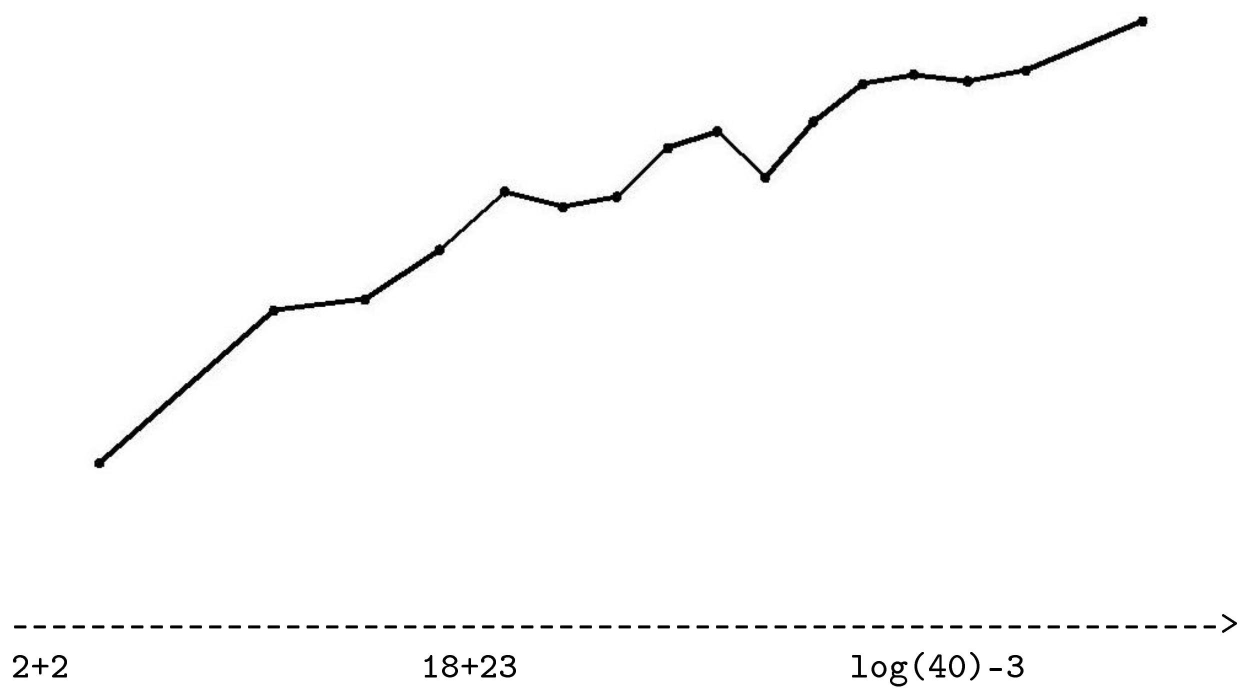
$$r \mapsto \frac{\sharp\{v|X_{iv}=1, R_v=r\}}{\sharp\{v|R_v=r\}}$$

or

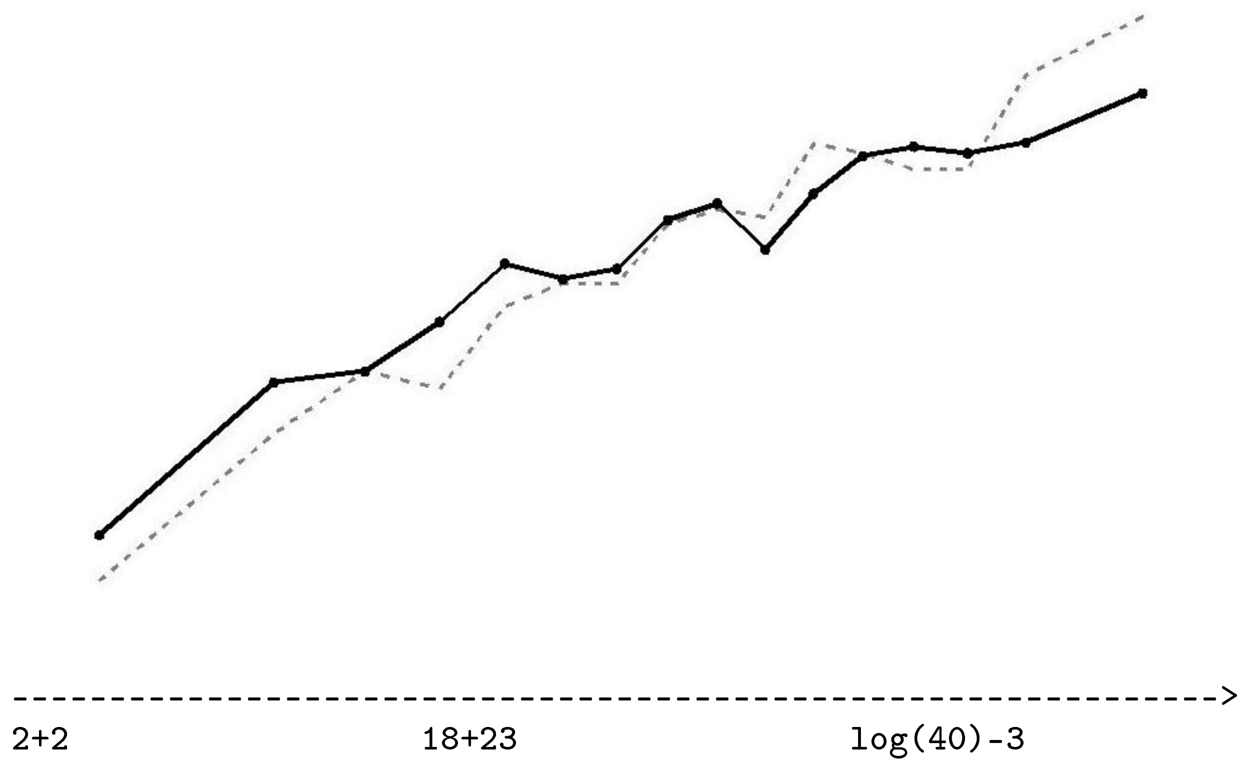
$$\widehat{\theta} \mapsto \frac{\sharp\{v|X_{iv}=1, \widehat{\theta}_v=\widehat{\theta}\}}{\sharp\{v|\widehat{\theta}_v=\widehat{\theta}\}}$$

Simulate data under the model. Compare the observed values to what we would expect if the model fitted the data.

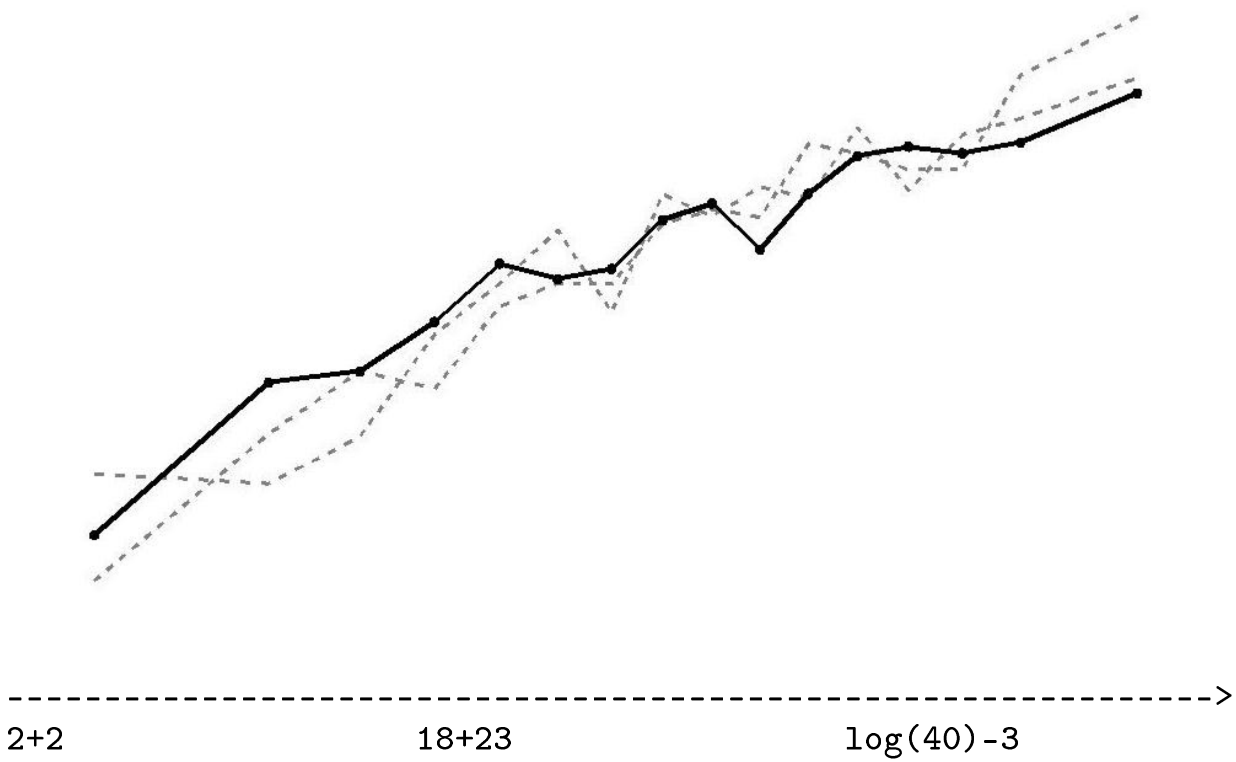
Item fit plot: Emiprical curves



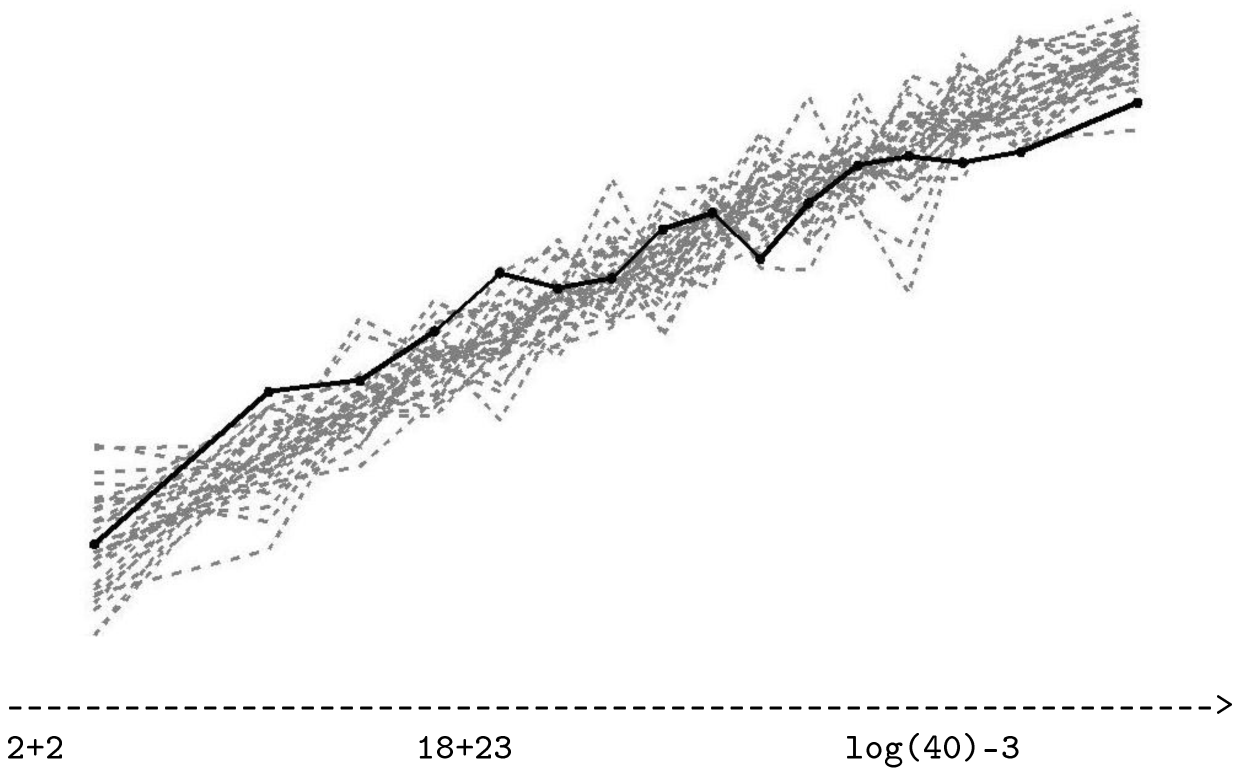
Item fit plot: Empirical curves with a single simulated curve



Item fit plot: Empirical curves with two simulated curves



Item fit plot: Observed and expected



Fit test example: Martin-Löf test

likelihood test based on two.dim. extended likel. function

$$I_1, I_2 \subset \{1, \ldots, I\},\ I_1 \cup I_2 = \emptyset.$$

Twodim. extended likel.

$$P(\overline{X}^{(1)} = \overline{x}^{(1)}, \overline{X}^{(2)} = \overline{x}^{(2)}) =$$

$$\prod_{d=1}^{2} P(\overline{X}^{(d)} = \overline{x}^{(d)} | R^{(d)} = r^{(d)}) \underbrace{P(R^{(1)} = r^{(1)}, R^{(2)} = t^{(2)})}_{\nu(r^{(1)}, r^{(2)})} \quad (13)$$

contribution to log likelihood

$$l_E(\beta, \nu) = L_C(\beta^{(1)}) + l_C(\beta^{(2)}) + \log(\nu(r^{(1)}, r^{(2)}) \quad (14)$$

Martin-Löf test

Under $H_0 : \theta_1 = \theta_2$ the contribution to the log likelihood is

$$l = l_C(\beta) + \log \kappa(r^{(1)} + r^{(2)}) \quad (15)$$

likelihood ratio test statistic

$$-2 \sum_{v=1}^{n} \left[l_C(\beta) - l_C(\beta^{(1)}) - l_C(\beta^{(2)}) + \log(\kappa(r)) - \log(\nu(r^{(1)}, r^{(2)})) \right]$$

Martin-Löf test, ctd.

Model: CML est. $(\widehat{\beta}_i)_{i=1,\ldots,I_1}$ and $(\widehat{\beta}_i)_{i=1,\ldots,I_1}$ (SAS, R) separately for the two dimensions. Estimate all ν's

Hypothesis: CML est. $(\widehat{\beta}_i)_{i=1,\ldots,I}$ (SAS, R). Estimate all κ's

Compute likelihood ratio test statistic. Simulate under hypothesis to evaluate significance.

Christensen, Kreiner. Appl Psych Meas, 2007,31:20-30.

Longitudinal Rasch models

Longitudinal data
Individual/group level effects
Assumptions/requirements
Attempts at modeling

Longitudinal Rasch models

Latent outcome variable.
Measurement part.
Longitudinal data.
Structural part

Longitudinal Rasch models

Latent outcome variable θ.
Measurement part.
Longitudinal data.
Structural part

Latent outcome variable θ.

Measurement part $Pr(X_i = 1|\theta)$, for $i = 1, \ldots, I$.

Longitudinal data.

Structural part

Latent outcome variable θ.
Measurement part $Pr(X_i = 1|\theta)$, for $i = 1, \ldots, I$.
Longitudinal data $\overline{X}_t = (X_{1,t}, \ldots, X_{I,t})$, for $t = 1, 2, ...$
Structural part

Latent outcome variable θ.
Measurement part $Pr(X_i = 1|\theta)$, for $i = 1, \ldots, I$.
Longitudinal data $\overline{X}_t = (X_{1,t}, \ldots, X_{I,t})$, for $t = 1, 2, ...$
Structural part $\theta = ..$

Longitudinal data

Items $\overline{X}_t = (X_{1,t}, \ldots, X_{I,t})$ at time points $t = 1, \ldots, T$.

$$Pr(X_{i,t} = 1|\theta_1, \ldots, \theta_t) = ...$$

Identifiability.
Test invariance in β's *
Structure: on θ's and/or on β's
Individual level or group level analysis ?

*cannot measure change if measurement instrument changes.

Data $\{\overline{X}_1|\overline{X}_2|...\}$

```
T=1                     T=2
id=1 101100010          id=1 1101100*0
id=2 001011100          id=2 111011100
id=3 10110***0          id=3 10000*1*1
```

Data $\{\overline{X}_1|\overline{X}_2|...\}$

```
T=1                  T=2
id=1 101100010       id=1 1101100*0
id=2 001011100       id=2 *********
id=3 10110***0       id=3 10000*1*1
```

Data

$$\left\{ \begin{array}{l} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \end{array} \right\}$$

```
T=1 id=1 101100010
    id=2 001011100
    id=3 10110***0
      :
T=2 id=1 101100010
    id=2 001011100
    id=3 10110***0
```

Data

$$\left\{ \begin{array}{l} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \end{array} \right\}$$

```
T=1 id=1 101100010
    id=2 001011100
    id=3 10110***0
      :
T=2 id=1 101100010
    id=2 *********
    id=3 10110***0
```

Linear logistic test model (Fischer, 1973)

Data $\{\underline{X}_1|\underline{X}_2|...\}$. Same items at $t = 1, 2$.

$$Pr(X_{i,t} = 1|\theta) = \frac{\exp(\theta - \beta_{i,t})}{1 + \exp(\theta - \beta_{i,t})} \tag{16}$$

Interpret change in β's as change in θ

$$(\theta - \beta_{i,1}) - (\theta - \beta_{i,2}) = (\theta^* - \beta_i) - (\theta - \beta_i) = (\theta^* - \theta)$$

conditional estimation $L_C(\beta; x) = P(\underline{X} = \underline{x}|R = r)$.
marginal estimation $L_M(\beta; x) = \int P(\underline{X} = \underline{x}|\theta)\varphi(\theta)d\theta$.

Not a model for change at individual level.

Multidim. IRT*

Data $\left\{ \begin{array}{c} \underline{X}_1 \\ \underline{X}_2 \\ \vdots \end{array} \right\}$ Andersen: $(\theta_1, \theta_2) \sim N(0, \Sigma)$, $\Sigma = \begin{pmatrix} 1 & \rho \\ & 1 \end{pmatrix}$

$$Pr(X_{i,t} = 1|\theta) = \frac{\exp(\theta_t - \beta_i)}{1 + \exp(\theta_t - \beta_i)} \tag{17}$$

Embretson: $(\theta_1, \theta_1 + \theta_{CHANGE})$

$$Pr(X_{i,t} = 1|\theta) = \frac{\exp(\sum_{h=1}^{t} \theta_h - \beta_i)}{1 + \exp(\sum_{h=1}^{t} \theta_h - \beta_i)} \tag{18}$$

Change at individual level.

*Andersen. Psychometrika 1985, 50:3-16.
Embretson. Psychometrika 1991, 56:495-516.

Multidim. IRT (Adams et al, 1997)

Data $\left\{ \begin{array}{l} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \end{array} \right\}$ generalization of (17)

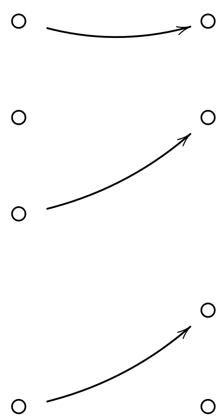
$$Pr(X_{i,t} = x_{i,t} | \underline{\theta}, \beta_i) = \frac{\exp(x_{i,t}\theta_t - \beta^{(t)}_{i,x,})}{1 + \exp(\sum_{l=1}^{k} \ldots)} \quad (19)$$

Abilities $\overline{\theta} = (\theta_1, \ldots, \theta_T)$
Item parameters $\overline{\beta}_i = (\beta_{i,t})_{t=1,\ldots,T}$

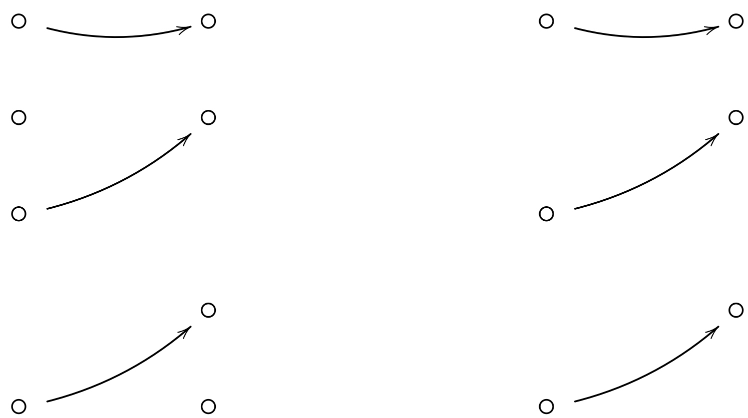
Change at individual level.

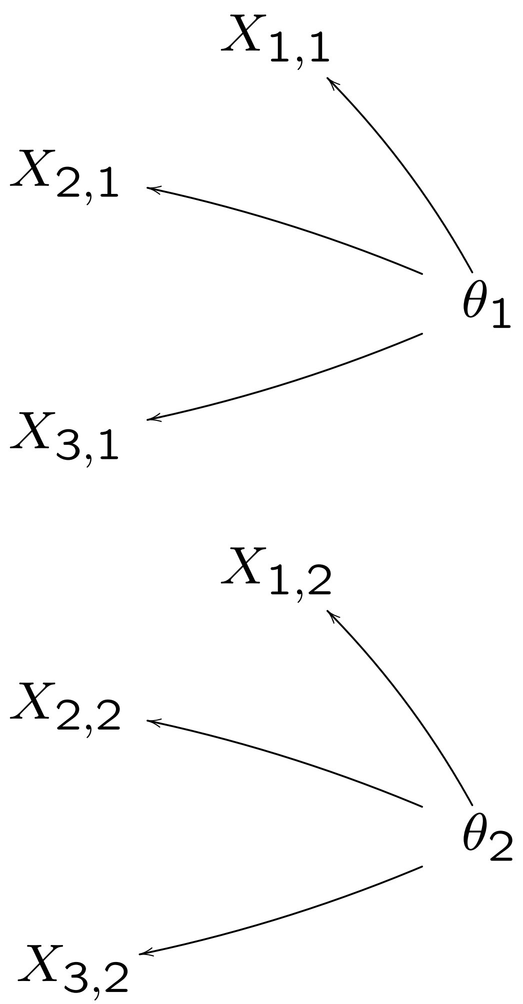
Adams, Wilson, Wang. Appl Psych Meas 1997, 21:1-23.

Open cohort

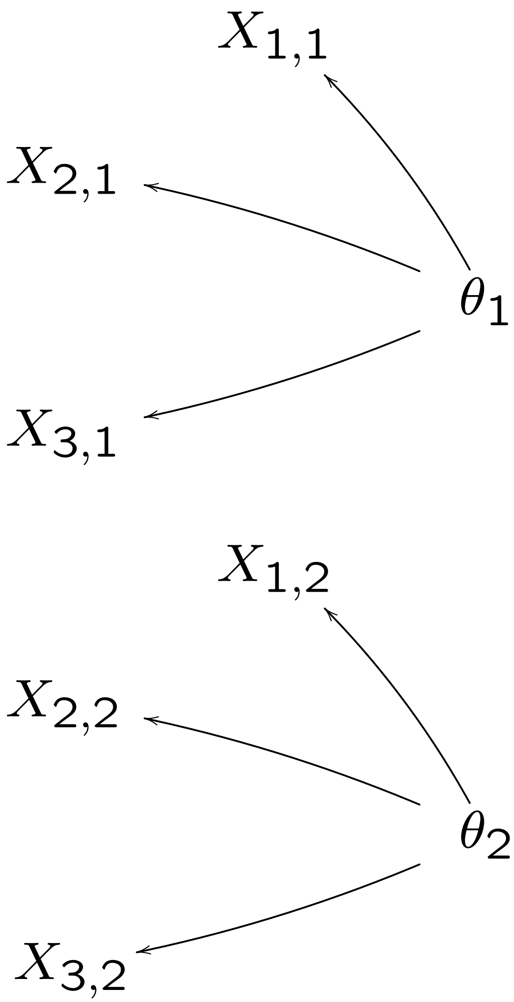


Open cohort / repeated measures only





$X_{1,1}$
$X_{2,1}$
θ_1
$X_{3,1}$
$X_{1,2}$
$X_{2,2}$
θ_2
$X_{3,2}$



INVARIANCE

ACROSS

TIME

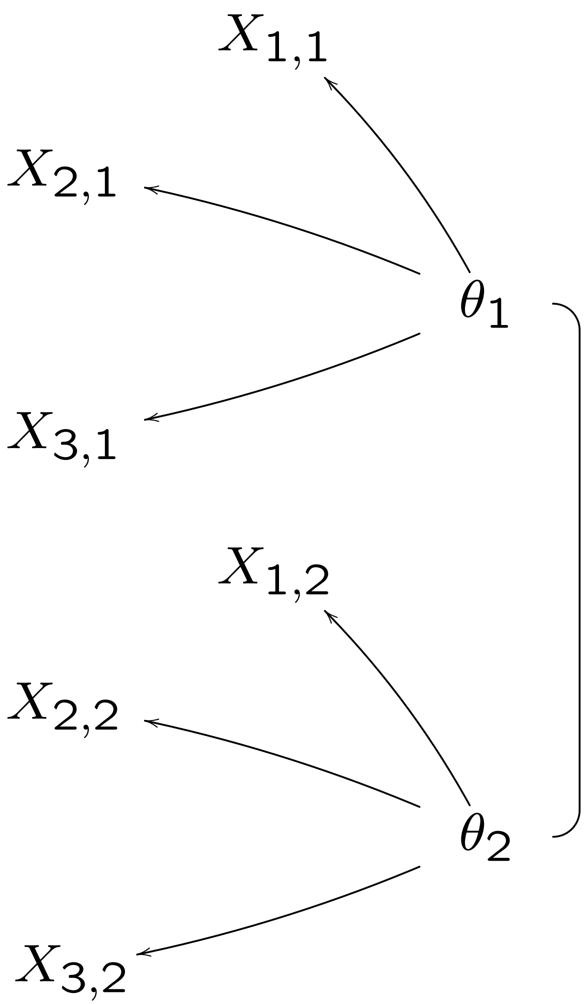
POINTS

test that items are invariant over time

	time 1			time 2		
id	it1	it2	it3	it1	it2	it3
1	:	:	:	*	*	*
2	:	:	:	*	*	*
:	:	:	:	*	*	*

MML: add interaction parameter(s) to 2-dim. model
CML: split, cross over, and test for DIF

MML: add interaction parameter(s) to 2-dim. model



randomly split subjects into two groups

```
                time 1               time 2
g       id      it1 it2 it3          it1 it2 it3
1       .        :   :   :            *   *   *
1       .        :   :   :            *   *   *
1       .        :   :   :            *   *   *
-       ------------------------------------
2       .        :   :   :            *   *   *
2       .        :   :   :            *   *   *
2       .        :   :   :            *   *   *
```

randomly split subjects into two groups

```
                time 1           time 2
g       id      it1 it2 it3      it1 it2 it3
1       .        :   :   :        *   *   *
1       .        :   :   :        *   *   *
1       .        :   :   :        *   *   *
-       ------------------------------------
2       .        :   :   :        *   *   *
2       .        :   :   :        *   *   *
2       .        :   :   :        *   *   *
```

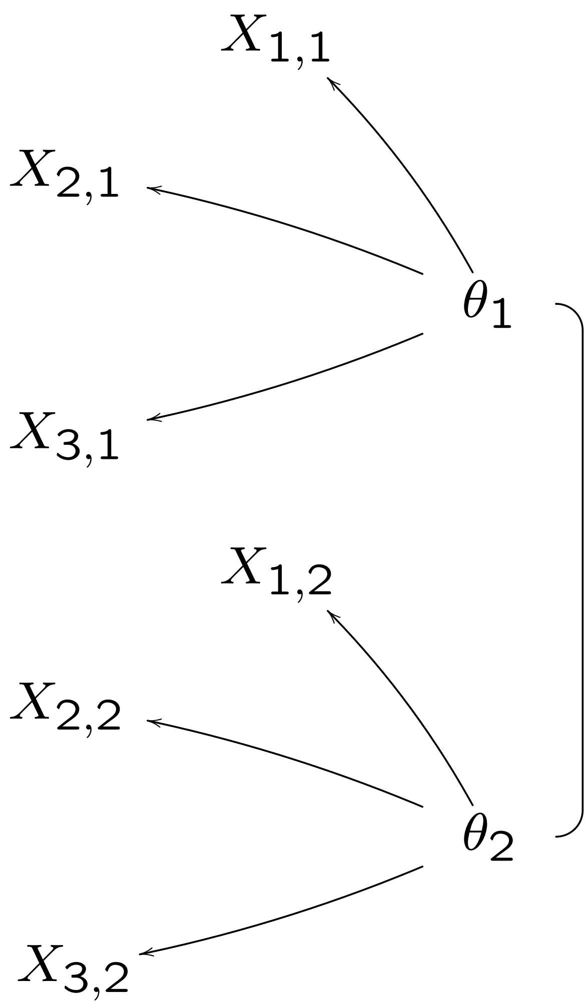
cross over

```
g       id      it1 it2 it3      it1 it2 it3
1       .        :   :   :        *   *   *
1       .        :   :   :        *   *   *
1       .        :   :   :        *   *   *
-       ------------------------------------
2       .        *   *   *        :   :   :
2       .        *   *   *        :   :   :
2       .        *   *   *        :   :   :
```

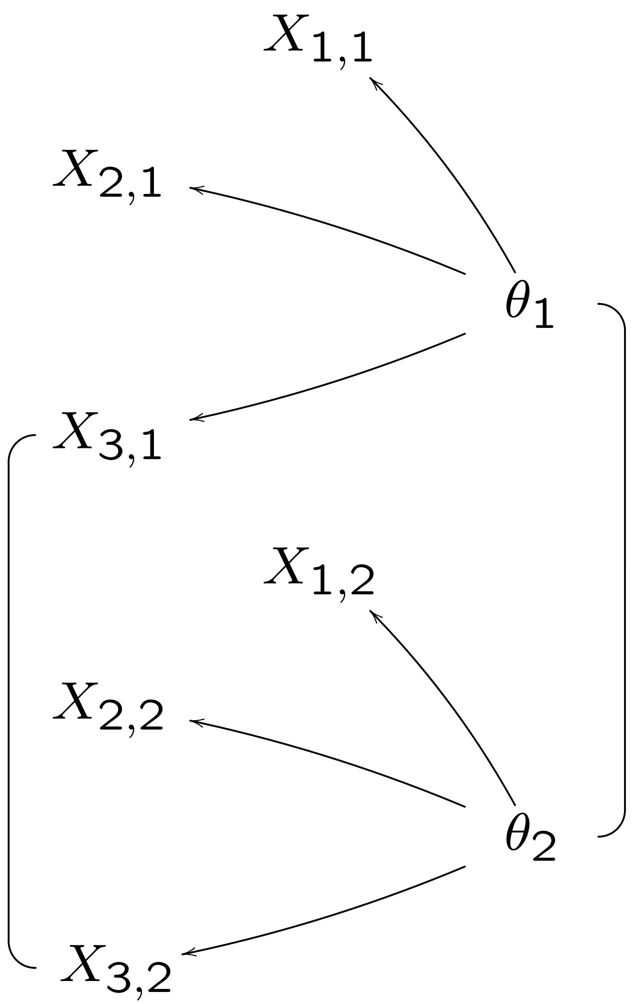
test DIF wrt. g in each item set (pool analyses to gain power)

```
g       id      it1 it2 it3
1       .        :   :   :
1       .        :   :   :
1       .        :   :   :
-       --------------------------
2       .        *   *   *
2       .        *   *   *
2       .        *   *   *
```

```
g       id                       it1 it2 it3
1       .                         *   *   *
1       .                         *   *   *
1       .                         *   *   *
-       ------------------------------------
2       .                         :   :   :
2       .                         :   :   :
2       .                         :   :   :
```



$X_{1,1}$
$X_{2,1}$
θ_1
$X_{3,1}$
$X_{1,2}$
$X_{2,2}$
θ_2
$X_{3,2}$



LD

$ACROSS$

$TIME$

$POINTS$

test LD across time

MML: add interaction parameter(s) to 2-dim. model

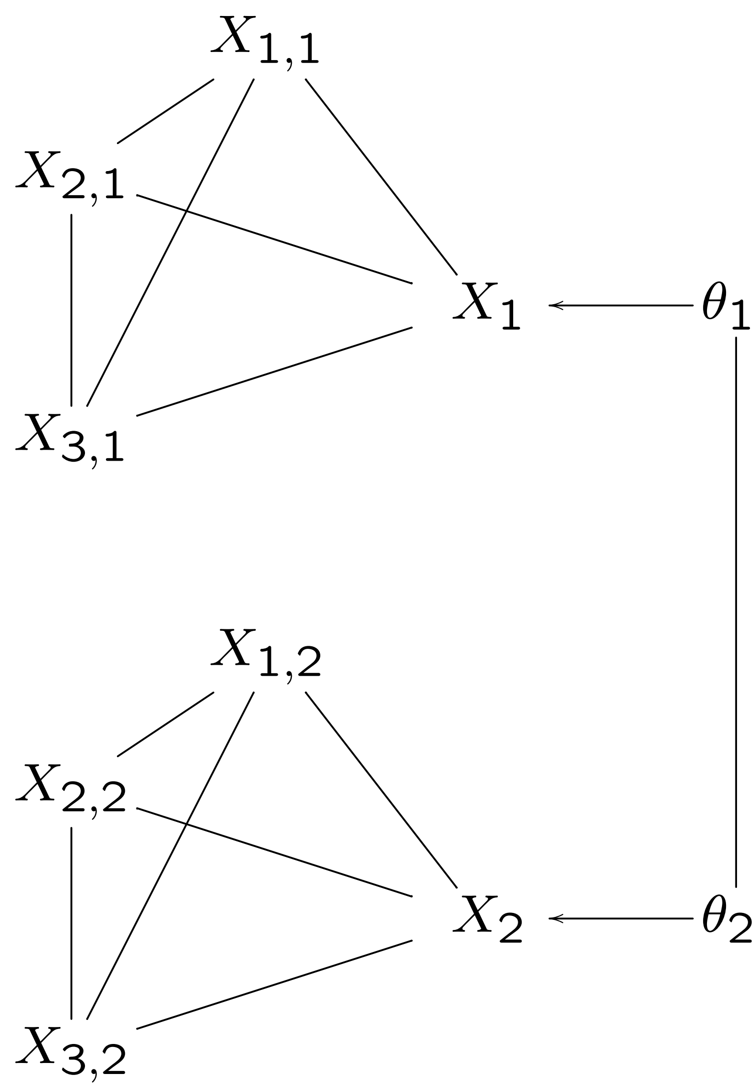
CML: look at the graphical Rasch model. Test hypotheses

$$X_{i,1} \perp X_{i,2} | R_1$$

and

$$X_{i,1} \perp X_{i,2} | R_2$$

Graphical Rasch model



LD across time points can be overcome

If items

```
t=1                          t=2
X1 X2 X3 X4 X5 X6 X7 X8      X1 X2 X3 X4 X5 X6 X7 X8
```

fit a Rasch model at each time point

LD across time points can overcome

If items

```
t=1                              t=2
X1 X2 X3 X4 X5 X6 X7 X8          X1 X2 X3 X4 X5 X6 X7 X8
```

fit a Rasch model at each time point: an incomplete design

```
t=1                              t=2
X1    X3    X5    X7                X2    X4    X6   X8
```

yields estimates of θ that are not confounded by LD.