

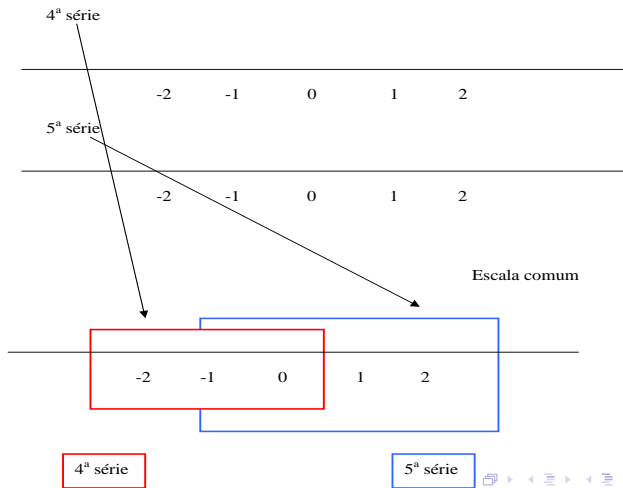
# Modelo de grupos múltiplos e Equalização

Prof. Caio Azevedo

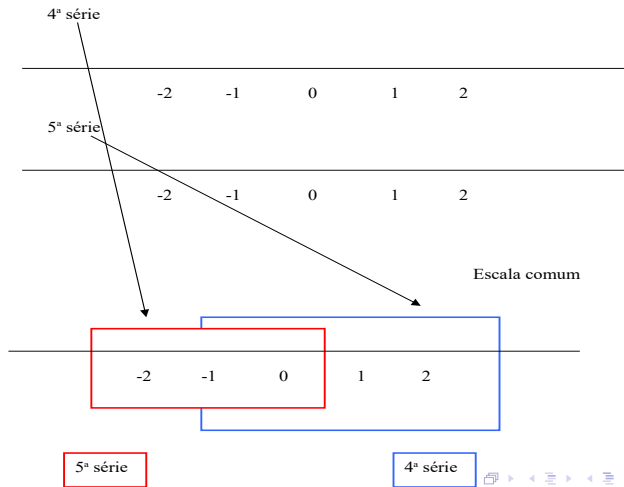
# Equalização

- Como comparar indivíduos submetidos a diferentes provas?
- Equalizar: tornar resultados comparáveis (oriundos de diferentes métricas).
- Em teoria se os testes forem paralelos (semelhantes em termos de dificuldade e discriminação e outras características que estejam relacionadas com a escala) eles poderiam ser, inclusive, totalmente diferentes (veja o Capítulo do livro Andrade et al, 2000, à respeito). Nesse caso, em princípio, a Equalização tornar-se-ia desnecessária.
- (Classes de) Processos de Equalização:
  - A posteriori: média desvio, regressão, .
  - Em concomitância com o ajuste do modelo: modelo de grupos múltiplos (Bock and Zimowski, 1997).

# Traços latentes



# Parâmetros de dificuldade



- Duas escalas (A e B).

- Temos que:

- $a_i^B = \frac{a_i^A}{\alpha}$

- $b_i^B = \alpha b_i^A + \beta$

- $\theta_j^B = \alpha \theta_j^A + \beta$

- O parâmetro  $c$  não depende da escala.

- Exercício: como ficaria o processo para o modelo de 4 parâmetros?

- Como determinar  $(\alpha, \beta)$ ?

## Exemplo 2: Dados do “International Project on Mathematical Attainment (IPMA)”

- Projeto Internacional de Desempenho (aquisição de conhecimento) em Matemática.
- Originalmente, é um longitudinal (com exclusões) com 4 instantes (séries).
- 568 alunos de primeiro ano foram selecionados de oito escolas primárias públicas (no primeiro instante). Ao longo das séries subsequentes, alguns alunos abandonaram o estudo por diferentes motivos. O conjunto de dados presente consiste no seguinte número de alunos, desde o primeiro até o quarto grau: 556, 556, 401 e 295.

## Exemplo 2

- Consideraremos as duas primeiras séries.
- Grupos independentes: veja artigo Azevedo et al (2012)  
<http://www.sciencedirect.com/science/article/pii/S0167947312001430> (também para maiores detalhes), onde se obteve evidências empíricas a respeito da nulidade das correlações intra indivíduos.
- Grupo 1: 1ª série - Grupo 2: 2ª série. Considera-se somente os 40 primeiros itens.
- Os estudantes estão agrupados em salas de aula e estas, nas escolas (estrutura hierárquica).

## Exemplo 2

- Informação disponível: itens dissertativos corretos como corretos / errados, idade do aluno, sala de aula, gênero, escola, professor. A análise apresentada em Azevedo et al (2012) revela que as correlações posteriores entre os traços latentes não são significativas e, portanto, um modelo IRT de múltiplos grupos pode ser considerado.
- Quatro testes (correspondentes a cada série): Teste 1 - 20 itens. Para os graus dois a quatro, são consideradas as respostas aos 20 novos itens e os 20 itens de teste anteriores, o que leva a 40 itens para cada série e um total de 80 itens de teste diferentes.



# Estrutura dos testes

Test	Item			
	1 - 20	21 - 40	41 - 60	61 - 80
1	■			
2	■	■		
3		■	■	
4			■	■

## Exemplo 2: Equalização a posteriori

- Ajusta-se um modelo de um único grupo para cada um dos grupos.
- Note que os resultados, em geral, não são comparáveis diretamente.
- Coloca-se os resultados de um dos grupos, na escala do outro, através de algum método de Equalização a posteriori.

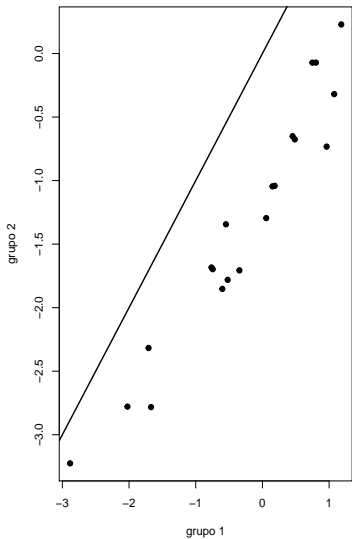
## Método da média-desvio

- Obter  $\alpha = \frac{S_{G_1}}{S_{G_2}}$  e  $\beta = M_{G_1} - \alpha M_{G_2}$ , em que  $S_{G_i}$  é o desvio-padrão das estimativas parâmetros de dificuldade dos itens comuns, obtidas na métrica do Grupo  $i$  e  $M_{G_i}$  é a média das estimativas parâmetros de dificuldade dos itens comuns, obtidas na métrica do Grupo  $i$ ,  $i=1,2$ .
- Obter as estimativas dos (de todos os) parâmetros do grupo 2, na métrica do grupo 1, através de:

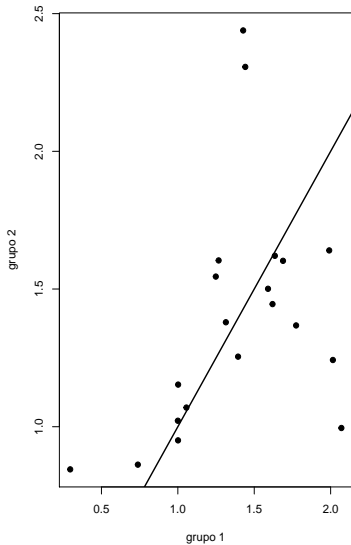
$$a_{G_2}^{G_1} = \frac{1}{\alpha} a_{G_2} \quad ; \quad b_{G_2}^{G_1} = \alpha b_{G_2} + \beta$$
$$\theta_{G_2}^{G_1} = \alpha \theta_{G_2} + \beta$$

- Este método é simétrico (a escolha de quem serão os grupos 1 e 2 são equivalentes).

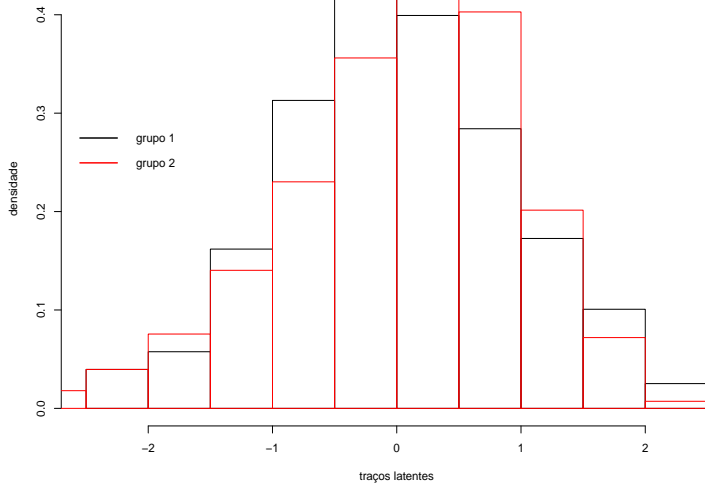
parâmetros de dificuldade

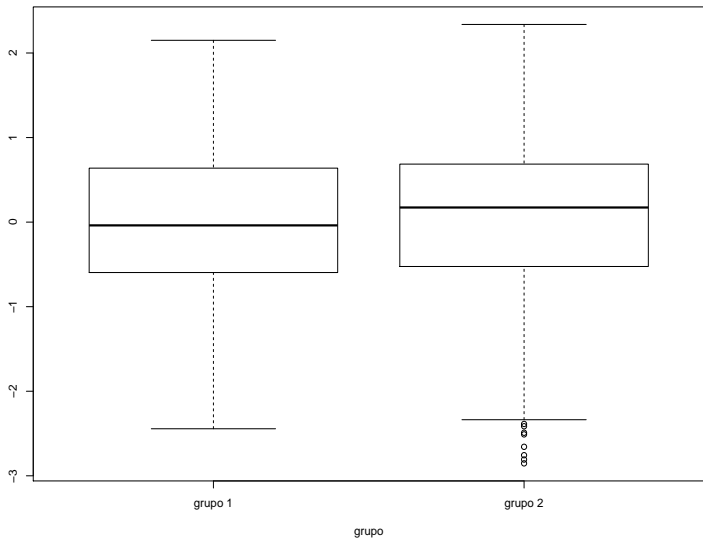


parâmetros de discriminação

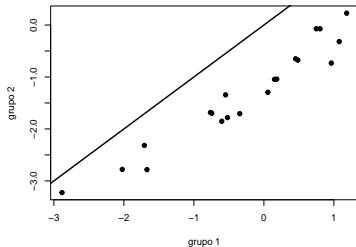


Histogram of rthetaG1

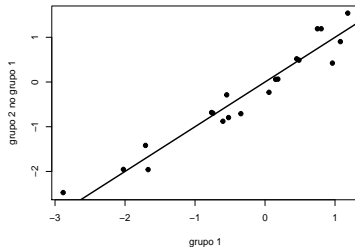




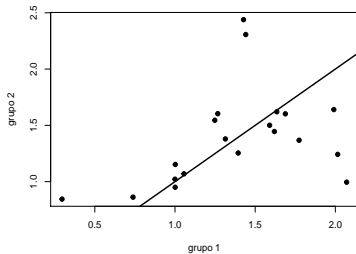
parâmetros de dificuldade – diferentes métricas



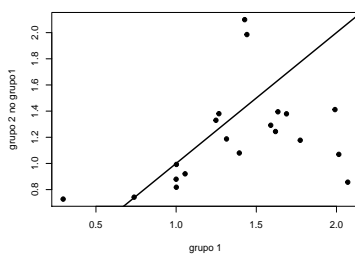
parâmetros de dificuldade – mesmas métricas



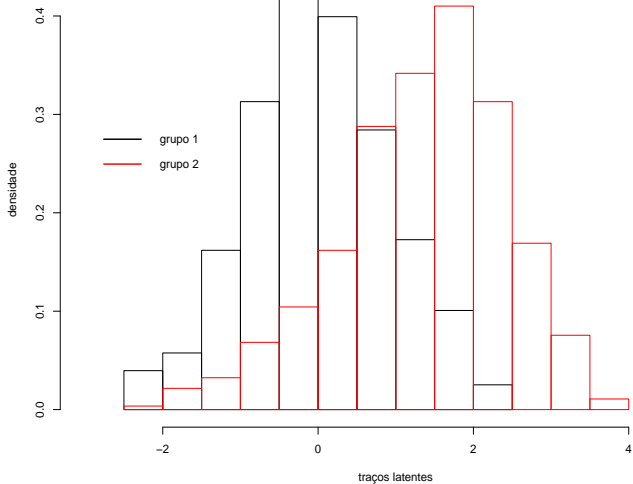
parâmetros de discriminação – diferentes métricas



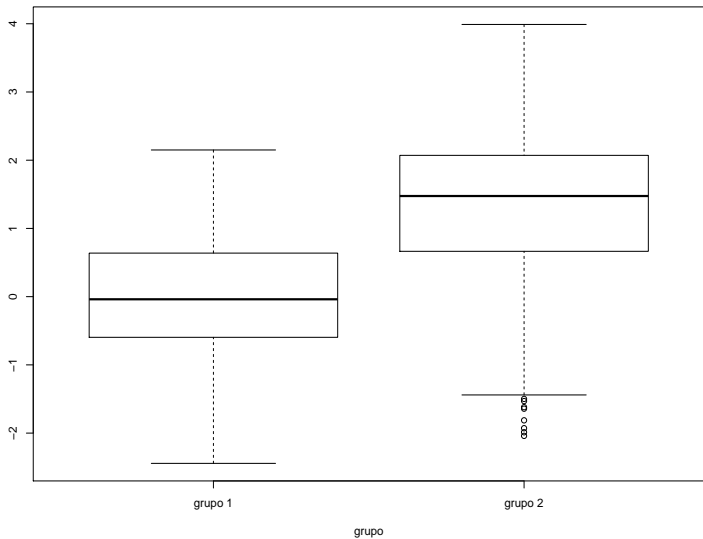
parâmetros de discriminação – diferentes métricas



### histogramas dos traços latentes







# Estimação no Modelo de grupos múltiplos

- Considera a situação em que temos indivíduos pertencentes a diferentes grupos: séries, turnos, regiões.
- Indivíduos pertencentes a um mesmo grupo possuem características comuns e são mais semelhantes entre si do que indivíduos pertencentes a outros grupos.
- Os grupos são previamente definidos  $\neq$  modelos de variáveis latentes.
- Modelo de grupos múltiplos Bock and Zimowski (1997).

# Modelo logístico unidimensional de 3 parâmetros (várias populações)

$$P(Y_{ijk} = 1 | (\theta_{jk}, \zeta_i)) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_{jk} - b_i)}}$$

$i = 1, \dots, I$  (item),  $j = 1, \dots, n_k$  (indivíduo),  $k = 1, \dots, K$ , (grupo)

- $Y_{ijk}$  : é a resposta do indivíduo  $j$ , do grupo  $k$  ao item  $i$ . É igual a 1 se o indivíduo responde corretamente e 0 caso contrário.
- $\theta_{jk}$  : é o traço latente (conhecimento, nível de depressão, etc) do indivíduo  $j$ .  $\zeta_i : (a_i, b_i, c_i)$ .
- $a_i$  : é o parâmetro de discriminação associado ao item  $i$ .
- $b_i$  : é o parâmetro de dificuldade associado ao item  $i$ .
- $c_i$  : é o parâmetro de “acerto casual” associado ao item  $i$ .

# Distribuição dos traços latentes

$$\theta_{jk} | \boldsymbol{\eta}_{\theta_k} \sim N(\mu_{\theta_k}, \psi_{\theta_k})$$

- $\boldsymbol{\eta}_{\theta_k}$ : parâmetros populacionais
- Identificabilidade  $\mu_{\theta_1} = 0, \psi_{\theta_1} = 1$ .
- Estimar  $(\mu_{\theta_k}, \psi_{\theta_k}), k = 2, \dots, K$ .
- Notações:  $\mathbf{y}_{.jk} = (y_{1jk}, \dots, y_{n_kjk})'$ ,  $\boldsymbol{\zeta} = (\boldsymbol{\zeta}'_1, \dots, \boldsymbol{\zeta}'_I)'$ ,  $\boldsymbol{\zeta}_i = (a_i, b_i, c_i)'$ ,  
 $\boldsymbol{\eta}_{\theta} = (\boldsymbol{\eta}'_{\theta_1}, \dots, \boldsymbol{\eta}'_{\theta_K})'$ ,  $I_{jk}$ : é o conjunto de itens respondidos pelo indivíduo  $j$  do grupo  $k$ .

## Probabilidade marginal

$$P(\mathbf{Y}_{.jk} = \mathbf{y}_{.jk} | \zeta, \boldsymbol{\eta}_k) = \int_{\mathfrak{R}} P(\mathbf{Y}_{.jk} = \mathbf{y}_{.jk} | \zeta, \theta) g(\theta | \boldsymbol{\eta}_{\theta_k}) d\theta$$

De modo que a verossimilhança é dada por

$$L(\zeta, \boldsymbol{\eta}_\theta) = \prod_{k=1}^K \prod_{j=1}^{n_k} P(\mathbf{Y}_{.jk} = \mathbf{y}_{.jk} | \zeta, \boldsymbol{\eta}_{\theta_k})$$

E a log-verossimilhança, por sua vez, é dada por

$$l(\zeta, \boldsymbol{\eta}_\theta) = \sum_{k=1}^K \sum_{j=1}^{n_k} \ln P(\mathbf{Y}_{.jk} = \mathbf{y}_{.jk} | \zeta, \boldsymbol{\eta}_{\theta_k}) = \sum_{k=1}^K \sum_{j=1}^{n_k} \ln P(\mathbf{Y}_{.jk} | \zeta, \boldsymbol{\eta}_{\theta_k})$$

- Estimadores de máxima verossimilhança marginal

$$\frac{\partial l(\zeta, \eta_\theta)}{\partial \zeta} = \mathbf{0}$$

$$\frac{\partial l(\zeta, \eta_\theta)}{\partial \eta_\theta} = \mathbf{0}$$

- Sem solução explícita.
- Problema: dessa forma teríamos que estimar todos os parâmetros simultaneamente, pois  $\frac{\partial^2 l(\zeta, \eta)}{\partial \zeta \partial \eta_\theta} \neq \mathbf{0}$

- Se  $\eta_\theta$  for conhecido, usa-se a abordagem de Bock & Aitkin.
- Se  $\zeta$  for conhecido, obter a estimativa do vetor  $\eta_\theta$  é um processo relativamente simples.
- Abordagem, de Bock & Aitkin, parâmetros dos itens e populacionais: Máxima verossimilhança marginal-perfilada (MVMP) ou Moda Marginal-Perfilada (MMP).

$$\frac{\partial l(\zeta, \hat{\eta}_\theta)}{\partial \zeta} = \mathbf{0} : \frac{\partial l(\hat{\zeta}, \eta_\theta)}{\partial \eta_\theta} = \mathbf{0}$$

- ou seja, obter as estimativas de cada um dos conjuntos de parâmetros de forma alternada.

Utilizando desenvolvimentos apresentados em [http://www.ime.unicamp.br/~cnaber/Aula\\_TRI\\_Intro\\_MI685\\_2S\\_2017\\_Met\\_Est.pdf](http://www.ime.unicamp.br/~cnaber/Aula_TRI_Intro_MI685_2S_2017_Met_Est.pdf) e [http://www.ime.unicamp.br/~cnaber/Aula\\_TRI\\_Ana\\_Dados\\_Bay\\_MI685\\_2S\\_2017.pdf](http://www.ime.unicamp.br/~cnaber/Aula_TRI_Ana_Dados_Bay_MI685_2S_2017.pdf), temos

$$\begin{aligned}
 \mathbf{S}(\zeta_i) &= \sum_{k=1}^K \sum_{j=1}^{n_k} \left\{ \frac{1}{P(\mathbf{Y}_{.jk}|\zeta, \boldsymbol{\eta}_{\theta_k})} \frac{\partial P(\mathbf{Y}_{.jk}|\zeta, \boldsymbol{\eta}_{\theta_k})}{\partial \zeta_i} \right\} \\
 &= \sum_{k=1}^K \sum_{l \in Q_k} (\bar{r}_{ilk} - \bar{f}_{ilk} P_{ilk}) W_{ilk} \mathbf{h}_{ilk} \quad (1)
 \end{aligned}$$

$$\mathbf{I}(\zeta_i) = \sum_{k=1}^K \sum_{l \in Q_k} \bar{f}_{ilk} P_{ilk}^* Q_{ilk}^* \mathbf{h}_{ilk} \mathbf{h}'_{ilk} \quad (2)$$



Maiores detalhes podem ser encontrados em Andrade, Tavares e Valle (2000) e também em Azevedo (2003) (versão estendida). Em pelo menos uma dessas referências é possível encontrar as fórmulas necessárias para a implementação dos métodos de estimação.

$$S(\mu_{\theta_k}) = \sigma_k^{-2} \sum_{j=1}^{n_k} \sum_{l \in Q_k} g_{jk}^*(\bar{\theta}_{lk}) (\bar{\theta}_{lk} - \mu_{\theta_k})$$

$$S(\psi_{\theta_k}) = (2\sigma_k^4)^{-1} \sum_{j=1}^{n_k} \sum_{l \in Q_k} g_{jk}^*(\bar{\theta}_{lk}) \left\{ (\bar{\theta}_{lk} - \mu_{\theta_k})^2 - \psi_{\theta_k} \right\}$$

$$\mu_{\theta_k} = \widehat{\mu}_{.k}$$

$$\widehat{\psi}_{\theta_k} = \widehat{\sigma}_{.k}^2 + \widehat{\delta}_{.k}^2$$

Com

$$\widehat{\mu}_{.k} = \frac{1}{n_k} \sum_{j=1}^{n_k} \widehat{\mu}_{jk}; \widehat{\sigma}_{.k}^2 = \frac{1}{n_k} \sum_{j=1}^{n_k} \widehat{\sigma}_{jk}^2; \widehat{\delta}_{.k}^2 = \frac{1}{n_k} \sum_{j=1}^{n_k} (\widehat{\mu}_{jk} - \mu_{\theta_k})^2$$

e

$$\widehat{\mu}_{jk} = \sum_{l \in Q_k} \bar{\theta}_{lk} g_{jk}^* (\bar{\theta}_{lk})$$

$$\widehat{\sigma}_{jk}^2 = \sum_{l \in Q_k} (\bar{\theta}_{lk} - \widehat{\mu}_{jk})^2 g_{jk}^* (\bar{\theta}_{lk})$$

## Adaptação do Pseudo Algoritmo EM

De maneira análoga ao caso de uma única população, podemos usar uma adaptação do **algoritmo EM**, de modo que os parâmetros populacionais de cada população sejam estimados, em separado, no **Passo E**, e que os parâmetros de cada item sejam estimados, em separado, no **Passo M**, como mostrado a seguir

## Adaptação do Pseudo Algoritmo EM

### Passo E

- 1 Usar os pontos de quadratura,  $\bar{\theta}_{lk}$ , os pesos associados  $A_{lk}^{(t)}$ ,  $l = 1, \dots, q_k$ , as estimativas no passo anterior dos parâmetros dos itens  $\hat{\zeta}_i^{(t)}$ ,  $i = 1, \dots, I$ , e dos parâmetros populacionais,  $\hat{\mu}_k^{(t)}$  e  $\hat{\sigma}_k^{2(t)}$ ,  $k = 1, \dots, K$  para gerar  $g_{jk}^* (\bar{\theta}_{lk})^{(t)}$  e, posteriormente,  $\bar{r}_{ilk}^{(t)}$  e  $\bar{f}_{ilk}^{(t)}$ ,  $i = 1, \dots, I$  e  $l = 1, \dots, q_k$ .
- 2 Usar os pontos de quadratura e  $g_{jk}^* (\bar{\theta}_{lk})^{(t)}$  para obter  $\bar{\mu}_{.k}^{(t+1)}$ ,  $\bar{\sigma}_{.k}^{2(t+1)}$  e  $\bar{\delta}_{.k}^{(t+1)}$  através de (3), e posteriormente,  $\hat{\mu}_k^{(t+1)}$  e  $\hat{\sigma}_k^{2(t+1)}$  por (3) e (3).

## Passo M

- Com  $\hat{\mathbf{r}}^{(t)}$ ,  $\hat{\mathbf{f}}^{(t)}$  e  $\hat{\boldsymbol{\eta}}^{(t+1)}$  obtidos no Passo E, resolver as equações de estimação para  $\zeta_i$ ,  $i = 1, \dots, I$ , usando Newton-Raphson ou Escore de Fisher através das expressões de (1) a (2).
- A estimação dos traços latentes pode ser feita, essencialmente, da mesma forma como no caso de um único grupo.
- Para se obter a moda marginal relativa a posteriori dos parâmetros dos itens, pode-se utilizar uma abordagem semelhante àquela apresentada para o caso de um único grupo. Neste caso, deve-se atribuir distribuições a priori para os parâmetros dos itens (e, também, para os parâmetros populacionais).

# Métodos bayesianos plenos

## ■ Priors

- $\theta_{jk} \stackrel{i.i.d.}{\sim} N(\mu_{\theta_k}, \psi_{\theta_k})$ .
- $(a_i, b_i) \stackrel{i.i.d.}{\sim} N_2(\boldsymbol{\mu}_\zeta, \boldsymbol{\Psi}_\zeta) I_{(a_i > 0)}$ .
- $c_i \stackrel{i.i.d.}{\sim} \text{beta}(r, s)$ .
- $\mu_{\theta_k} \stackrel{i.i.d.}{\sim} N(\mu_0, \psi_0)$ .
- $\psi_{\theta_k} \stackrel{i.i.d.}{\sim} \text{IG}(\kappa_0, \nu_0)$ .

# Métodos bayesianos plenos

- Amostrador de Gibbs: posteriori via verossimilhança aumentada

$$\begin{aligned} p(\zeta, \theta, \eta_\theta | \mathbf{y}) &\propto p(\mathbf{y} | \mathbf{z}, \mathbf{u}) p(\mathbf{z} | \mathbf{a}, \mathbf{b}, \boldsymbol{\theta}) p(\mathbf{u} | \mathbf{c}) p(\zeta) p(\boldsymbol{\theta}) \\ &\propto \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} \exp \left\{ -0.5 (z_{ijk} - a_i \theta_j + b_i)^2 \right\} I_{\{z_{ijk}, y_{ijk}, u_{ijk}\}}^{(z)} \\ &\times \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} u_{ijk}^{c_{ij}} (1 - u_{ijk})^{1 - c_{ij}} I_{\{z_{ijk}, y_{ijk}, u_{ijk}\}}^{(u)} \left( \prod_{k=1}^K \prod_{j=1}^{n_k} e^{-0.5 \frac{(\theta_{jk} - \mu_{\theta_k})^2}{\psi_{\theta_k}}} \right) \\ &\times \left( \prod_{i=1}^I \exp \left[ -0.5 (\zeta_i - \mu_\zeta)^t \boldsymbol{\Psi}_\zeta^{-1} (\zeta_i - \mu_\zeta) \right] I_{(a_i > 0)} \right) \\ &\times \prod_{i=1}^I c_i^{r-1} (1 - c_i)^{s-1} \\ &\times \left\{ \prod_{k=1}^K \exp \left( -\frac{1}{\psi_{\theta_0}} (\mu_{\theta_k} - \mu_0)^2 \right) \right\} \left\{ \prod_{k=1}^K \psi_{\theta_k}^{-v_0/2-1} \exp \left( -\frac{\kappa_0}{2\psi_{\theta_k}} \right) \right\} \end{aligned}$$

# Métodos bayesianos plenos

- Metropolis-Hastings (posteriori via verossimilhança original):

$$\begin{aligned} p(\zeta, \theta | \mathbf{y}) &\propto L(\zeta, \theta) p(\zeta) p(\theta) \\ &\propto \left( \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} p_{ijk}^{y_{ijk}} (1 - p_{ijk})^{1 - y_{ijk}} \right) \left( \prod_{k=1}^K \prod_{j=1}^{n_k} e^{-0.5 \frac{(\theta_{jk} - \mu_{\theta_k})^2}{\psi_{\theta_k}}} \right) \\ &\times \left( \prod_{i=1}^I \exp \left[ -0.5 (\zeta_i - \mu_{\zeta})^t \Psi_{\zeta}^{-1} (\zeta_i - \mu_{\zeta}) \right] I_{(a_i > 0)} \right) \\ &\times \prod_{i=1}^I c_i^{r-1} (1 - c_i)^{s-1} \\ &\times \left\{ \prod_{k=1}^K \exp \left( -\frac{1}{\psi_{\theta_0}} (\mu_{\theta_k} - \mu_0)^2 \right) \right\} \left\{ \prod_{k=1}^K \psi_{\theta_0}^{-v_0/2-1} \exp \left( -\frac{\kappa_0}{2\psi_{\theta_k}} \right) \right\} \end{aligned}$$



# Condicionais completas

- Sob a verossimilhança original, independentemente da escolha da priori, as condicionais completas não possuem forma conhecida (distribuição catalogada).
- Sob a verossimilhança aumentada, consoante a priori escolhida, é possível de obter todas as condicionais completas com forma conhecida (distribuição catalogada).

# Modelo de 2 parâmetros

Denotando por  $(\cdot)$  o conjunto de todos os outros parâmetros, temos que os passos do Amostrador de Gibbs são dados por :

- Iniciar a simulação das cadeias através de valores convenientes.
- Simular  $Z_{ijk}$  de forma mutuamente independente através de  $Z_{ij} | (\cdot)$  (normais univariadas truncadas).
- Simular  $\theta_{jk}$  de forma mutuamente independente através de  $\theta_j | (\cdot)$  (normais univariadas).
- Simular  $\eta_{\theta_k}$  de forma mutuamente independente através de  $\eta_{\theta_k} | (\cdot)$  (normais univariadas e inversas gama).
- Simular  $\zeta_i$  de forma mutuamente independente através de  $\zeta_i | (\cdot)$  (normais bivariadas truncadas).

# Modelo de 3 parâmetros, Sahu ou Beguin and Glas

Denotando por  $(\cdot)$  o conjunto de todos os outros parâmetros, temos que os passos do Amostrador de Gibbs são dados por :

- Iniciar a simulação das cadeias através de valores convenientes.
- Simular  $Z_{ijk}$  de forma mutuamente independente através de  $Z_{ijk} | (\cdot)$  (normais univariadas truncadas).
- Simular  $U_{ijk}$  de forma mutuamente independente através de  $U_{ijk} | (\cdot)$  (bernoullis univariadas).
- Simular  $\theta_{jk}$  de forma mutuamente independente através de  $\theta_{jk} | (\cdot)$  (normais univariadas).
- Simular  $\eta_{\theta_k}$  de forma mutuamente independente através de  $\eta_{\theta_k} | (\cdot)$  (normais univariadas e inversas gama).
- Simular  $\zeta_i$  de forma mutuamente independente através de  $\zeta_i | (\cdot)$  (normais bivariadas truncadas).
- Simular  $c_i$  de forma mutuamente independente através de  $c_i | (\cdot)$  (betas univariadas)

# Condicionais completas: AG - M2P

- **Passo 1** : Simular as variáveis não observáveis utilizando

$Z_{ijk} | (\cdot) \sim N(a_i \theta_{jk} - b_i, 1) \mathbb{1}_{\{y_{ijk}, z_{ijk}\}}$ , em que

$\mathbb{1}_{\{y_{ijk}, z_{ijk}\}} = I_{\{z_{ijk} > 0, y_{ijk} = 1\}} + I_{\{z_{ijk} < 0, y_{ijk} = 0\}}$  de forma mutuamente independente.

- **Passo 2** : Simular os traços latentes utilizando  $\theta_{jk} | (\cdot) \sim N(\hat{\psi}_{\theta_{jk}} \hat{\theta}_{jk}, \hat{\psi}_{\theta_{jk}})$ , mutuamente independentes, em que ( $I_{jk}$  conjunto de itens que respondidos pelo indivíduo  $j$  do grupo  $k$ )

$$\hat{\theta}_{jk} = \sum_{i \in I_{jk}} z_{ijk} a_i + \sum_{i \in I_{jk}} a_i b_i + \frac{\mu_{\theta_{jk}}}{\psi_{\theta_{jk}}},$$
$$\hat{\psi}_{\theta_{jk}} = \left( \frac{1}{\psi_{\theta_{jk}}} + \sum_{i \in I_{jk}} a_i^2 \right)^{-1}.$$

## Condicionais completas: AG - M2P

- **Passo 3** : Simular os parâmetros dos itens através de

$\zeta_i | (\cdot) \sim N(\widehat{\Psi}_{\zeta_i} \widehat{\zeta}_i, \widehat{\Psi}_{\zeta_i})$ , de forma mutuamente independente, em que

$$\begin{aligned}\widehat{\zeta}_i &= \mathbf{H}_{i..}^t \mathbf{z}_{i..} + \Psi_{\zeta}^{-1} \boldsymbol{\mu}_{\zeta}, \\ \widehat{\Psi}_{\zeta_i} &= \left( \mathbf{H}_{i..}^t \mathbf{H}_{i..} + \Psi_{\zeta}^{-1} \right)^{-1}, \\ \mathbf{H}_{i..} &= [\boldsymbol{\theta} \quad \mathbf{1}] \bullet \mathbb{1}_i,\end{aligned}$$

$\mathbb{1}_i$  é uma matriz indicadora de dimensão  $n \times 2$  cuja linhas assumem o valor do vetor com 0's ou 1's consoante o indivíduo  $j$ , do grupo  $k$  responde ou não ao item  $i$  e  $\bullet$  denota o produto de Hadamard, o qual multiplica de forma pontual os elementos correspondentes de matrizes de mesma dimensão, Horn and Johnson (1991).

## Condicionalis completas: AG - M2P

- **Passo 4** : Simular as médias populacionais através de

$\mu_{\theta_k} | (\cdot) \sim N(\hat{\psi}_{\mu_k} \hat{\mu}_{\theta_k}, \hat{\psi}_{\mu_k})$ , de forma mutuamente independente, em que

$$\hat{\mu}_{\theta_k} = \sum_{j=1}^{n_k} \frac{\theta_{jk}}{\psi_{\theta_k}} + \frac{\mu_0}{\psi_0},$$
$$\hat{\psi}_{\mu_k} = \left( \frac{n_k}{\psi_{\theta_k}} + \frac{1}{\psi_0} \right)^{-1}.$$

- **Passo 5** : Simular as variâncias populacionais  $\psi_{\theta_k} | (\cdot) \sim IG(\hat{v}_k, \hat{\kappa}_k)$ , de forma mutuamente independente, em que

$$\hat{v}_k = \frac{n_k + v_0}{2},$$
$$\hat{\kappa}_k = \frac{\sum_{j=1}^{n_k} (\theta_{jk} - \mu_{\theta_k})^2 + \kappa_0}{2},$$

# Modelo de 3 parâmetros, Sahu e Beguin and Glas

- Beguin and Glas: Simular  $U_{ijk}$  de Bernoullis como em [http://www.ime.unicamp.br/~cnaber/Aula\\_TRI\\_Intro\\_MI685\\_2S\\_2017\\_Full\\_MCMC.pdf](http://www.ime.unicamp.br/~cnaber/Aula_TRI_Intro_MI685_2S_2017_Full_MCMC.pdf), página 16. Depois, dado  $U_{ijk}$  simular  $Z_{ijk} \stackrel{ind.}{\sim} N(a_i\theta_{jk} - b_i, 1)I_{\{z_{ijk}, u_{ijk}\}}$ , em que  $I_{\{z_{ijk}, u_{ijk}\}} = I_{\{z_{ijk} \geq 0, u_{ijk} = 1\}} + I_{\{z_{ijk} < 0, u_{ijk} = 0\}}$ .

# Condicionalis completas: AG - M3P

- Variáveis Z e U (slides seguintes)
- Traços latentes e parâmetros dos itens: como para o modelo M2P.
- Parâmetro de acerto casual:

$c_i | (.) \text{beta}(r + \sum_{k=1}^K \sum_{j \in I_{ik}} u_{ijk}, s + \sum_{k=1}^K \sum_{j \in I_{ik}} (1 - u_{ijk}))$  (Sahu) e  
 $c_i | (.) \text{beta}(s + \sum_{k=1}^K \sum_{j \in I_{ik} | u_{ijk}=0} y_{ijk}, r + \sum_{k=1}^K \sum_{j \in I_{ik}} u_{ijk} - \sum_{j \in I_i | u_{ijk}=0} y_{ijk})$  (Beguin and Glas), em que  $I_{ik}$  é o conjunto de indivíduos do grupo  $k$ , que respondeu ao item  $i$ .



# Modelo de 3 parâmetros, Sahu e Beguin and Glas

- Sahu: Proceder da seguinte forma:

Simulando	$Y_{ijk}$	$U_{ijk}$	$Z_{ijk}$
$Z_{ijk}$	0	0	$N(a_i\theta_{jk} - b_i, 1)I_{(z_{ijk} < 0)}$
$Z_{ijk}$	1	0	$N(a_i\theta_{jk} - b_i, 1)I_{(z_{ijk} > 0)}$
$Z_{ijk}$	1	1	$N(a_i\theta_{jk} - b_i, 1)$
$U_{ijk}$	0	0	qualquer valor
$U_{ijk}$	1	1	$< 0$
$U_{ijk}$	1	Bernoulli( $c_i$ )	$\geq 0$

# Algoritmo de Metropolis-Hastings dentro do Gibbs

- Simular  $\theta_{jk}^{(t)} \sim g(\theta_j | \zeta^{(t-1)}, \mathbf{y}_{..})$  (condicional completa), para  $j = 1, \dots, n$   
independentemente, considerando como priori uma  $N(\theta_{jk} | \mu_{\theta_{jk}}, \sigma_{\theta_{jk}}^2)$  através de :

- (a) Simular  $\theta_{jk}^{(*)} \sim N(\theta_{jk}^{(t-1)}, \psi_{\theta_{jk}})$
- (b) Calcular o vetor de probabilidades de aceitação  $\theta_{jk}^{(t)} = \theta_{jk}^{(*)}$

$$\pi_j(\theta_{jk}^{(t-1)}, \theta_{jk}^{(*)}) = \min \left\{ \frac{L(\zeta^{(t-1)}, \theta_{jk}^{(*)}) \exp \left\{ -\frac{(\theta_{jk}^{(*)} - \mu_{\theta_{jk}})^2}{2\sigma_{\theta_{jk}}^2} \right\}}{L(\zeta^{(t-1)}, \theta_{jk}^{(t-1)}) \exp \left\{ -\frac{(\theta_{jk}^{(t-1)} - \mu_{\theta_{jk}})^2}{2\sigma_{\theta_{jk}}^2} \right\}}, 1 \right\}$$

- (c) Aceitar cada  $\theta_{jk}^{(t)} = \theta_{jk}^{(*)}$  com probabilidade  $\pi_{jk}$ , caso contrário  $\theta_{jk}^{(t)} = \theta_{jk}^{(t-1)}$

# Algoritmo de Metropolis-Hastings dentro do Gibbs (M2P)

- Simular  $\zeta_i^{(t)} \sim g(\zeta_i | \theta^{(t)}, \mathbf{y}_{..})$  (condicional completa), para  $i = 1, \dots, I$  independentemente, considerando como priori

$f(\zeta_i | \tau_i) \equiv \log - normal(a_i | \mu_{a_i}, \sigma_{a_i}^2) \times N(b_i | \mu_{b_i}, \sigma_{b_i}^2)$  através de :

- Simular  $a_i^{(*)} \sim \log - normal(a_i^{(*)} | a_i^{(t-1)}, \psi_{a_i})$ ,  
 $b_i^{(*)} \sim N(b_i^{(*)} | b_i^{(t-1)}, \psi_{b_i})$
- Calcular o vetor de probabilidades de aceitação  $\zeta^{(t)} = \zeta^{(*)}$

$$\pi_i(\zeta^{(t-1)}, \zeta^{(*)}) = \min\{R_{\zeta_i}, 1\}$$

# Algoritmo de Metropolis-Hastings dentro do Gibbs (M2P)

com

$$R_{\zeta_i} = \frac{L(\zeta_i^{(*)}, \theta^{(t)}) \exp \left\{ -\frac{(\ln a_i^{(*)} - \mu_{a_i})^2}{2\sigma_{a_i}^2} \right\} \exp \left\{ -\frac{(b_i^{(*)} - \mu_{b_i})^2}{2\sigma_{b_i}^2} \right\}}{L(\zeta_i^{(t-1)}, \theta^{(t)}) \exp \left\{ -\frac{(\ln a_i^{(t-1)} - \mu_{a_i})^2}{2\sigma_{a_i}^2} \right\} \exp \left\{ -\frac{(b_i^{(t-1)} - \mu_{b_i})^2}{2\sigma_{b_i}^2} \right\}} \\ \times \frac{[a_i^{(*)}]^2 \exp \left\{ -\frac{(\ln a_i^{(*)} - a_i^{(t-1)})^2}{2\psi_{a_1}^2} \right\}}{[a_i^{(t-1)}]^2 \exp \left\{ -\frac{(\ln a_i^{(t-1)} - a_i^{(*)})^2}{2\psi_{a_1}^2} \right\}}$$

(c) Aceitar cada  $\zeta_i^{(t)} = \zeta_i^{(*)}$  com probabilidade  $\pi_i$ , caso contrário  $\zeta_i^{(t)} = \zeta_i^{(t-1)}$

# Algoritmo de Metropolis-Hastings dentro do Gibbs (M3P)

- Simular  $\zeta_i^{(t)} \sim g(\zeta_i | \theta^{(t)}, \mathbf{y}_{..})$  (condicional completa), para  $i = 1, \dots, I$  independentemente, considerando como priori

$$f(\zeta_i | \boldsymbol{\tau}_i) \equiv \log - normal(a_i | \mu_{a_i}, \sigma_{a_i}^2) \times N(b_i | \mu_{b_i}, \sigma_{b_i}^2) \times Beta(c_i | \alpha_i - 1, \beta_i - 1)$$

através de :

- (a) Simular  $a_i^{(*)} \sim \log - normal(a_i^{(*)} | a_i^{(t-1)}, \psi_{a_i})$ ,  
 $b_i^{(*)} \sim N(b_i^{(*)} | b_i^{(t-1)}, \psi_{b_i})$  e  
 $c_i^{(*)} \sim Beta(c_i^{(*)} | s_i^{(t-1)} - 1, r_i^{(t-1)} - 1)$

- (b) Calcular o vetor de probabilidades de aceitação  $\zeta^{(t)} = \zeta^{(*)}$

$$\pi_i(\zeta^{(t-1)}, \zeta^{(*)}) = \min\{R_{\zeta_i}, 1\}$$

# Algoritmo de Metropolis-Hastings dentro do Gibbs (M3P)

com

$$\begin{aligned}
 R_{\zeta_i} = & \frac{L(\zeta_i^{(*)}, \theta^{(t)}) \exp \left\{ -\frac{(\ln a_i^{(*)} - \mu_{a_i})^2}{2\sigma_{a_i}^2} \right\} \exp \left\{ -\frac{(b_i^{(*)} - \mu_{b_i})^2}{2\sigma_{b_i}^2} \right\}}{L(\zeta_i^{(t-1)}, \theta^{(t)}) \exp \left\{ -\frac{(\ln a_i^{(t-1)} - \mu_{a_i})^2}{2\sigma_{a_i}^2} \right\} \exp \left\{ -\frac{(b_i^{(t-1)} - \mu_{b_i})^2}{2\sigma_{b_i}^2} \right\}} \\
 & \times \frac{[c_i^{(*)}]^{\alpha_i - 2 + s_i^{(t-1)} - 1} [a_i^{(*)}]^2 \exp \left\{ -\frac{(\ln a_i^{(*)} - a_i^{(t-1)})^2}{2\psi_{a_1}^2} \right\}}{[c_i^{(t-1)}]^{\alpha_i - 2 + s_i^{(*)} - 1} [a_i^{(t-1)}]^2 \exp \left\{ -\frac{(\ln a_i^{(t-1)} - a_i^{(*)})^2}{2\psi_{a_1}^2} \right\}} \\
 & \times \frac{\beta(s_i^{(*)} - 1, r_i^{(*)} - 1) [1 - c_i^{(*)}]^{\beta_i - 2 + r_i^{(t-1)} - 1}}{\beta(s_i^{(t-1)} - 1, r_i^{(t-1)} - 1) [1 - c_i^{(t-1)}]^{\beta_i - 2 + r_i^{(*)} - 1}}
 \end{aligned}$$

(c) Aceitar cada  $\zeta_i^{(t)} = \zeta_i^{(*)}$  com probabilidade  $\pi_i$ , caso contrário  $\zeta_i^{(t)} = \zeta_i^{(t-1)}$

# Algoritmo de Metropolis-Hastings dentro do Gibbs (M3P)

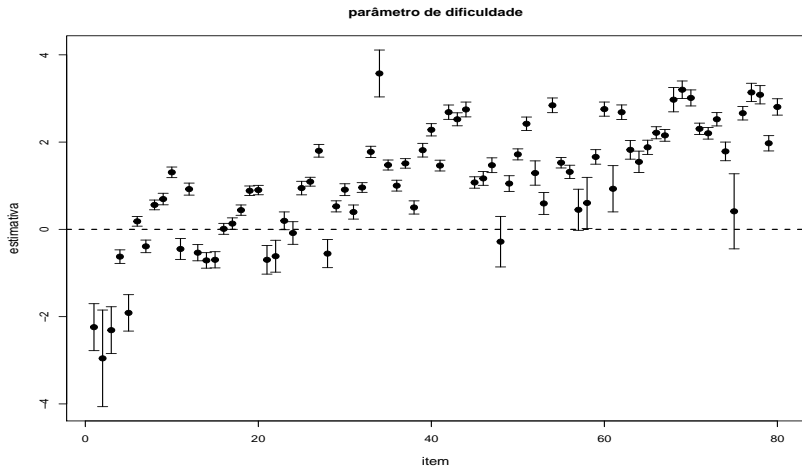
- Os parâmetros populacionais podem ser simulados tal como para o amostrador de Gibbs.
- Verificação da qualidade de ajuste do modelo:
  - Semelhante ao caso de um único grupo. Além dos níveis global, indivíduos, e item, temos agora o nível grupo.
  - Nível grupo: distribuição dos escores, proporção (e probabilidades) de respostas corretas, p-valores etc.

# Análise de dados

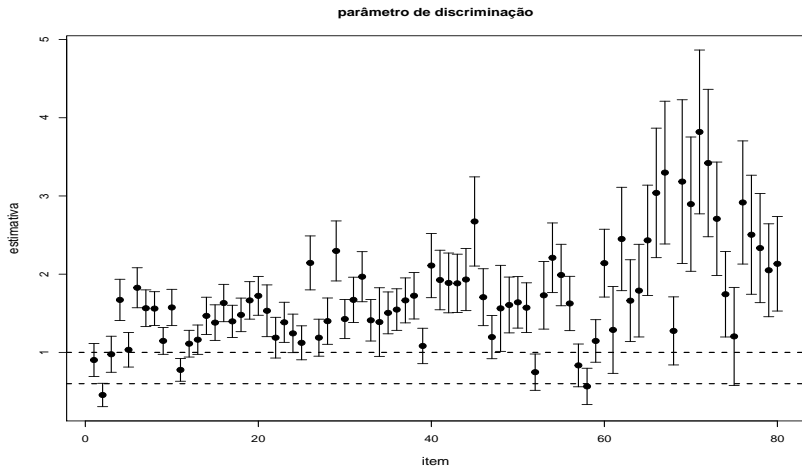
- Resultados 1: Estimação dos itens e parâmetros populacionais via MVM e dos traços latentes via EAP via mirt. Prioris para os traços latentes:  $\theta_{jk} \overset{i.i.d.}{\sim} N(\mu_{\theta_k}, \psi_{\theta_k})$ .
- Resultados 2: Estimação via métodos bayesianos plenos via WinBUGS. Prioris:  $a_i \overset{i.i.d.}{\sim} LN(0, 4891; 12, 0531)$ ,  $b_i \overset{i.i.d.}{\sim} N(0; 0, 5)$ ,  $c_i \overset{i.i.d.}{\sim} \text{beta}(100, 340)$ ,  $\theta_{jk} \overset{i.i.d.}{\sim} N(\mu_{\theta_k}, \psi_{\theta_k})$ .



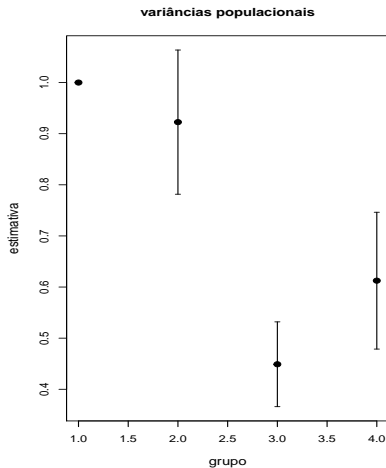
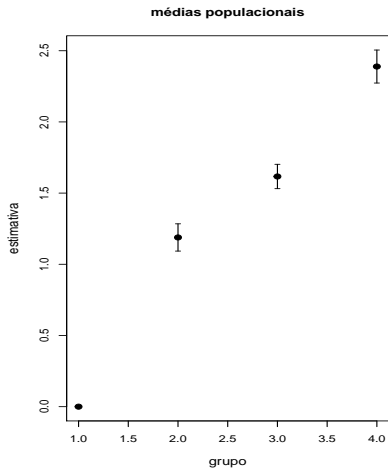
# mirt : parâmetros dos itens



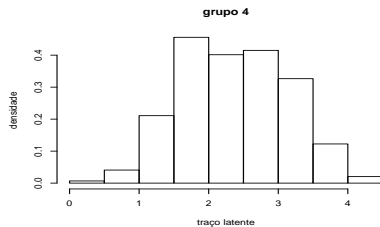
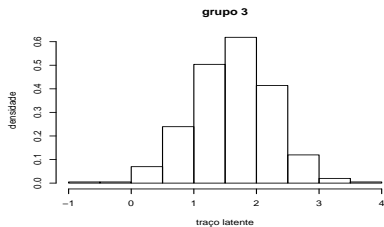
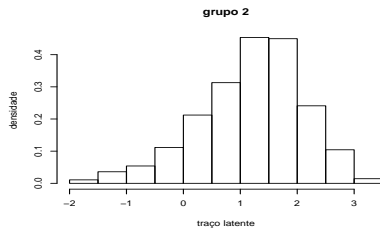
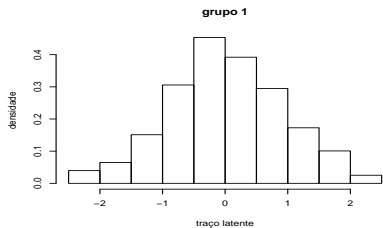
# mirt: parâmetros dos itens



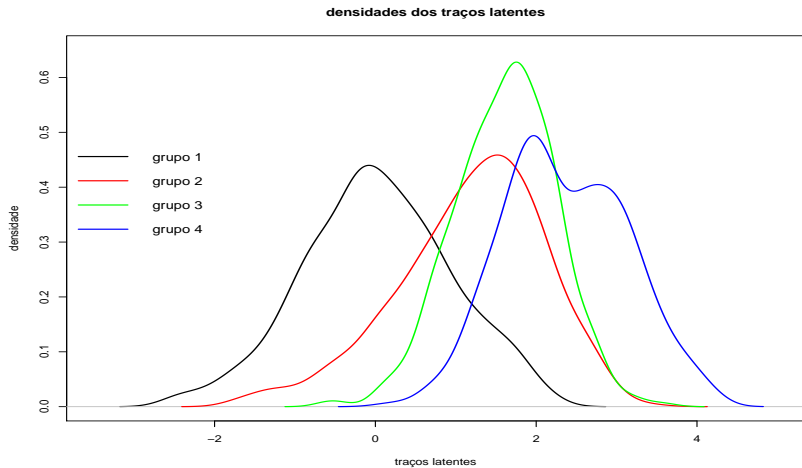
# mirt: parâmetros populacionais



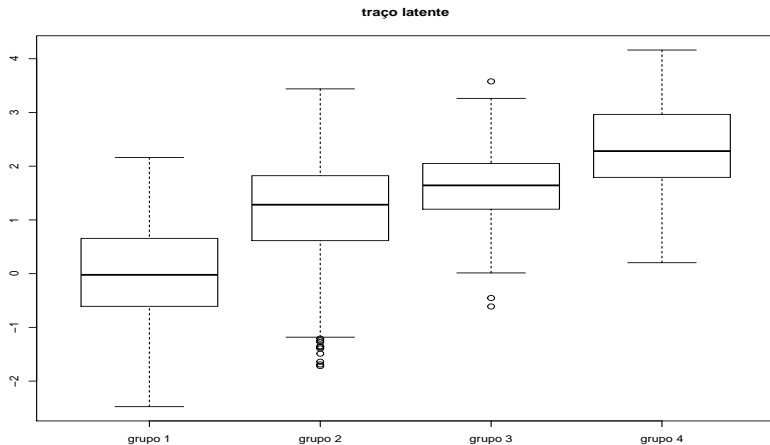
# mirt: traços latentes



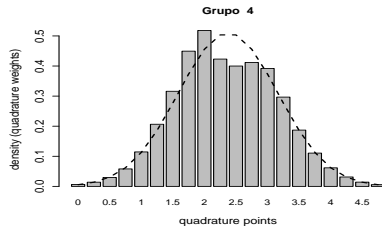
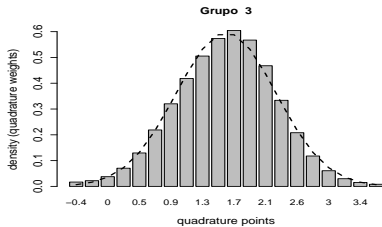
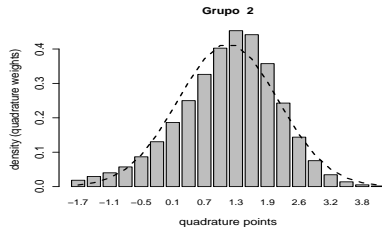
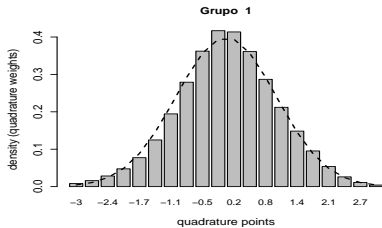
# mirt: traços latentes



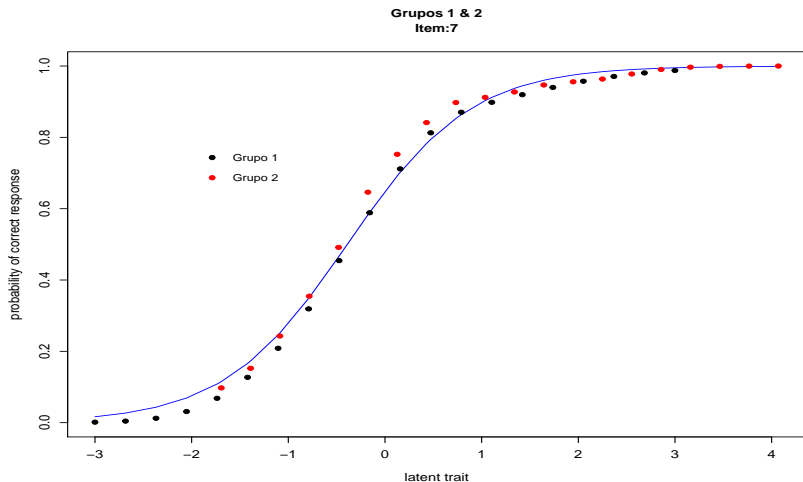
# mirt: traços latentes



# mirt: distribuições latentes

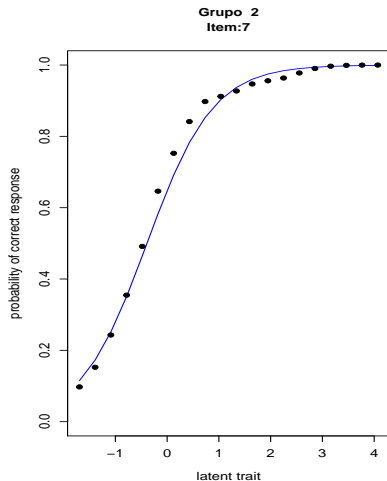
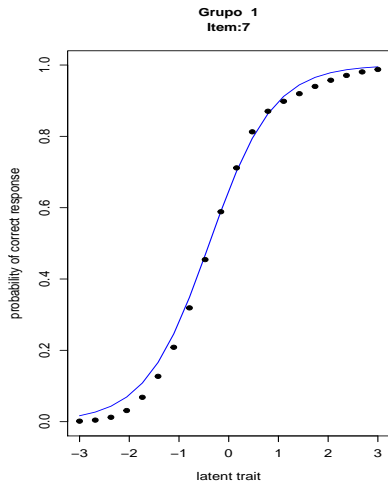


# mirt: probabilidades de acerto

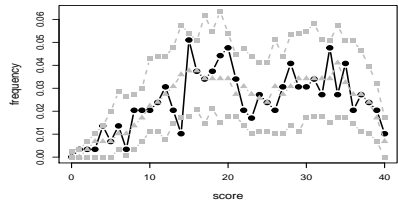
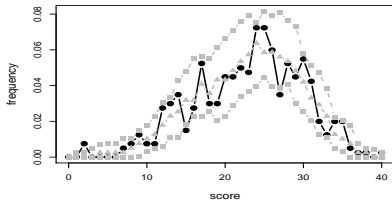
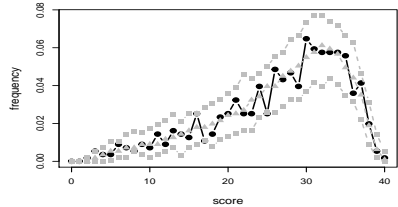
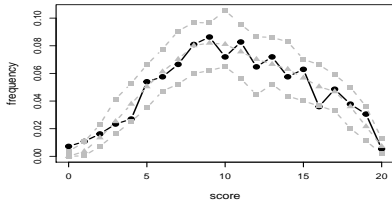




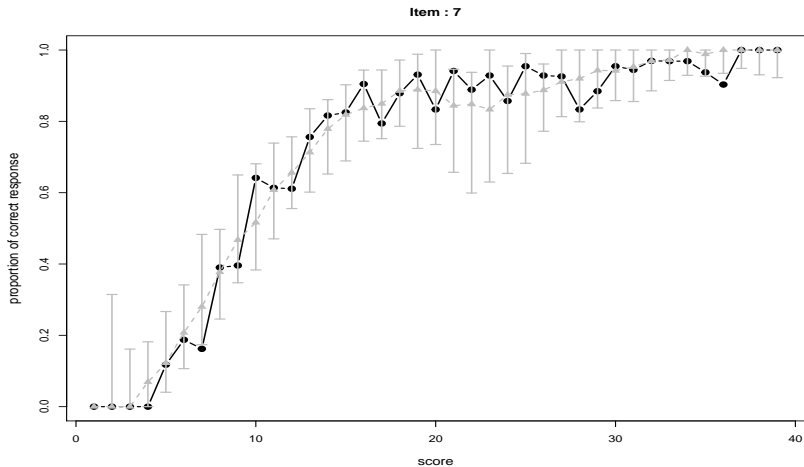
# mirt: probabilidades de acerto



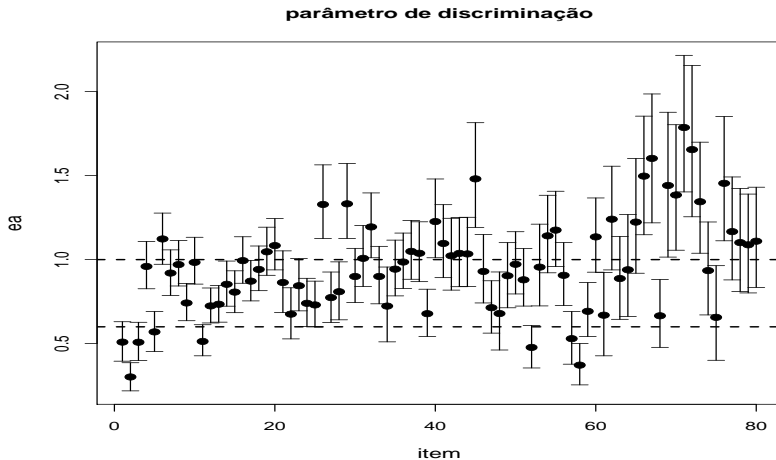
# mirt: distribuições dos escores



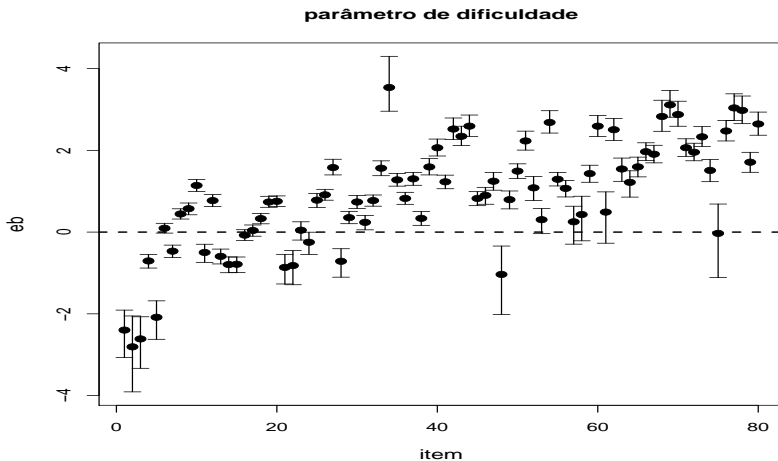
# mirt: proporções de acerto



# MCMC : parâmetros dos itens

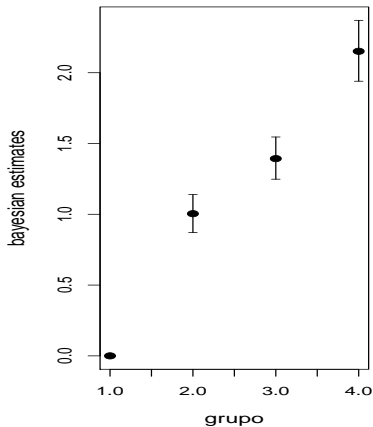


# MCMC: parâmetros dos itens

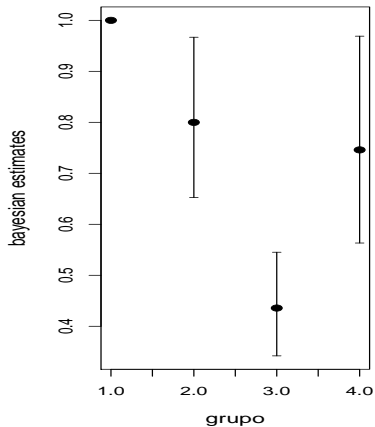


# MCMC: parâmetros populacionais

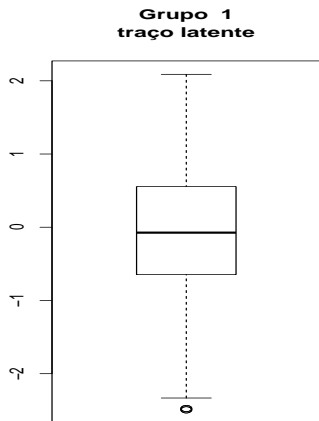
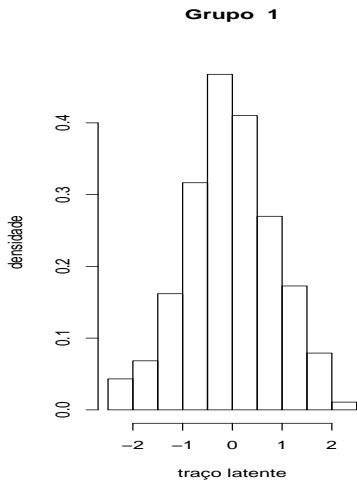
**médias populacionais**



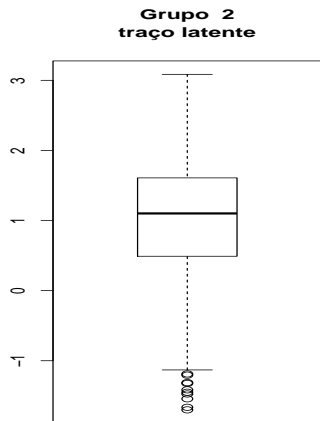
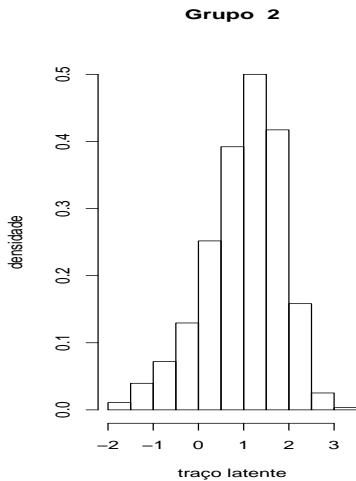
**variâncias populacionais**



# MCMC: traços latentes

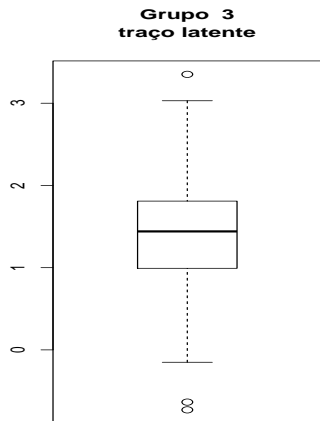
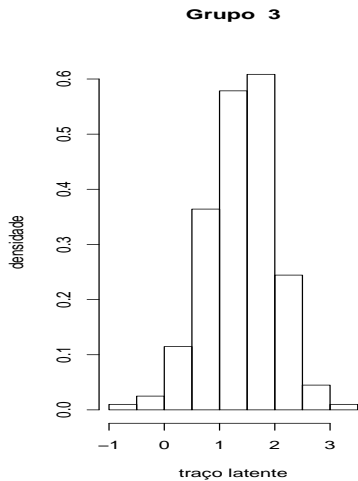


# MCMC: traços latentes

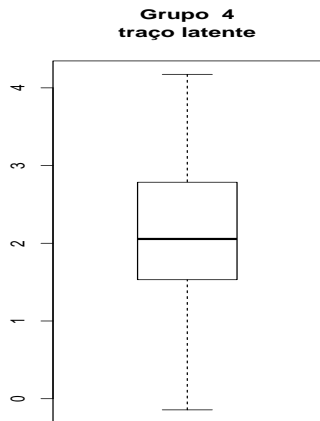
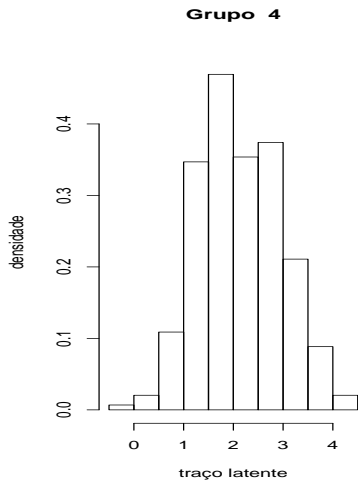




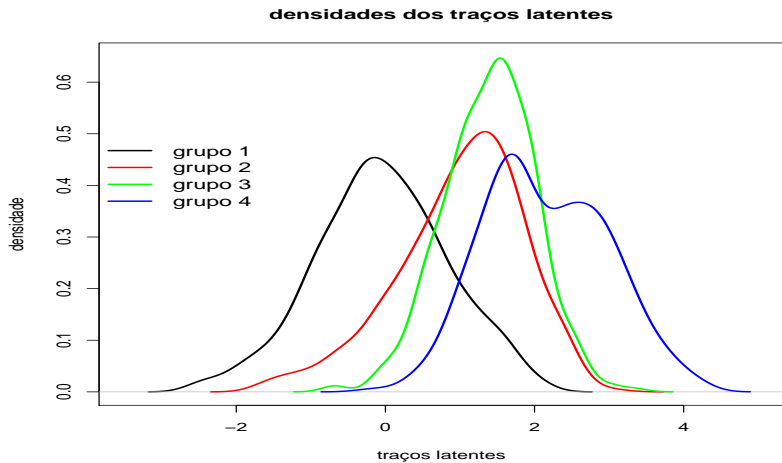
# MCMC: traços latentes



# MCMC: traços latentes

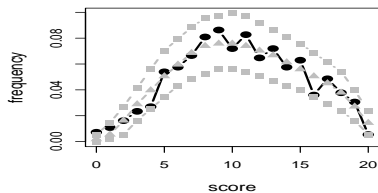


# MCMC: traços latentes

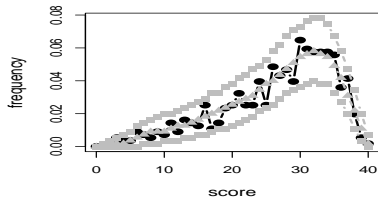


# MCMC: distribuições dos escores

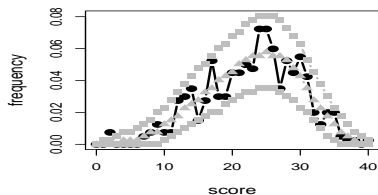
**Grupo 1**



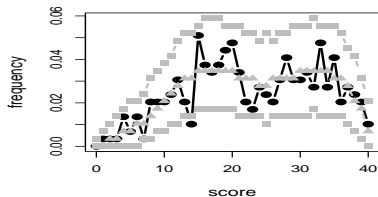
**Grupo 2**



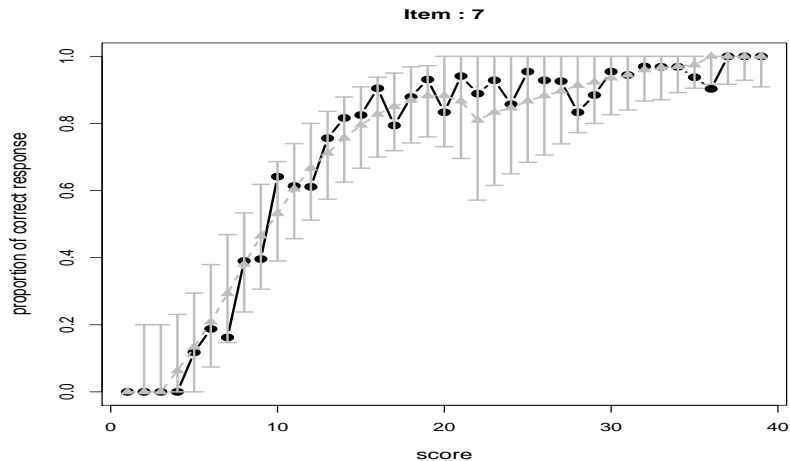
**Grupo 3**



**Grupo 4**



# MCMC: proporções de respostas corretas



# MCMC: proporções de respostas corretas

