Teoria de Resposta ao Item

Métodos de estimação: bayesianos e marginal-perfilados bayesianos

Caio L. N. Azevedo, IMECC/Unicamp

Inferência bayesiana

- Sejam A e $B_1, B_2, \ldots$ eventos associados à um mesmo espaço de probabilidade $(\Omega, \mathcal{A}, \mathcal{P})$
- Pelo teorema da probabilidade total, temos que $P(A) = \sum_{i=1}^{\infty} P(A|B_i)P(B_i)$
- Pelo teorema de Bayes, temos que

$$P(B_i|A) = \frac{P(A \cap B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^{\infty} P(A|B_i)P(B_i)}$$

Inferência bayesiana

- Seja $\mathbf{X}$ um vetor aleatório com densidade $p(\mathbf{x}|\boldsymbol{\beta})$ (verossimilhança) em que β é um vetor de parâmetros (deconhecidos), $\boldsymbol{\beta} \subset \mathcal{R}^p$.
- Adimita que a ingnorância a respeito de $\boldsymbol{\beta}$ pode ser traduzida por uma densidade $p(\boldsymbol{\beta}|\boldsymbol{\eta}_{\boldsymbol{\beta}}) \equiv p(\boldsymbol{\beta})$. As componentes de $\boldsymbol{\eta}_{\boldsymbol{\beta}}$ são conhecidas como hiperparâmetros.
- A inferência bayesiana baseia-se na utilização da distribuição a posteriori de $\boldsymbol{\beta}$, ou seja:

$$
\begin{aligned}
p(\boldsymbol{\beta}|\mathbf{x}) &= \frac{p(\mathbf{x},\boldsymbol{\beta})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\boldsymbol{\beta})p(\boldsymbol{\beta})}{\int p(\mathbf{x}|\boldsymbol{\beta})p(\boldsymbol{\beta})d\boldsymbol{\beta}} \\
&\propto p(\mathbf{x}|\boldsymbol{\beta})p(\boldsymbol{\beta})
\end{aligned}
$$

- As estimativas de MV não estão definidas para escores nulos ou perfeitos.
- Escore nulo $L(\theta_j) = \prod_{i=I}(1 - P_{ij})$
- Escore perfeito $L(\theta_j) = \prod_{i=I}(P_{ij})$
- Solução: utilização de métodos bayesianos.

Parâmetros dos itens conhecidos

$$
\begin{aligned}
p(\boldsymbol{\theta}|\mathbf{y},\zeta) &= \frac{p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\boldsymbol{\theta})}{\int p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\boldsymbol{\theta})d\boldsymbol{\theta}} \propto p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\boldsymbol{\theta}) \\
&= \left(\prod_{i=1}^{I}\prod_{j=1}^{n} P_{ij}^{y_{ij}} Q_{ij}^{1-y_{ij}}\right)\left(\prod_{j=1}^{n} p(\theta_j)\right) \\
&= \prod_{j}^{n}\left(p(\mathbf{y}_{.j}|\theta_j,\zeta)p(\theta_j)\right) = \prod_{j=1}^{n} p(\theta_j|\mathbf{y}_{.j},\zeta_i)
\end{aligned}
$$

Moda a posteriori

- Maximizar $\ln p(\boldsymbol{\theta}|\mathbf{y},\zeta) = \sum_{j=1}^{n} \ln p(\theta_j|\mathbf{y}_{.j},\zeta)$

$$S(\theta_j)_B = \frac{\partial l_j^*(\theta_j)}{\partial\theta_j} = \frac{\partial l(\theta_j)}{\partial\theta_j} + \frac{\partial \ln g(\theta_j|\boldsymbol{\eta})}{\partial\theta_j}.$$

$$S(\theta_j)_B = \sum_{i=1}^{I}(y_{ij} - P_{ij}) - \frac{\theta_j - \mu_\theta}{\sigma_\theta^2}.$$

- Informação de Fisher

$$I(\theta_j) = \sum_{i=1}^{I} P_{ij}Q_{ij} + \frac{1}{\sigma_\theta^2}$$

- Erro-padrão a posteriori (assintótico)

$$\mathcal{EP}(\theta_j) = \sqrt{I(\widehat{\theta}_j)^{-1}}$$

Média a posteriori

$$E\left[\theta_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}\right] = \frac{\displaystyle\int_{\Re}\theta p(\mathbf{y}_{.j}|\theta_j,\zeta)p(\theta_j)d\theta}{\displaystyle\int_{\Re}p(\mathbf{y}_{.j}|\theta_j,\zeta)p(\theta_j)d\theta}.$$

$$\begin{aligned} E\left[\theta_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}\right] &\approx E\left[\overline{\theta}_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}\right] = \frac{\sum_{l=1}^{q}\overline{\theta}_l P(\mathbf{Y}_{.j}|\overline{\theta}_l,\zeta)g(\overline{\theta}_l|\boldsymbol{\eta})}{\sum_{l=1}^{q}P(\mathbf{Y}_{.j}|\overline{\theta}_l,\zeta)g(\overline{\theta}_l|\boldsymbol{\eta})} \\ &= \frac{\sum_{l=1}^{q}\overline{\theta}_l P(\mathbf{Y}_{.j}|\overline{\theta}_l,\zeta)A_l}{\sum_{l=1}^{q}P(\mathbf{Y}_{.j}|\overline{\theta}_l,\zeta)A_l}. \end{aligned}$$

Variância a posteriori

$$Var\left[\theta_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}\right] = \frac{\int_{\Re}\left[\theta - \boldsymbol{E}\left(\theta\right)\right]^2 P(\mathbf{Y}_{.j}|\theta,\zeta)g(\theta|\boldsymbol{\eta})d\theta}{\int_{\Re} P(\mathbf{Y}_{.j}|\theta,\zeta)g(\theta|\boldsymbol{\eta})d\theta}$$

Em termos dos pontos de quadratura, temos que

$$Var\left[\overline{\theta}_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}\right] = \frac{\sum_{l=1}^{q}\left\{\overline{\theta}_l - \boldsymbol{E}\left[\overline{\theta}_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}\right]\right\}^2 P(\mathbf{Y}_{.j}|\overline{\theta}_l,\zeta)g(\overline{\theta}_l|\boldsymbol{\eta})}{\sum_{l=1}^{q} P(\mathbf{Y}_{.j}|\overline{\theta}_l,\zeta)g(\overline{\theta}_l|\boldsymbol{\eta})}$$

Erro-padrão a posteriori

$$\mathcal{EP}[\overline{\theta}_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}] = \sqrt{Var\left[\overline{\theta}_j|\mathbf{y}_{.j},\zeta,\boldsymbol{\eta}\right]}$$

- Máxima Verossimilhança - MV :
  - ⊕ Para testes "longos" produz estimadores não viciados;
  - ⊖ Não está definido para alguns padrões de resposta.
- Bayesiano - EAP :
  - ⊕ Definido para qualquer padrão de resposta;
  - ⊕ Possui o menor erro médio;
  - ⊕ Sua implementação é relativamente simples;
  - ⊖ Viciado;
  - ⊖ Necessidade de uma distribuição a priori para θ.
- Bayesiano - MAP :
  - ⊕ Definido para qualquer padrão de resposta;
  - ⊖ Viciado.
  - ⊖ Exige cálculos mais complexos do que o método de MV;
  - ⊖ Necessidade de uma distribuição a priori para θ.

Traços latentes conhecidos

$$
\begin{aligned}
p(\mathbf{b}|\mathbf{y},\zeta) &= \frac{p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\mathbf{b})}{\int p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\mathbf{b})d\mathbf{b}} \propto p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\mathbf{b}) \\
&= \left(\prod_{i=1}^{I}\prod_{j=1}^{n} P_{ij}^{y_{ij}} Q_{ij}^{1-y_{ij}}\right)\left(\prod_{i=1}^{n} p(b_i)\right) \\
&= \prod_{i}^{I}\left(p(\mathbf{y}_{i.}|\theta_j,\zeta)p(b_i)\right) = \prod_{i=1}^{I} p(b_i|\mathbf{y}_{i.},\boldsymbol{\theta})
\end{aligned}
$$

- Que priori utilizar para os parâmetros de dificuldade?
- $b_i \sim N(\mu_b, \sigma_b^2)$ (dificuldades são "habilidades").

Traços latentes conhecidos

$$
\begin{aligned}
p(\mathbf{b}|\mathbf{y},\zeta) &= \frac{p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\mathbf{b})}{\int p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\mathbf{b})d\mathbf{b}} \propto p(\mathbf{y}|\boldsymbol{\theta},\zeta)p(\mathbf{b}) \\
&= \left(\prod_{i=1}^{I}\prod_{j=1}^{n} P_{ij}^{y_{ij}} Q_{ij}^{1-y_{ij}}\right)\left(\prod_{i=1}^{n} p(b_i)\right) \\
&= \prod_{i}^{I} \left(p(\mathbf{y}_{i.}|\theta_j,\zeta)p(b_i)\right) = \prod_{i=1}^{I} p(b_i|\mathbf{y}_{i.},\boldsymbol{\theta})
\end{aligned}
$$

- Que priori utilizar para os parâmetros de dificuldade?
- $b_i \sim N(\mu_b, \sigma_b^2)$ (dificuldades são "habilidades").

Moda a posteriori

- Maximizar $\ln p(\mathbf{b}|\mathbf{y},\boldsymbol{\theta}) = \sum_{i=1}^{I} \ln p(b_i|\mathbf{y}_{.j},\boldsymbol{\theta})$

$$
\begin{aligned}
S\left(b_i\right)_B &= \frac{\partial l_j^*\left(b_i\right)}{\partial b_i} = \frac{\partial l\left(b_i\right)}{\partial b_i} + \frac{\partial \ln g\left(b_i\right)}{\partial b_i}. \\
S\left(b_i\right)_B &= \sum_{j=1}^{n}\left(y_{ij} - P_{ij}\right) - \frac{b_i - \mu_b}{\sigma_b^2}.
\end{aligned}
$$

- Informação de Fisher

$$
I(b_i) = \sum_{j=1}^{n} P_{ij} Q_{ij} + \frac{1}{\sigma_b^2}
$$

- Erro-padrão a posteriori (assintótico)

$$
\mathcal{EP}(b_i) = \sqrt{I(\widehat{b}_i)^{-1}}
$$

Posteriori

$$p(\boldsymbol{\zeta}, \boldsymbol{\eta} | \mathbf{y}_{..}) \propto \left\{ \prod_{j=1}^{n} \int_{\Re} \prod_{i=1}^{I} P\left(\mathbf{Y}_{ij} | \theta, \boldsymbol{\zeta}_i\right) g\left(\theta, \boldsymbol{\eta}\right) d\theta \right\} \left\{ \prod_{i=1}^{I} p(\boldsymbol{\zeta}_i | \boldsymbol{\tau}_i) \right\}$$

Log-posteriori

$$\ln p(\boldsymbol{\zeta}, \boldsymbol{\eta} | \mathbf{y}_{..}) \propto \left\{ \sum_{j=1}^{n} \ln \int_{\Re} \prod_{i=1}^{I} P\left(\mathbf{Y}_{ij} | \theta, \boldsymbol{\zeta}_i\right) g\left(\theta, \boldsymbol{\eta}\right) d\theta \right\} + \left\{ \sum_{i=1}^{I} \ln p(\boldsymbol{\zeta}_i | \boldsymbol{\tau}_i) \right\}$$

Maximizar a log-posteriori

$$\frac{\partial}{\partial \boldsymbol{\zeta}_i} \ln p(\boldsymbol{\zeta}, \boldsymbol{\eta} | \mathbf{y}_{..}) \propto \left\{ \sum_{j=1}^{n} \frac{\partial}{\partial \boldsymbol{\zeta}_i} \ln \int_{\Re} \prod_{i=1}^{I} P\left(\mathbf{Y}_{ij} | \theta, \boldsymbol{\zeta}_i\right) g\left(\theta, \boldsymbol{\eta}\right) d\theta \right\} + \left\{ \frac{\partial}{\partial \boldsymbol{\zeta}_i} \ln p(\boldsymbol{\zeta}_i | \boldsymbol{\tau}_i) \right\}$$

Expressões

$$
\begin{aligned}
S(b_i)_{BM} &= \sum_{l=1}^{q}\left[\left(\sum_{j=1}^{n} y_{ij} g_j^{*}\left(\overline{\theta}_l\right) - P_{il}\sum_{j=1}^{n} g_j^{*}\left(\overline{\theta}_l\right)\right)\right] - \frac{b_i - \mu_b}{\sigma_b^2} \\
&= \sum_{l=1}^{q}\left[\left(\overline{r}_{il} - \overline{f}_{il} P_{il}\right)\right] - \frac{b_i - \mu_b}{\sigma_b^2}\,, \\
I(b_i)_{BM} &= \sum_{l=1}^{q} \overline{f}_{il} P_{il} Q_{il} + \frac{1}{\sigma_b^2}
\end{aligned}
$$

Algoritmo EM - distribuição a posteriori

- Tomando a esperança da log-posteriori, condicionada a $(\mathbf{Y}'_{..}, \mathbf{b}')'$, para os dados completos, temos que

$$E[\ln p(\mathbf{b}|(\mathbf{Y}'_{..}, \mathbf{b}')')] = \overline{C} + \sum_{i=1}^{I} \sum_{l=1}^{q} \left\{ \overline{r}_{il} \ln P_{il} + (\overline{f}_{il} - \overline{r}_{il}) \ln Q_{il} \right\} + \ln p(\mathbf{b}),$$

em que

$$\overline{r}_{il} = E[r_{il}|\mathbf{Y}_{..}, \mathbf{b}], \quad \overline{f}_{il} = E[f_{il}|\mathbf{Y}_{..}, \mathbf{b}] \quad \text{e} \quad \overline{C} = E[C|\mathbf{Y}_{..}, \mathbf{b}].$$

- Dessa forma, os passos E e M são :
  - **Passo E**
    Usar os pontos de quadratura $\overline{\theta}_l$, os pesos A_l, $l = 1, \cdots, q$ e estimativas iniciais dos parâmetros dos itens, $\widehat{\mathbf{b}}_i$, $i = 1, \cdots, I$, para gerar $g_j^*(\overline{\theta}_l)$ e, posteriormente, $\overline{r}_{il}$ e $\overline{f}_{il}$, $i = 1, \cdots, I$ e $l = 1, \cdots, q$.
  - **Passo M**
    Com $\mathbf{r}$ e $\mathbf{f}$ obtidos no Passo E, resolver as equações de estimação para ζ_i, $i = 1, \cdots, I$, usando o algoritmo de Newton-Raphson ou Escore de Fisher.

Passo M

- Passo M

$$\widehat{b}_i^{(t+1)} = \widehat{b}_i^{(t)} + I(\widehat{b}_i^{(t)})^{-1} S(\widehat{b}_i^{(t)})$$

em que

$$\begin{aligned} S(b_i) &= \sum_{l=1}^{q} \left[\left(\sum_{j=1}^{n} y_{ij} g_j^*\left(\overline{\theta}_l\right) - P_{il} \sum_{j=1}^{n} g_j^*\left(\overline{\theta}_l\right) \right) \right] - \frac{b_i - \mu_b}{\sigma_b^2} \\ &= \sum_{l=1}^{q} \left[\left(\overline{r}_{il} - \overline{f}_{il} P_{il} \right) \right] , \\ I(b_i) &= \sum_{l=1}^{q} \overline{f}_{il} P_{il} Q_{il} + \frac{1}{\sigma_b^2} \end{aligned}$$

Máxima Verossimilhança Marginal - MVM:

- ⊕ Possui propriedades assintóticas: as estimativas dos parâmetros a_i, b_i e c_i são consistentes;
- ⊕ Uma vez estimados os parâmetros dos itens, pode-se estimar as habilidades através de métodos simples;
- ⊕ Permite resolver o problema de indeterminação (métrica) relativo ao modelo,uma vez que se atribui um parâmetro de escala e de posição para a distribuição das habilidades;
- ⊕ Permite caracterizar empiricamente a distribuição das habilidades;
- ⊖ Não está definido para itens com acerto total ou erro total;
- ⊖ É bastante trabalhoso computacionalmente;
- ⊖ Necessidade do estabelecimento de uma distribuição para θ;
- ⊖ Apresenta problemas na estimação do parâmetro c_i em alguns casos; deve ser usado somente com um número suficientemente grande de respondentes.

Moda marginal a Posteriori - MMAP:

- ⊕ Definido para qualquer padrão de resposta;
- ⊕ Uma vez estimados os parâmetros dos itens, pode-se estimar as habilidades através de métodos simples;
- ⊕ Permite resolver o problema de indeterminação (métrica) relativo ao modelo,uma vez que se atribui um parâmetro de escala e de posição para a distribuição das habilidades;
- ⊕ Permite caracterizar empiricamente a distribuição das habilidades;
- ⊖ É mais trabalhoso computacionalmente do que o MVM;
- ⊖ Necessidade de distribuições a priori para os parâmetros dos itens.

Posteriori

$$p(\boldsymbol{\zeta}, \boldsymbol{\eta}|\mathbf{y}_{..}) \quad \propto \quad \left\{ \prod_{j=1}^{n} \int_{\Re} \prod_{i=1}^{I} P\left(\mathbf{Y}_{ij}|\theta, \boldsymbol{\zeta}_i\right) g\left(\theta, \boldsymbol{\eta}\right) d\theta \right\} \left\{ \prod_{i=1}^{I} p(\boldsymbol{\zeta}_i|\boldsymbol{\tau}_i) \right\} \left\{ \prod_{i=1}^{I} p(\boldsymbol{\tau}_i) \right\}$$

Log-posteriori

$$\begin{aligned} \ln p(\boldsymbol{\zeta}, \boldsymbol{\eta}|\mathbf{y}_{..}) \quad \propto \quad & \left\{ \sum_{j=1}^{n} \ln \int_{\Re} \prod_{i=1}^{I} P\left(\mathbf{Y}_{ij}|\theta, \boldsymbol{\zeta}_i\right) g\left(\theta, \boldsymbol{\eta}\right) d\theta \right\} + \left\{ \sum_{i=1}^{I} \ln p(\boldsymbol{\zeta}_i|\boldsymbol{\tau}_i) \right\} \\ & + \left\{ \sum_{i=1}^{I} \ln p(\boldsymbol{\tau}_i) \right\} \end{aligned}$$

Maximizar a log-posteriori

$$\frac{\partial}{\partial \boldsymbol{\zeta}_i} \ln p(\boldsymbol{\zeta}, \boldsymbol{\eta}|\mathbf{y}_{..}) \quad \propto \quad \left\{ \sum_{j=1}^{n} \frac{\partial}{\partial \boldsymbol{\zeta}_i} \ln \int_{\Re} \prod_{i=1}^{I} P\left(\mathbf{Y}_{ij}|\theta, \boldsymbol{\zeta}_i\right) g\left(\theta, \boldsymbol{\eta}\right) d\theta \right\} + \left\{ \frac{\partial}{\partial \boldsymbol{\zeta}_i} \ln p(\boldsymbol{\zeta}_i|\boldsymbol{\tau}_i) \right\}$$

Parâmetro a

$$p(a_i|\mu_{a_i}, \sigma^2_{a_i}) = \frac{1}{\sqrt{2\pi} a_i \sigma_{a_i}} \exp\left[-\frac{1}{2\sigma^2_{a_i}}(\ln a_i - \mu_{a_i})^2\right].$$

Parâmetro b

$$p(b_i|\mu_{b_i}, \sigma^2_{b_i}) = \frac{1}{\sqrt{2\pi}\sigma_{b_i}} \exp\left\{\frac{-(b_i - \mu_{b_i})^2}{2\sigma^2_{b_i}}\right\}$$

Parâmetro c

$$p(c_i|\alpha_i, \beta_i) = \frac{\Gamma(\alpha_i + \beta_i - 2)}{\Gamma(\alpha_i - 1)\Gamma(\beta_i - 1)} c_i^{\alpha_i - 2}(1 - c_i)^{\beta_i - 2}.$$

$$\mathbf{S}(\zeta_i)_B \quad = \quad \sum_{l=1}^{q} (\bar{r}_{il} - f_{il} P_{il})\, \mathbf{h}_{il} + \boldsymbol{\lambda}_i\,,$$

com

$$\mathbf{h}_{ij} = \left(P_{ij}^{*} Q_{ij}^{*}\right)^{-1} \left(\frac{\partial P_{ij}}{\partial \zeta_i}\right) = \begin{pmatrix} D\,(1 - c_i)\,(\theta_j - b_i) \\ -Da_i\,(1 - c_i) \\ \dfrac{1}{P_{ij}^{*}} \end{pmatrix}$$

$$\boldsymbol{\lambda}_i \quad = \quad \left[\frac{1}{a_i}\left[1 + \frac{\ln a_i - \mu_{a_i}}{\sigma_{a_i}^2}\right]\, ; \, -\frac{(b_i - \mu_{b_i})}{\sigma_{b_i}^2}\, ; \, \frac{\alpha_i - 2}{c_i} - \frac{\beta_i - 2}{1 - c_i}\right]' .$$

$$\frac{\partial^2}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \ln p(\boldsymbol{\zeta}, \boldsymbol{\eta} | \mathbf{y}_{..}) \propto \left\{ \sum_{j=1}^{n} \frac{\partial^2}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \ln \int_{\Re} \prod_{i=1}^{I} P\left(\mathbf{Y}_{ij} | \theta, \boldsymbol{\zeta}_i\right) g\left(\theta, \boldsymbol{\eta}\right) d\theta \right\} \left\{ \frac{\partial^2}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \ln p(\boldsymbol{\zeta}_i | \boldsymbol{\tau}_i) \right\}$$

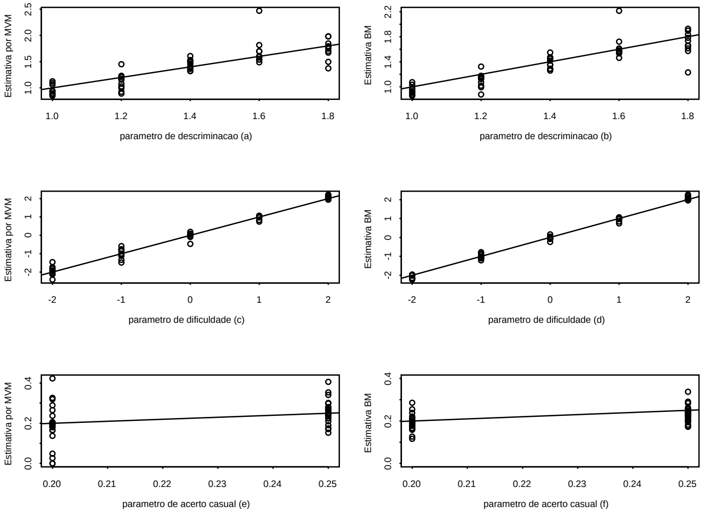
$$\mathbf{I}(\boldsymbol{\zeta}_i)_{BM} = \sum_{l=1}^{q} \overline{f}_{il} P_{il}^* Q_{il}^* \mathbf{h}_{il} \mathbf{h}_{il}^{'} - \boldsymbol{\Lambda}_i \, .$$

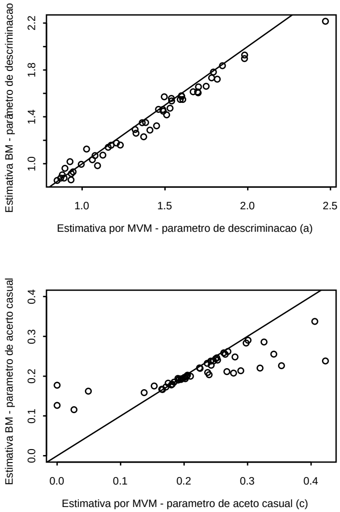
$$\boldsymbol{\Lambda}_i = \begin{bmatrix} \dfrac{\left[\sigma_{a_i}^2 + \ln a_i - \mu_{a_i} - 1\right]}{\mu_{a_i}^2 \sigma_{a_i}^2} & . & . \\ 0 & -\dfrac{1}{\sigma_{b_i}^2} & . \\ 0 & 0 & -\left[\dfrac{\alpha_i - 2}{c_i^2}\right] + \dfrac{\beta_i - 2}{(1 - c_i)^2} \end{bmatrix} .$$

- Foram gerados n = 5000 valores independentes de uma distribuição N(0,1) a fim de servirem como as habilidades dos indivíduos.
- I = 50 itens de maneira a cobrir os valores apropriados para os parâmetros: a variando de 1,0 a 1,8, b variando de -2,0 a 2,0 e c = 0,20 e 0,25.
- Para a geração das respostas foi construído um programa em linguagem Ox.
- Para a obtenção das estimativas dos parâmetros (itens e habilidades) foi usado o programa Bilog

Tabela: Estatísticas relativas as estimativas dos parâmetros dos itens

Estatística	**Parâmetro**					
	discriminação(a)		**dificuldade(b)**		**acerto casual (c)**	
	MVM	**BM**	**MVM**	**BM**	**MVM**	**BM**
SQR	1,74	1,30	1,81	0,65	0,31	0,07
SVar	1,27	0,92	3,10	0,89	0,75	0,14
SEQM	3,01	2,22	4,91	1,54	1,06	0,21


Estimativa por MVM
parametro de descriminacao (a)
Estimativa BM
parametro de descriminacao (b)
Estimativa por MVM
parametro de dificuldade (c)
Estimativa BM
parametro de dificuldade (d)
Estimativa por MVM
parametro de acerto casual (e)
Estimativa BM
parametro de acerto casual (f)


Estimativa BM - parâmetro de descriminacao
2.2
1.8
1.4
1.0
1.0
1.5
2.0
2.5
Estimativa por MVM - parametro de descriminacao (a)
Estimativa BM - parâmetro de dificuldade
2
1
0
-1
-2
-2
-1
0
1
2
Estimativa por MVM - parametro de dificuldade (b)
Estimativa BM - parametro de acerto casual
0.4
0.3
0.2
0.1
0.0
0.0
0.1
0.2
0.3
0.4
Estimativa por MVM - parametro de aceto casual (c)

Log-posteriori

$$\ln g_j^*(\theta_j) = Const + \ln P(\mathbf{Y}_{.j}|\theta_j, \boldsymbol{\zeta}) + \ln g(\theta_j|\boldsymbol{\eta}). \tag{1}$$

Equação de estimação bayesiana

$$\frac{\partial \ln g_j^*(\theta_j)}{\partial \theta_j} = \frac{\partial \ln P(\mathbf{Y}_{.j}|\theta_j, \boldsymbol{\zeta})}{\partial \theta_j} + \frac{\partial \ln g(\theta_j|\boldsymbol{\eta})}{\partial \theta_j} = 0.$$

Informação de Fisher

$$I(\theta_j) = \sum_{i=1}^{I} P_{ij}^* Q_{ij}^* W_{ij} h_{ij}^2 - \frac{1}{\sigma^2}. \tag{2}$$

Scoring de Fisher

$$\widehat{\theta}_j^{(t+1)} = \widehat{\theta}_j^{(t)} + I(\theta_j^{(t)})^{-1} S(\theta_j^{(t)})$$

$$g(\theta|\mathbf{y}_{.j},\boldsymbol{\zeta},\boldsymbol{\eta}) = \frac{P(\mathbf{y}_{.j}|\theta,\boldsymbol{\zeta})g(\theta|\boldsymbol{\eta})}{P(\mathbf{y}_{.j}|\boldsymbol{\zeta},\boldsymbol{\eta})}. \tag{3}$$

Segue que a esperança da posteriori é

$$\widehat{\theta}_j \equiv E[\theta|\mathbf{y}_{.j},\boldsymbol{\zeta},\boldsymbol{\eta}] = \frac{\int_{\Re} \theta g(\theta|\boldsymbol{\eta})P(\mathbf{y}_{.j}|\theta,\boldsymbol{\zeta})d\theta}{\int_{\Re} g(\theta|\boldsymbol{\eta})P(\mathbf{y}_{.j}|\theta,\boldsymbol{\zeta})d\theta}.$$

$$\begin{aligned} \boldsymbol{E}\left[\theta_j|\mathbf{y}_{.j},\boldsymbol{\zeta},\boldsymbol{\eta}\right] \quad &\approx \quad \boldsymbol{E}\left[\overline{\theta}_j|\mathbf{y}_{.j},\boldsymbol{\zeta},\boldsymbol{\eta}\right] = \frac{\sum_{l=1}^{q}\overline{\theta}_l P(\mathbf{Y}_{.j}|\overline{\theta}_l,\boldsymbol{\zeta})g(\overline{\theta}_l|\boldsymbol{\eta})}{\sum_{l=1}^{q} P(\mathbf{Y}_{.j}|\overline{\theta}_l,\boldsymbol{\zeta})g(\overline{\theta}_l|\boldsymbol{\eta})} \\ &= \quad \frac{\sum_{l=1}^{q}\overline{\theta}_l P(\mathbf{Y}_{.j}|\overline{\theta}_l,\boldsymbol{\zeta})A_l}{\sum_{l=1}^{q} P(\mathbf{Y}_{.j}|\overline{\theta}_l,\boldsymbol{\zeta})A_l}. \end{aligned}$$

E a variância a posteriori

$$Var\left(\theta_j\right) \equiv Var[\theta|\mathbf{y}_{.j},\boldsymbol{\zeta},\boldsymbol{\eta}] = \frac{\int_{\Re}\left(\theta - E\left(\theta\right)\right)^2 g(\theta|\boldsymbol{\eta})P(\mathbf{y}_{.j}|\theta,\boldsymbol{\zeta})d\theta}{\int_{\Re} g(\theta|\boldsymbol{\eta})P(\mathbf{y}_{.j}|\theta,\boldsymbol{\zeta})d\theta}.$$

$$Var\left[\overline{\theta}_j|\mathbf{y}_{.j},\boldsymbol{\zeta},\boldsymbol{\eta}\right] = \frac{\sum_{l=1}^{q}\left\{\overline{\theta}_l - \boldsymbol{E}\left[\overline{\theta}_j|\mathbf{y}_{.j},\boldsymbol{\zeta},\boldsymbol{\eta}\right]\right\}^2 P(\mathbf{Y}_{.j}|\overline{\theta}_l,\boldsymbol{\zeta})g(\overline{\theta}_l|\boldsymbol{\eta})}{\sum_{l=1}^{q} P(\mathbf{Y}_{.j}|\overline{\theta}_l,\boldsymbol{\zeta})g(\overline{\theta}_l|\boldsymbol{\eta})}$$

Esperança a posteriori

- MV com MVM : estimativa de máxima verossimilhança das habilidades usando a estimativa de máxima verossimilhança marginal dos parâmetros dos itens.
- MV com MMAP : estimativa de máxima verossimilhança das habilidades usando a estimativa bayesiana marginal dos parâmetros dos itens.
- EAP com MVM : estimativa da esperança a posteriori das habilidades usando a estimativa de máxima verossimilhança margianl dos parâmetros dos itens.
- EAP com MMAP : estimativa da esperança a posteriori das habilidades usando a estimativa de bayesiana marginal dos parâmetros dos itens.
- MAP com MVM : estimativa da moda a posteriori das habilidades usando a estimativa de máxima verossimilhança marginal dos parâmetros dos itens.
- MAP com MMAP : estimativa da moda a posteriori das habilidades usando a estimativa bayesiana marginal dos parâmetros dos itens.

Tabela: Estatísticas relativas as estimativas das habilidades

Métodos de Estimação	Estatísticas		
	SQR	**SQVar**	**SQM**
MV com MVM	831,08	30938766,88	30939597,96
MV com MMAP	807,59	22954776,39	22955583,98
EAP com MVM	626,85	632,70	1259,55
EAP com MMAP	624,10	638,09	1262,20
MAP com MVM	628,80	617,02	1245,82
MAP com MMAP	625,44	623,37	1248,81

Esperança a posteriori