Random Item Effects Modeling (Part 1)

Jean-Paul Fox

University of Twente
Department of Behavioural Science
Enschede, Netherlands

April 27, 2011

Random Item Effects Modeling

Measurement Invariance

Random Item Effects Approach

Identification

MCMC

Testing Measurement Invariance

Measurement Invariance: Longitudinal Data

Longitudinal Invariance

Measurement Invariance

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)

Measurement Invariance

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items

Measurement Invariance

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items
  2. Dependence on the score distribution

Measurement Invariance

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items
  2. Dependence on the score distribution
- Conditional methods (Scheuneman, 1979; Lord, 1980; Mellenbergh, 1982)

Measurement Invariance

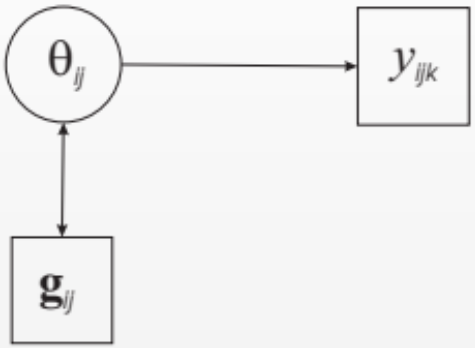
DIF:

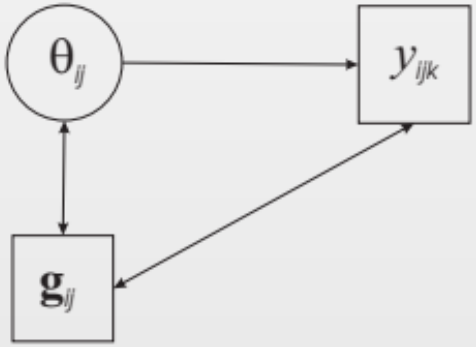
- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items
  2. Dependence on the score distribution
- Conditional methods (Scheuneman, 1979; Lord, 1980; Mellenbergh, 1982)

$$P(Y = 1 \mid \theta, g) \quad = \quad P(Y = 1 \mid \theta)$$

for all g, θ

Graphical display of DIF





Conditional DIF Detection

1. Specify item response model (likelihood model)

Conditional DIF Detection

1. Specify item response model (likelihood model)
2. Specify distribution of the latent variable θ

Conditional DIF Detection

1. Specify item response model (likelihood model)
2. Specify distribution of the latent variable θ
3. Estimate the ICCs in each group (θ not on the same scale)

Conditional DIF Detection

1. Specify item response model (likelihood model)
2. Specify distribution of the latent variable θ
3. Estimate the ICCs in each group (θ not on the same scale)
4. Assess the degree of bias (Mellenbergh, Hambleton)

Conditional DIF Detection

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group

Conditional DIF Detection

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group
2. Different population distributions, same item curves, simultaneous analysis

Conditional DIF Detection

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group
2. Different population distributions, same item curves, simultaneous analysis
3. Full measurement invariance versus partial measurement invariance

Conditional DIF Detection

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group
2. Different population distributions, same item curves, simultaneous analysis
3. Full measurement invariance versus partial measurement invariance
4. Inspection of item parameters for DIF detection

Many groups, ...

1. Specifying the population distribution (number of parameters)

Many groups, ...

1. Specifying the population distribution (number of parameters)
2. Finding anchor items

Many groups, ...

1. Specifying the population distribution (number of parameters)
2. Finding anchor items
3. DIF in difficulty and discrimination parameters. (Tests for partial measurement invariance)

Random Item Parameters

Group-Specific Items

$$P(Y_{ijk} = 1 \mid \theta_{ij}, \tilde{a}_{kj}, \tilde{b}_{kj}) = \begin{cases} \frac{\exp(d(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}))}{1+\exp(d(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}))} \\ \Phi(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}) \end{cases}$$

National and International Item Parameters

National

$$\tilde{\boldsymbol{\xi}}_{kj} = \left(\tilde{a}_{kj}, \tilde{b}_{kj}\right)^t \sim \mathcal{N}\left((a_k, b_k)^t, \boldsymbol{\Sigma}_{\tilde{\xi}}\right)$$

National and International Item Parameters

National

$$\tilde{\boldsymbol{\xi}}_{kj} = \left(\tilde{a}_{kj}, \tilde{b}_{kj}\right)^t \sim \mathcal{N}\left((a_k, b_k)^t, \boldsymbol{\Sigma}_{\tilde{\xi}}\right)$$

International

$$(a_k, b_k)^t \sim \mathcal{N}\left(\boldsymbol{\mu}_{\xi}, \boldsymbol{\Sigma}_{\xi}\right)$$

Hyper Prior

$$
\begin{aligned}
\boldsymbol{\Sigma}_\xi &\sim \mathcal{IW}(\nu, \boldsymbol{\Sigma}_0) \\
\boldsymbol{\mu}_\xi \mid \boldsymbol{\Sigma}_\xi &\sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_\xi / K_0)
\end{aligned}
$$

Hyper Prior

Hyper Prior

$$
\begin{aligned}
\boldsymbol{\Sigma}_\xi &\sim \mathcal{IW}(\nu, \boldsymbol{\Sigma}_0) \\
\boldsymbol{\mu}_\xi \mid \boldsymbol{\Sigma}_\xi &\sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_\xi / K_0)
\end{aligned}
$$

Correlation Structure

$$
Cov\left(\tilde{\boldsymbol{\xi}}_{kj}, \tilde{\boldsymbol{\xi}}_{k'j'}\right) = \begin{cases} \boldsymbol{\Sigma}_\xi + \boldsymbol{\Sigma}_{\tilde{\xi}} & k = k', j = j' \\ \boldsymbol{\Sigma}_\xi & k = k', j \neq j' \\ 0 & k \neq k'. \end{cases}
$$

Population Model

Independent Sampling Design

$$\theta_i \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$$

Population Model

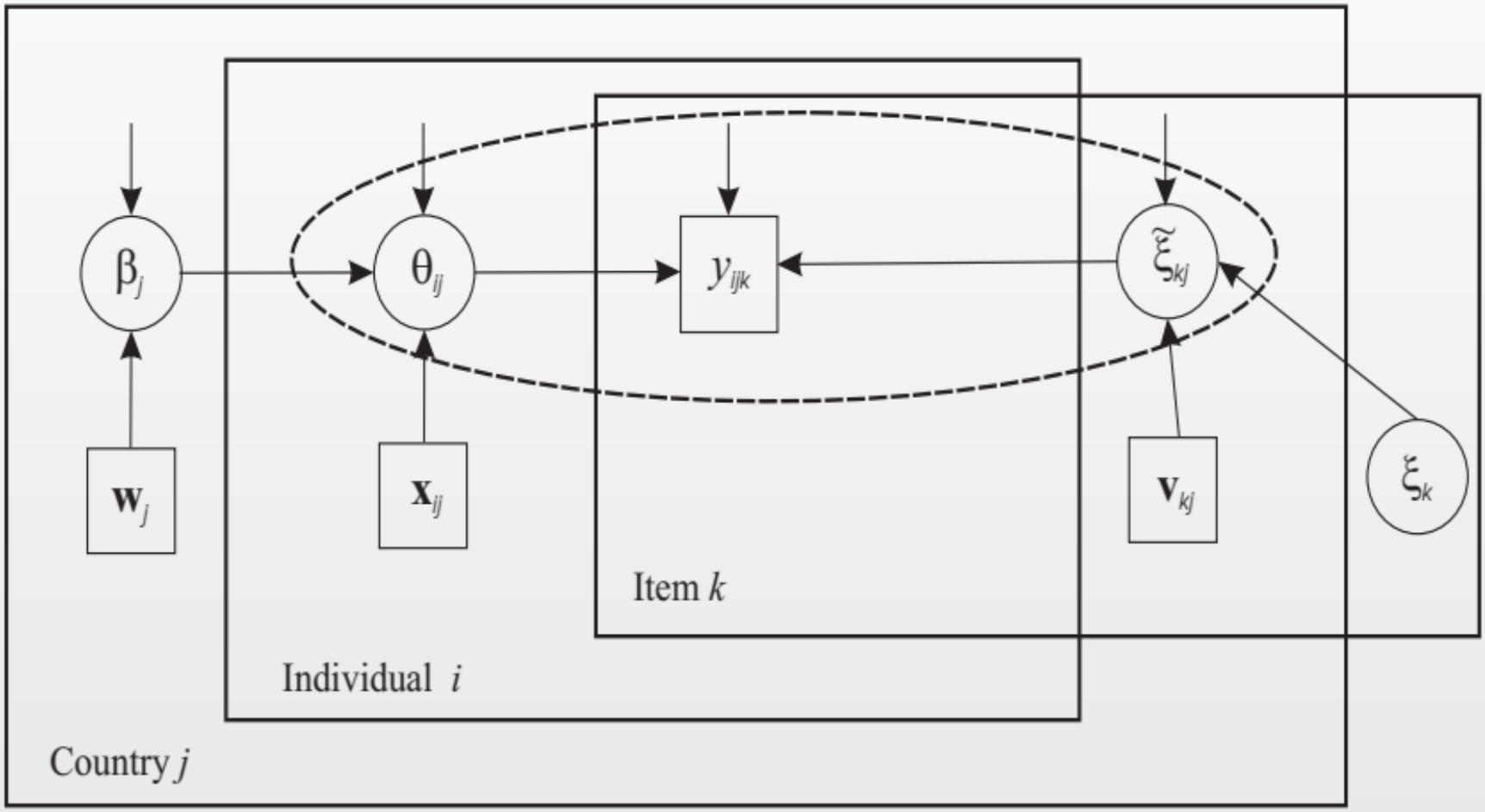
Independent Sampling Design

$$\theta_i \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$$

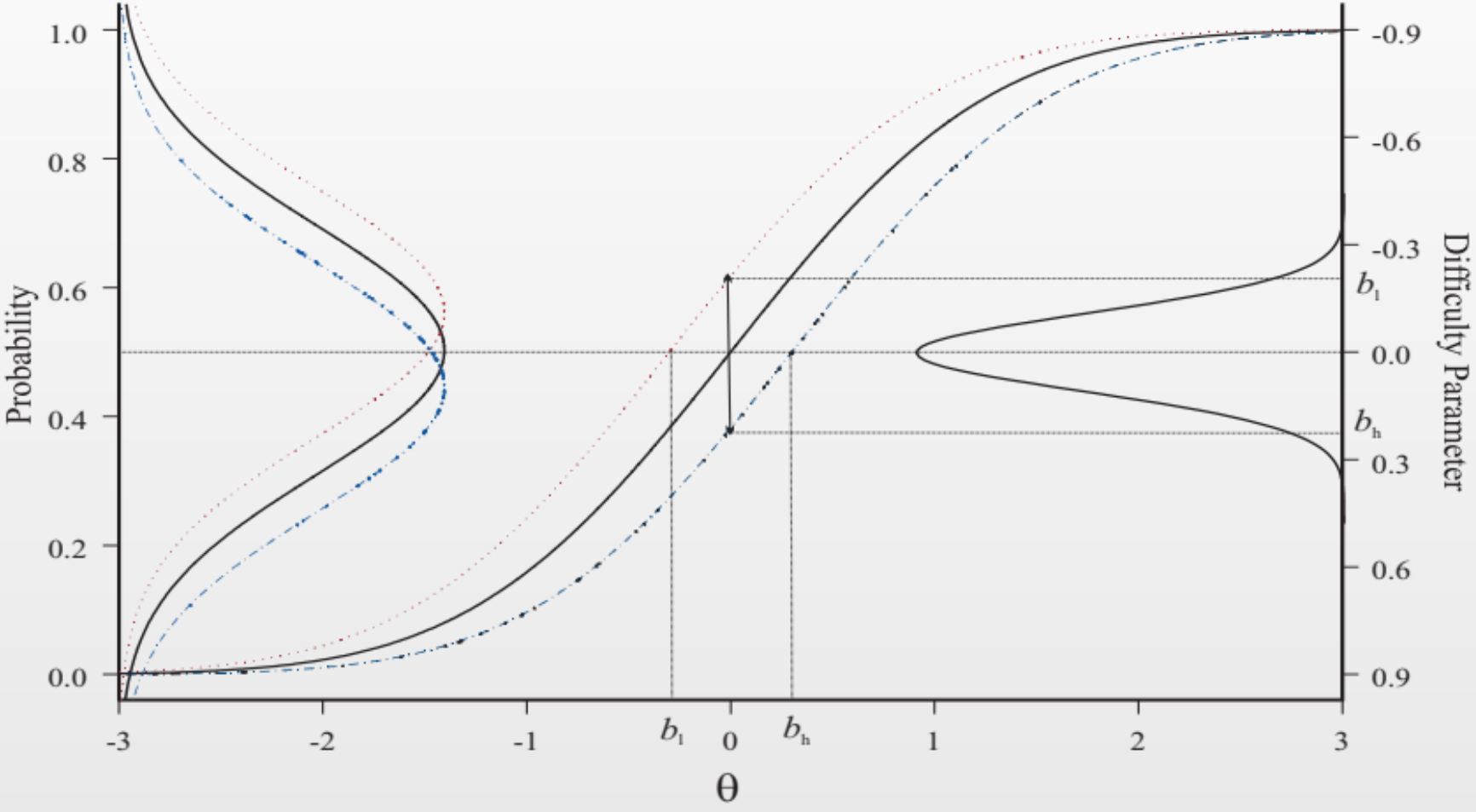
Multilevel Design

$$\begin{aligned} \theta_{ij} &\sim \mathcal{N}\left(\boldsymbol{x}_{ij}^t \boldsymbol{\beta}_j, \sigma_j^2\right) \\ \boldsymbol{\beta}_j &\sim \mathcal{N}\left(\boldsymbol{w}_j^t \boldsymbol{\gamma}, \tau^2\right) \end{aligned}$$

Graphical Display of Random Item Effects Model



Random Item Characteristic Curves



Implied (Marginal) IRT Model

$$P\left(Y_{ijk} = 1 \mid \theta_i, \boldsymbol{\xi}_k\right) = E\left(\Phi\left(a_k\theta_i - \tilde{b}_{kj}\right)\right)$$

Implied (Marginal) IRT Model

$$
\begin{aligned}
P\left(Y_{ijk} = 1 \mid \theta_i, \boldsymbol{\xi}_k\right) &= E\left(\Phi\left(a_k\theta_i - \tilde{b}_{kj}\right)\right) \\
&= E\left(P\left(Z_{ijk} \leq a_k\theta_i - \tilde{b}_{kj} \mid \tilde{b}_{kj}\right)\right)
\end{aligned}
$$

Implied (Marginal) IRT Model

$$
\begin{aligned}
P\left(Y_{ijk} = 1 \mid \theta_i, \boldsymbol{\xi}_k\right) &= E\left(\Phi\left(a_k\theta_i - \tilde{b}_{kj}\right)\right) \\
&= E\left(P\left(Z_{ijk} \leq a_k\theta_i - \tilde{b}_{kj} \mid \tilde{b}_{kj}\right)\right) \\
&= P\left(Z_{ijk} \leq a_k\theta_i - b_k - \epsilon_{b_{kj}}\right)
\end{aligned}
$$

Implied (Marginal) IRT Model

$$
\begin{aligned}
P\left(Y_{ijk} = 1 \mid \theta_i, \boldsymbol{\xi}_k\right) &= E\left(\Phi\left(a_k\theta_i - \tilde{b}_{kj}\right)\right) \\
&= E\left(P\left(Z_{ijk} \leq a_k\theta_i - \tilde{b}_{kj} \mid \tilde{b}_{kj}\right)\right) \\
&= P\left(Z_{ijk} \leq a_k\theta_i - b_k - \epsilon_{b_{kj}}\right) \\
&= P\left(Z_{ijk} + \epsilon_{b_{kj}} \leq a_k\theta_i - b_k\right)
\end{aligned}
$$

Implied (Marginal) IRT Model

$$
\begin{aligned}
P\left(Y_{ijk}=1 \mid \theta_i, \boldsymbol{\xi}_k\right) &= E\left(\Phi\left(a_k\theta_i - \tilde{b}_{kj}\right)\right) \\
&= E\left(P\left(Z_{ijk} \leq a_k\theta_i - \tilde{b}_{kj} \mid \tilde{b}_{kj}\right)\right) \\
&= P\left(Z_{ijk} \leq a_k\theta_i - b_k - \epsilon_{b_{kj}}\right) \\
&= P\left(Z_{ijk} + \epsilon_{b_{kj}} \leq a_k\theta_i - b_k\right) \\
&= \Phi\left(\frac{a_k\theta_i - b_k}{\sqrt{1+\sigma^2_{b_k}}}\right)
\end{aligned}
$$

Implied (Marginal) IRT Model

$$
\begin{aligned}
P\left(Y_{ijk}=1 \mid \theta_i, \boldsymbol{\xi}_k\right) &= E\left(\Phi\left(a_k\theta_i - \tilde{b}_{kj}\right)\right) \\
&= E\left(P\left(Z_{ijk} \leq a_k\theta_i - \tilde{b}_{kj} \mid \tilde{b}_{kj}\right)\right) \\
&= P\left(Z_{ijk} \leq a_k\theta_i - b_k - \epsilon_{b_{kj}}\right) \\
&= P\left(Z_{ijk} + \epsilon_{b_{kj}} \leq a_k\theta_i - b_k\right) \\
&= \Phi\left(\frac{a_k\theta_i - b_k}{\sqrt{1+\sigma^2_{b_k}}}\right) \\
&= \Phi\left(a^*_k\theta_i - b^*_k\right)
\end{aligned}
$$

Polytomous Response Data

Graded Response Model

$$P\left(Y_{ijk} = c \mid \theta_i, \tilde{\boldsymbol{\xi}}_{kj}\right) = \left\{ \begin{array}{l} \Phi\left(\tilde{a}_{kj}\theta_i - \tilde{\kappa}_{kj,c-1}\right) - \Phi\left(\tilde{a}_{kj}\theta_i - \tilde{\kappa}_{kj,c}\right) \\ \Psi\left(\tilde{a}_{kj}\theta_i - \tilde{\kappa}_{kj,c-1}\right) - \Psi\left(\tilde{a}_{kj}\theta_i - \tilde{\kappa}_{kj,c}\right) \end{array} \right.$$

Latent Variable Framework

Latent Augmented Responses

$$Z_{ijk} \mid Y_{ijk} = c, \theta_i, \tilde{\boldsymbol{\xi}}_{kj} \sim \mathcal{N}\left(\tilde{a}_{kj}\theta_i, 1\right) I\left(\tilde{\kappa}_{kj,c-1} \leq Z_{ijk} \leq \tilde{\kappa}_{kj,c}\right).$$

Latent Variable Framework

Latent Augmented Responses

$$Z_{ijk} \mid Y_{ijk} = c, \theta_i, \tilde{\boldsymbol{\xi}}_{kj} \sim \mathcal{N}\left(\tilde{a}_{kj}\theta_i, 1\right) I\left(\tilde{\kappa}_{kj,c-1} \leq Z_{ijk} \leq \tilde{\kappa}_{kj,c}\right).$$

Graded Response Model

$$\begin{aligned} P\left(Y_{ijk} = c \mid \theta_i, \tilde{\boldsymbol{\xi}}_{kj}\right) &= P\left(\tilde{a}_{kj}\theta_i - \tilde{\kappa}_{kj,c} \leq \tilde{Z}_{ijk} \leq \tilde{a}_{kj}\theta_i - \tilde{\kappa}_{kj,c-1}\right) \\ &= P\left(\tilde{\kappa}_{kj,c-1} - \tilde{a}_{kj}\theta_i \leq \tilde{Z}_{ijk} \leq \tilde{\kappa}_{kj,c} - \tilde{a}_{kj}\theta_i\right) \\ &= P_{ijk}(c) - P_{ijk}(c-1). \end{aligned}$$

Implied (Marginal) IRT Model

$$
\begin{aligned}
P\left(Y_{ijk} \leq c \mid \theta_i, \boldsymbol{\xi}_k\right) &= E\left(\Phi\left(a_k\theta_i - \tilde{\kappa}_{kj,c-1}\right)\right) \\
&= E\left(P\left(\tilde{Z}_{ijk} \leq a_k\theta_i - \tilde{\kappa}_{kj,c-1} \mid \tilde{\kappa}_{kj,c-1}\right)\right) \\
&= P\left(\tilde{Z}_{ijk} \leq a_k\theta_i - \kappa_{k,c-1} - \epsilon_{\kappa_{kj,c-1}}\right) \\
&= P\left(\tilde{Z}_{ijk} + \epsilon_{\kappa_{kj,c-1}} \leq a_k\theta_i - \kappa_{k,c-1}\right) \\
&= \Phi\left(\frac{a_k\theta_i - \kappa_{k,c-1}}{\sqrt{\left(1+\sigma^2_{\kappa_k}\right)}}\right) \\
&= \Phi\left(a^*_k\theta_i - \kappa^*_{k,c-1}\right).
\end{aligned}
$$

Implied (Marginal) IRT Model

Expected conditional success probability

$$P\left(Y_{ijk} = c \mid \theta_i, \boldsymbol{\xi}_k\right) = \Phi\left(a_k^* \theta_i - \kappa_{k,c-1}^*\right) - \Phi\left(a_k^* \theta_i - \kappa_{k,c}^*\right).$$

Implied (Marginal) IRT Model

Expected conditional success probability

$$P\left(Y_{ijk} = c \mid \theta_i, \boldsymbol{\xi}_k\right) = \Phi\left(a_k^* \theta_i - \kappa_{k,c-1}^*\right) - \Phi\left(a_k^* \theta_i - \kappa_{k,c}^*\right).$$

International Item Parameters

$$\begin{aligned} a_k^* &= a_k / \sqrt{1 + \sigma_{\kappa_k}^2}, \\ \kappa_{k,c}^* &= \kappa_{k,c} / \sqrt{1 + \sigma_{\kappa_k}^2}. \end{aligned}$$

Identification: linkage between countries

- Identify scale; mean and variance latent variable

Identification: linkage between countries

- Identify scale; mean and variance latent variable
- Restrictions on item parameters or population parameters latent variable

Identification: linkage between countries

- Identify scale; mean and variance latent variable
- Restrictions on item parameters or population parameters latent variable
- Detect any indeterminancies

Identification: linkage between countries

Latent mean versus country-specific difficulties

$$P\left(Y_{ijk} = 1 \mid \beta_{0j}, \tilde{b}_{kj}\right) = E\left(\Phi\left(\theta_{ij} - \tilde{b}_{kj}\right)\right)$$

Identification: linkage between countries

Latent mean versus country-specific difficulties

$$P\left(Y_{ijk} = 1 \mid \beta_{0j}, \tilde{b}_{kj}\right) = E\left(\Phi\left(\theta_{ij} - \tilde{b}_{kj}\right)\right)$$
$$= E\left(P\left(Z_{ijk} \leq \theta_{ij} - \tilde{b}_{kj} \mid \theta_{ij}, \tilde{b}_{kj}\right)\right)$$

Identification: linkage between countries

Latent mean versus country-specific difficulties

$$
\begin{aligned}
P\left(Y_{ijk} = 1 \mid \beta_{0j}, \tilde{b}_{kj}\right) &= E\left(\Phi\left(\theta_{ij} - \tilde{b}_{kj}\right)\right) \\
&= E\left(P\left(Z_{ijk} \leq \theta_{ij} - \tilde{b}_{kj} \mid \theta_{ij}, \tilde{b}_{kj}\right)\right) \\
&= P\left(Z_{ijk} \leq \beta_{0j} + e_{ij} - \tilde{b}_{kj} \mid \beta_{0j}, \tilde{b}_{kj}\right)
\end{aligned}
$$

Identification: linkage between countries

Latent mean versus country-specific difficulties

$$
\begin{aligned}
P\left(Y_{ijk}=1 \mid \beta_{0j}, \tilde{b}_{kj}\right) &= E\left(\Phi\left(\theta_{ij}-\tilde{b}_{kj}\right)\right) \\
&= E\left(P\left(Z_{ijk} \leq \theta_{ij}-\tilde{b}_{kj} \mid \theta_{ij}, \tilde{b}_{kj}\right)\right) \\
&= P\left(Z_{ijk} \leq \beta_{0j}+e_{ij}-\tilde{b}_{kj} \mid \beta_{0j}, \tilde{b}_{kj}\right) \\
&= \Phi\left(\frac{\beta_{0j}-\tilde{b}_{kj}}{\sqrt{1+\sigma_{\theta}^{2}}}\right)
\end{aligned}
$$

Identification: linkage between countries

Latent variance versus country-specific discriminations

$$
\begin{aligned}
&P\left(Y_{ijk}=1 \mid \beta_{0j}, \tilde{b}_{kj}\right) = E\left(E\left(\Phi\left(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj} \mid \theta_{ij}, \tilde{\vec{\xi}}_{kj}\right)\right)\right) \\
&= E\left(P\left(\tilde{Z}_{ijk} \leq \tilde{a}_{kj}\left(\beta_{0j} + e_{ij}\right) - \tilde{b}_{kj} \mid \beta_{0j}, \tilde{\vec{\xi}}_{kj}\right)\right) \\
&= P\left(\tilde{Z}_{ijk} \leq a_k\left(\beta_{0j} + e_{ij}\right) + \epsilon_{a_{k_j}}\left(\beta_{0j} + e_{ij}\right) - \tilde{b}_{kj} \mid \beta_{0j}, \tilde{b}_{kj}\right).
\end{aligned}
$$

Estimation; MCMC

- (Straightforward) Implementation M-H within Gibbs Algorithm

Estimation; MCMC

- (Straightforward) Implementation M-H within Gibbs Algorithm
- Sampling (Constrained) threshold parameters

Sampling Thresholds; M-H

Nation-Specific

$$\tilde{\kappa}^*_{kj,c} \sim \mathcal{N}\left(\tilde{\kappa}^{(m)}_{kj,c}, \sigma^2_{mh}\right) I\left(\tilde{\kappa}^*_{kj,c-1} < \tilde{\kappa}^*_{kj,c} < \tilde{\kappa}^{(m)}_{kj,c+1}\right)$$

Sampling Thresholds; M-H

Nation-Specific

$$\tilde{\kappa}^*_{kj,c} \sim \mathcal{N}\left(\tilde{\kappa}^{(m)}_{kj,c}, \sigma^2_{mh}\right) I\left(\tilde{\kappa}^*_{kj,c-1} < \tilde{\kappa}^*_{kj,c} < \tilde{\kappa}^{(m)}_{kj,c+1}\right)$$

Acceptance Ratio

$$R = \frac{p\left(\tilde{\boldsymbol{\kappa}}^*_{kj} \mid \theta_{ij}, \boldsymbol{\kappa}_k, \sigma^2_{\kappa_k}, \boldsymbol{y}\right)}{p\left(\tilde{\boldsymbol{\kappa}}^{(m)}_{kj} \mid \theta_{ij}, \boldsymbol{\kappa}_k, \sigma^2_{\kappa_k}, \boldsymbol{y}\right)} \prod_{c=1}^{C_k-1} \frac{\Phi\left(\frac{\tilde{\kappa}^{(m)}_{kj,c+1}-\tilde{\kappa}^{(m)}_{kj,c}}{\sigma_{mh}}\right) - \Phi\left(\frac{\tilde{\kappa}^*_{kj,c-1}-\tilde{\kappa}^{(m)}_{kj,c}}{\sigma_{mh}}\right)}{\Phi\left(\frac{\tilde{\kappa}^*_{kj,c+1}-\tilde{\kappa}^*_{kj,c}}{\sigma_{mh}}\right) - \Phi\left(\frac{\tilde{\kappa}^{(m)}_{kj,c-1}-\tilde{\kappa}^*_{kj,c}}{\sigma_{mh}}\right)}$$

Sampling Thresholds; M-H

International

$$p\left(\kappa_{k,c} \mid \tilde{\boldsymbol{\kappa}}_{k,c}, \sigma^2_{\kappa_k}\right) \quad \propto \quad \prod_j p\left(\tilde{\kappa}_{kj,c} \mid \kappa_{k,c}, \sigma^2_{\kappa_k}\right),$$

Sampling Thresholds; M-H

International

$$p\left(\kappa_{k,c} \mid \tilde{\boldsymbol{\kappa}}_{k,c}, \sigma^2_{\kappa_k}\right) \quad \propto \quad \prod_j p\left(\tilde{\kappa}_{kj,c} \mid \kappa_{k,c}, \sigma^2_{\kappa_k}\right),$$

Proposal

$$\kappa^*_{k,c} \sim \mathcal{N}\left(\kappa^{(m)}_{k,c}, \sigma^2_{mh}\right) I\left(\kappa^*_{k,c-1} < \kappa^*_{k,c} < \kappa^{(m)}_{k,c+1}\right)$$

Sampling Thresholds; Gibbs Sampling

Nation-Specific Threshold

$\mathcal{A} = (\tilde{\kappa}_{kj,c-1}, \tilde{\kappa}_{kj,c+1})$

$$\tilde{\kappa}_{kj,c} = \nu_{kj,c} I_{\mathcal{A}}(\nu_{kj,c}),$$

$$I_{\mathcal{A}}(\nu_{kj,c}) = \begin{cases} 1 & \nu_{kj,c} \in \mathcal{A} \\ 0 & \text{otherwise,} \end{cases}$$

where $\nu_{kj,c} \sim \mathcal{N}\left(\kappa_{k,c}, \sigma^2_{\kappa_k}\right)$.

Sampling Thresholds; Gibbs Sampling

Distribution Unrestricted Nation-Specific Threshold

$$F(\nu_{kj,c}) = \Phi\left((\nu_{kj,c} - \kappa_{k,c})/\sigma_{\kappa_k}\right)$$

Sampling Thresholds; Gibbs Sampling

Distribution Unrestricted Nation-Specific Threshold

$$F(\nu_{kj,c}) = \Phi\left((\nu_{kj,c} - \kappa_{k,c})/\sigma_{\kappa_k}\right)$$

Distribution Restricted Nation-Specific Threshold

$$G\left(\tilde{\kappa}_{kj,c}\right) = \frac{F\left(\tilde{\kappa}_{kj,c}\right) - F\left(\tilde{\kappa}_{kj,c-1}\right)}{F\left(\tilde{\kappa}_{kj,c+1}\right) - F\left(\tilde{\kappa}_{kj,c-1}\right)}$$

Sampling Thresholds; Gibbs Sampling

Distribution Restricted Nation-Specific Threshold

$$
\begin{aligned}
\nu_{kj,c} &= F^{-1}\left(G\left(\tilde{\kappa}_{kj,c}\right)\right) = F^{-1}\left(p\right) = \kappa_{k,c} + \sigma_{\kappa_k}\Phi^{-1}\left(p\right) \\
&= \kappa_{k,c} + \sigma_{\kappa_k}\Phi^{-1}\left(\frac{F\left(\tilde{\kappa}_{kj,c}\right) - F\left(\tilde{\kappa}_{kj,c-1}\right)}{F\left(\tilde{\kappa}_{kj,c+1}\right) - F\left(\tilde{\kappa}_{kj,c-1}\right)}\right)
\end{aligned}
$$

Sampling Thresholds; Gibbs Sampling

Distribution Restricted Nation-Specific Threshold

$$
\begin{aligned}
\nu_{kj,c} &= F^{-1}\left(G\left(\tilde{\kappa}_{kj,c}\right)\right) = F^{-1}\left(p\right) = \kappa_{k,c} + \sigma_{\kappa_k}\Phi^{-1}\left(p\right) \\
&= \kappa_{k,c} + \sigma_{\kappa_k}\Phi^{-1}\left(\frac{F\left(\tilde{\kappa}_{kj,c}\right) - F\left(\tilde{\kappa}_{kj,c-1}\right)}{F\left(\tilde{\kappa}_{kj,c+1}\right) - F\left(\tilde{\kappa}_{kj,c-1}\right)}\right)
\end{aligned}
$$

Sample Higher-level Parameters

$$
\begin{aligned}
p\left(\boldsymbol{\nu}_{k,c}, \kappa_{k,c}, \sigma^2_{\kappa_k} \mid \tilde{\boldsymbol{\kappa}}_{k,c}\right) &\propto p\left(\tilde{\boldsymbol{\kappa}}_{k,c} \mid \boldsymbol{\nu}_{k,c}, \kappa_{k,c}, \sigma^2_{\kappa_k}\right) p\left(\boldsymbol{\nu}_{k,c}, \kappa_{k,c}, \sigma^2_{\kappa_k}\right) \\
&\propto p\left(\tilde{\boldsymbol{\kappa}}_{k,c} \mid \boldsymbol{\nu}_{k,c}\right) p\left(\boldsymbol{\nu}_{k,c}, \kappa_{k,c}, \sigma^2_{\kappa_k}\right) \\
&\propto p\left(\boldsymbol{\nu}_{k,c} \mid \kappa_{k,c}, \sigma^2_{\kappa_k}\right) p\left(\kappa_{k,c}\right) p\left(\sigma^2_{\kappa_k}\right),
\end{aligned}
$$

Testing for Measurement Invariance

- Testing cross-national item variances $\sigma^2_{a_k}$, $\sigma^2_{b_k}$

Testing for Measurement Invariance

- Testing cross-national item variances $\sigma^2_{a_k}$, $\sigma^2_{b_k}$
- Testing country means μ_j and variances σ^2_j

Bayes Factor

$$p\left(\sigma^2_{b_k} = 0 \mid \mathbf{y}\right) = \frac{p\left(\mathbf{y} \mid \sigma^2_{b_k} = 0\right) p\left(\sigma^2_{b_k} = 0\right)}{\int p\left(\mathbf{y} \mid \sigma^2_{b_k}\right) p\left(\sigma^2_{b_k}\right) d\sigma^2_{b_k}}$$

Bayes Factor

$$p\left(\sigma^2_{b_k} = 0 \mid \mathbf{y}\right) = \frac{p\left(\mathbf{y} \mid \sigma^2_{b_k} = 0\right) p\left(\sigma^2_{b_k} = 0\right)}{\int p\left(\mathbf{y} \mid \sigma^2_{b_k}\right) p\left(\sigma^2_{b_k}\right) d\sigma^2_{b_k}}$$

Encompassing Prior

$$\text{BF} = \frac{p\left(\sigma^2_{b_k} = 0 \mid \mathbf{y}\right)}{p\left(\sigma^2_{b_k} = 0\right)} = \frac{p\left(\mathbf{y} \mid H_0\right)}{p\left(\mathbf{y} \mid H_1\right)}$$

Bayes Factor, Common Likelihood

Common Likelihood

$$
\begin{aligned}
\mathrm{BF} &= \frac{\int_{\Theta_0} p\left(\sigma^2_{b_k} \mid H_0\right) p\left(\mathbf{y} \mid \sigma^2_{b_k}\right) d\sigma^2_{b_k}}{\int_{\Theta_1} p\left(\sigma^2_{b_k} \mid H_1\right) p\left(\mathbf{y} \mid \sigma^2_{b_k}\right) d\sigma^2_{b_k}} \\
&= \int_{\Theta_1} \left[\frac{p\left(\sigma^2_{b_k} \mid H_0\right)}{p\left(\sigma^2_{b_k} \mid H_1\right)}\right] \frac{p\left(\sigma^2_{b_k} \mid H_1\right) p\left(\mathbf{y} \mid \sigma^2_{b_k}\right)}{p\left(\mathbf{y} \mid H_1\right)} d\sigma^2_{b_k} \\
&= \int_{\Theta_1} \left[\frac{p\left(\sigma^2_{b_k} \mid H_0\right)}{p\left(\sigma^2_{b_k} \mid H_1\right)}\right] p\left(\mathbf{y} \mid \sigma^2_{b_k}, H_1\right) d\sigma^2_{b_k} \\
&= E\left[\frac{p\left(\sigma^2_{b_k} \mid H_0\right)}{p\left(\sigma^2_{b_k} \mid H_1\right)} \mid \mathbf{y},\right]
\end{aligned}
$$

HPD Region Testing

The point $(\tilde{\boldsymbol{\xi}}_{k1} = \tilde{\boldsymbol{\xi}}_{k2} = \ldots = \tilde{\boldsymbol{\xi}}_{kJ} = \boldsymbol{\xi}_k)$

$$P\left(p(\tilde{\boldsymbol{\xi}}_k \mid \mathbf{y}) > p(\tilde{\boldsymbol{\xi}}_{k1} = \tilde{\boldsymbol{\xi}}_{k2} = \ldots = \tilde{\boldsymbol{\xi}}_{kJ} = \boldsymbol{\xi}_k \mid \mathbf{y}) \mid \mathbf{y}\right) < 1 - \alpha,$$

HPD Region Testing

The point $\left(\tilde{\boldsymbol{\xi}}_{k1} = \tilde{\boldsymbol{\xi}}_{k2} = \ldots = \tilde{\boldsymbol{\xi}}_{kJ} = \boldsymbol{\xi}_k\right)$

$$P\left(\chi^2_{2J} \leq \sum_j \left(\boldsymbol{\xi}_k - \boldsymbol{\mu}_{\tilde{\xi}_{kj}}\right)^t \boldsymbol{\Omega}^{-1}_{\tilde{\xi}_{kj}} \left(\boldsymbol{\xi}_k - \boldsymbol{\mu}_{\tilde{\xi}_{kj}}\right) \mid \mathbf{z}_{kj}, \boldsymbol{\theta}_j, \boldsymbol{\Sigma}_{\tilde{\xi}}, \boldsymbol{\xi}_k\right) < 1-\alpha.$$

HPD Region Testing

The point $\left(\tilde{\boldsymbol{\xi}}_{k1} = \tilde{\boldsymbol{\xi}}_{k2} = \ldots = \tilde{\boldsymbol{\xi}}_{kJ} = \boldsymbol{\xi}_k\right)$

$$P\left(\chi^2_{2J} \leq \sum_j \left(\boldsymbol{\xi}_k - \boldsymbol{\mu}_{\tilde{\xi}_{kj}}\right)^t \boldsymbol{\Omega}^{-1}_{\tilde{\xi}_{kj}} \left(\boldsymbol{\xi}_k - \boldsymbol{\mu}_{\tilde{\xi}_{kj}}\right) \mid \mathbf{z}_{kj}, \boldsymbol{\theta}_j, \boldsymbol{\Sigma}_{\tilde{\xi}}, \boldsymbol{\xi}_k\right) < 1-\alpha.$$

DIC: Comparing Constrained and Unconstrained Models

DIC

$$
\begin{aligned}
\text{DIC} &= \overline{D\left(\boldsymbol{\Lambda}\right)} + \overline{D\left(\boldsymbol{\Lambda}\right)} - D\left(\hat{\boldsymbol{\Lambda}}\right) \\
&= \overline{D\left(\boldsymbol{\Lambda}\right)} + p_D,
\end{aligned}
$$

DIC: Comparing Constrained and Unconstrained Models

DIC

$$\begin{aligned} \text{DIC} &= \overline{D(\mathbf{\Lambda})} + \overline{D(\mathbf{\Lambda})} - D(\hat{\mathbf{\Lambda}}) \\ &= \overline{D(\mathbf{\Lambda})} + p_D, \end{aligned}$$

Deviance Function

$$D(\mathbf{\Lambda}) = -2 \log p(\mathbf{\Lambda})$$

DIC: Integrated Data Likelihood

$$
\begin{aligned}
p\left(\boldsymbol{y} \mid \boldsymbol{\theta}, \boldsymbol{\xi}, \boldsymbol{\Sigma}_{\tilde{\xi}}\right) &= \int p(\boldsymbol{z} \mid \boldsymbol{\theta}, \boldsymbol{\xi}, \boldsymbol{\Sigma}_{\tilde{\xi}}) d\boldsymbol{z} \\
&= \int \int p(\boldsymbol{z} \mid \boldsymbol{\theta}, \tilde{\boldsymbol{\xi}}) p(\tilde{\boldsymbol{\xi}} \mid \boldsymbol{\xi}, \boldsymbol{\Sigma}_{\tilde{\xi}}) d\tilde{\boldsymbol{\xi}} d\boldsymbol{z} \\
&= \int \frac{p(\boldsymbol{z}, \tilde{\boldsymbol{\xi}} \mid \boldsymbol{\theta}, \boldsymbol{\xi}, \boldsymbol{\Sigma}_{\tilde{\xi}})}{p(\tilde{\boldsymbol{\xi}} \mid \boldsymbol{z}, \boldsymbol{\theta}, \boldsymbol{\xi}, \boldsymbol{\Sigma}_{\tilde{\xi}})} d\boldsymbol{z}.
\end{aligned}
$$

Generalized Partial Credit Model - Group Specific

For item k, group j, person i, the probability of scoring category $c = 0, ..., C$ is denoted by:

$$P\left(Y_{ijk} = c \mid \theta_{ij}, \tilde{a}_{kj}, \tilde{b}_{ckj}\right) = \frac{exp[\sum_0^c(\tilde{a}_{kj}\theta_{ij}) - (\tilde{b}_{ckj})]}{\sum_0^C(exp[\sum_0^c(\tilde{a}_{kj}\theta_{ij}) - (\tilde{b}_{ckj})])},$$

where $\sum_0^{c=0}(\tilde{a}_{kj}\theta_{ij}) - (\tilde{b}_{ckj}) = 0$

Restrictions:

$$\text{for each } j, \sum_k \sum_c \tilde{b}_{ckj} = 0, \prod_k \tilde{a}_{kj} = 1$$

Random item parameters

The random group-specific item parameters are assumed to be normally distributed:

$$\tilde{\boldsymbol{\xi}}_{kj} \mid \boldsymbol{\xi}_k, \boldsymbol{\Sigma}_{\xi_k} \sim \mathcal{N}(\boldsymbol{\xi}_k, \boldsymbol{\Sigma}_{\xi_k})$$

$$\tilde{\boldsymbol{\xi}}_{kj} = (\tilde{a}_{kj}, \tilde{b}_{1kj}, ..., \tilde{b}_{Ckj})$$

$$\boldsymbol{\xi}_k = (a_k, b_{1k}, ..., b_{Ck})$$

$$\boldsymbol{\Sigma}_{\xi_k} = \begin{bmatrix} \sigma^2_{a_k} & \sigma_{a_k b_{1k}} & \cdots & \sigma_{a_k b_{Ck}} \\ \sigma_{b_{1k} a_k} & \sigma^2_{b_{1k}} & \cdots & \sigma_{b_{1k} b_{Ck}} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{b_{Ck} a_k} & \sigma_{b_{Ck} b_{1k}} & \cdots & \sigma^2_{b_{Ck}} \end{bmatrix}$$

Random item parameters

Population model for the international item parameters:

$$\boldsymbol{\xi}_k \mid \boldsymbol{\mu}_\xi, \boldsymbol{\Sigma}_\xi \sim \mathcal{N}(\boldsymbol{\mu}_\xi, \boldsymbol{\Sigma}_\xi)$$

$$\boldsymbol{\mu}_\xi = (a_0, b_{10}, ..., b_{C0})$$

$$\boldsymbol{\Sigma}_\xi = \begin{bmatrix} \sigma_a^2 & \sigma_{ab_1} & \cdots & \sigma_{ab_C} \\ \sigma_{b_1 a} & \sigma_{b_1}^2 & \cdots & \sigma_{b_1 b_C} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{b_C a} & \sigma_{b_C b_1} & \cdots & \sigma_{b_C}^2 \end{bmatrix}$$

Hyper priors:

$$\boldsymbol{\Sigma}_\xi \sim \mathcal{IW}(\nu, \boldsymbol{\Sigma}_0)$$
$$\boldsymbol{\mu}_\xi \mid \boldsymbol{\Sigma}_\xi \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_\xi / K_0)$$

Metropolis-Hastings step

- Metropolis-Hastings step necessary for sampling of the group-specific item parameters
- For each k and j and for $c = 1, C$, sample a proposal $\tilde{\xi}^*_{ckj}$ from the proposal distribution $N(\tilde{\xi}^{(m)}_{ckj}, \sigma^2_{mh}) I(\tilde{a}^*_{kj} > 0)$.
- The acceptance ratio R is now defined as:

$$R = \frac{p(\mathbf{y} \mid \tilde{\xi}^*_{ckj}, \boldsymbol{\theta}^{(m)}, \tilde{\boldsymbol{\xi}}^{(m+1)}_{kj}) p(\tilde{\xi}^*_{ckj} \mid (\xi^{(m)}_{ck}, \sigma^{(m)}_{\xi_{ck}})}{p(\mathbf{y} \mid \tilde{\xi}^{(m)}_{ckj}, \boldsymbol{\theta}^{(m)}, \tilde{\boldsymbol{\xi}}^{(m+1)}_{kj}) p(\tilde{\xi}^{(m)}_{ckj} \mid (\xi^{(m)}_{ck}, \sigma^{(m)}_{\xi_{ck}})} \quad (1)$$

- A random uniform number u_{ckj} is drawn, and the proposal is accepted when $u_{ckj} \leq R$.

Invariance Test

- Diagonal elements of a $\mathcal{IW}$ distributed matrix are inverse gamma distributed.
-

$$\boldsymbol{\Sigma}_{\xi_k} = \begin{bmatrix} \sigma^2_{a_k} & \sigma_{a_k b_{1k}} & \dots & \sigma_{a_k b_{Ck}} \\ \sigma_{b_{1k} a_k} & \sigma^2_{b_{1k}} & \dots & \sigma_{b_{1k} b_{Ck}} \\ \dots & \dots & \dots & \dots \\ \sigma_{b_{Ck} a_k} & \sigma_{b_{Ck} b_{1k}} & \dots & \sigma^2_{b_{Ck}} \end{bmatrix}$$

- Bayes factor for nested models can be applied in the same way as for the 2 parameter model.
- Compute Bayes factor for each of the diagonal elements:

$$\text{BF} = \frac{P(\sigma^2_{\xi_{ck}} < \delta | \mathbf{y}, M1)}{P(\sigma^2_{\xi_{ck}} < \delta | M1)}$$

Longitudinal invariance - causes of non-invariance

- Change of a person's understanding of items over time can result in "item parameter drift".
- Lengthening or shortening of the intervals between scale points over time
- Change caused by f.e.: experiences between time 1 and time 2, learning effects, testing effects

Longitudinal invariance - model on the latent variable

- The change over time in the latent variable for person i is modeled by a (random) linear time effect

$$\theta_{ij} = \theta_i + \lambda_i(T_{ij} - T_{i0}) + \beta_j + e_{ij}$$

- Multilevel model:

$$\begin{aligned}
\theta_{ij} &\sim \mathcal{N}\left(\boldsymbol{\lambda}_i \mathbf{t}_i + \beta_j, \sigma^2_{\theta_{ij}}\right) \\
\boldsymbol{\lambda}_i &\sim \mathcal{N}\left(\boldsymbol{\kappa}, \Phi\right) \\
\boldsymbol{\beta}_j &\sim \mathcal{N}\left(\gamma, \tau^2\right)
\end{aligned}$$

$\mathbf{t}_i = (1, (T_{ij} - T_{i0}))$

Longitudinal invariance - model on item parameters

- The change over time in item parameters is modeled by a linear time effect of the average time on occasion j, $t_j = \sum_i T_{ij}/I$

$$\tilde{\xi}_{ckj} = \xi_{ck} + \delta_{ck} t_j + \epsilon_{\xi_{ck}}$$

- A multivariate regression model is formulated for the time specific item parameters.

$$\boldsymbol{\Xi_k} \mid \boldsymbol{\Delta}_k, \Sigma_k \quad \sim \quad \mathcal{N}(\mathbf{T}\boldsymbol{\Delta}_k, \Sigma_k)$$

$$\begin{bmatrix} a_{k1} & b_{1k1} & \dots & b_{Ck1} \\ \dots & \dots & \dots & \dots \\ a_{kJ} & b_{1kJ} & \dots & b_{CkJ} \end{bmatrix} = \begin{bmatrix} 1 & t_1 \\ \dots & \dots \\ 1 & t_J \end{bmatrix} \begin{bmatrix} a_k & b_{1k} & \dots & b_{Ck} \\ \delta_{a_k} & \delta_{b_{1k}} & \dots & \delta_{b_{Ck}} \end{bmatrix} +$$

$$\mathbf{u}_{kj} \quad \sim \quad \mathcal{N}(0, \Sigma_k)$$

References

- De Jong, M.G., Steenkamp, J.B.E.M., and Fox, J.-P. (2007). Relaxing crossnational measurement invariance using a hierarchical IRT model. JCR, 34, 260–278.
- De Jong, M.G., Steenkamp, J.B.E.M., Fox, J.-P., and Baumgartner, H. (2008). Using item response theory to measure extreme response style in marketing research: A global investigation. JMR, 45, 104–115.
- Glas, C.A.W. and van der Linden, W.J. (2003). Computerized adaptive testing with item cloning. APM, 27, 247–261.
- Fox, J.-P. (2010) Bayesian Item Response Modeling: Theory and Applications. New York: Springer.
- Fox, J.-P. (2007). Multilevel IRT modeling in practice. JSS, 20 (5).
- Janssen, R., Tuerlinckx, F., Meulders, M., and De Boeck, P. (2000). A hierarchical IRT model for criterion-referenced measurement. JEBS, 25, 285–306.
- Verhagen, A.J. and Fox, J.-P. (2011). Bayesian Tests of Measurement Invariance. Psychometrika (under review).
- A. J. Verhagen and Fox, J.-P. (2010). Random item effects modeling for cross-national survey data. In E. Davidov & P. Schmidt, and J. Billiet (Eds.), Cross-cultural Analysis: Methods and Applications (pp), London: Routeledge Academic.