

M4: Enfoque Bayesiano nos modelos de resposta ao item



XXI SINAPE NATAL 2014

Simpósio Nacional de Probabilidade e Estatística
20 a 25 de julho de 2014

Praiamar Hotel - Natal-RN

PROMOÇÃO E REALIZAÇÃO



APOIO



PATROCÍNIO



ORGANIZAÇÃO



Embratur:
20.008452.89.0001-6

Associação Brasileira de Estatística
21° SINAPE - Simpósio Nacional de Probabilidade e Estatística - 2014

**ENFOQUE BAYESIANO NOS MODELOS DE
TEORIA DA RESPOSTA AO ITEM**

Dr. Jorge Luis Bazán

jlbazan@icmc.usp.br

<http://www.icmc.usp.br/pessoas/jlbazan/>

De 20 a 25 de julho de 2014 Hotel Praiamar, Ponta Negra, Natal-RN.

1. Introdução: Regressão binária	pág 3
2. Modelos de resposta dicotômica da TRI	pág 11
3. Novas ligações nos modelos de Resposta dicotômica	pág 32
4. Inferência Bayesiana na TRI	pág 64
5. Extensões dos Modelos TRI	pág 85
6. Comentários finais	pág 130
7. Outras Referências	pág 133

1. INTRODUÇÃO: REGRESSÃO BINÁRIA

- A regressão Logística tem sido o modelo mais usado na regressão binária em diferentes áreas de interesse.
- Todo tipo de respostas (categóricas, discretas e contínuas) podem ser transformadas em respostas dicotômicas.
- É uma ferramenta poderosa, bastante desenvolvida da perspectiva estatística e com muitas contribuições acadêmicas e inúmeras aplicações.
- No entanto, existem evidências de situações em que ela pode estar mal especificada e os dados podem estar mal ajustados.
- Assim, muitas propostas de modelos alternativos de regressão binária foram desenvolvidas, em especial considerando ligações assimétricas.

Considere: $y = (y_1, y_2, \dots, y_n)'$ v.a. dependente com n observações, $x_i = (x_{i1}, \dots, x_{in})'$ covariáveis, e $\beta = (\beta_1, \dots, \beta_k)'$ coeficientes de regressão.

Também $y_i = 1$ com probabilidade p_i e $y_i = 0$ com probabilidade $1 - p_i$.

Na modelagem de dados binários assume-se que:

$$p_i = F(\eta_i) = F(x_i'\beta), \quad i = 1, \dots, n,$$

em que $F(\cdot)$ denota a função de distribuição acumulada (f.d.a.) e F^{-1} é chamada a função de ligação em que $\eta_i = x_i'\beta$ é o i -ésimo preditor linear.

Na regressão Logística temos que $F(\cdot) = \frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)}$ em que $F(\cdot)$ corresponde a f.d.a. da distribuição Logística padrão. Neste caso $F^{-1}(\cdot)$ é chamada função de ligação LOGIT.

Quando $F(\cdot)$ corresponde a f.d.a. da distribuição Normal padrão, isto é $F(\cdot) = \Phi(x'_i\beta)$, tem-se a regressão probito e neste caso $F^{-1}(\cdot)$ é chamada de função de ligação PROBIT.

Quando $F(\cdot) = 1 - \exp(-\exp(x))$ corresponde á f.d.a. da Gumbel, tem-se a regressão de valor extremo. Neste caso $F^{-1}(\cdot)$ é chamada de função de ligação LOG LOG COMPLEMENTAR.

Quando $F(\cdot) = \exp(-\exp(-x))$ corresponde a f.d.a. da Gumbel reversa, tem-se a regressão de valor extremo reversa e neste caso $F^{-1}(\cdot)$ é chamada de função de ligação LOGLOG.

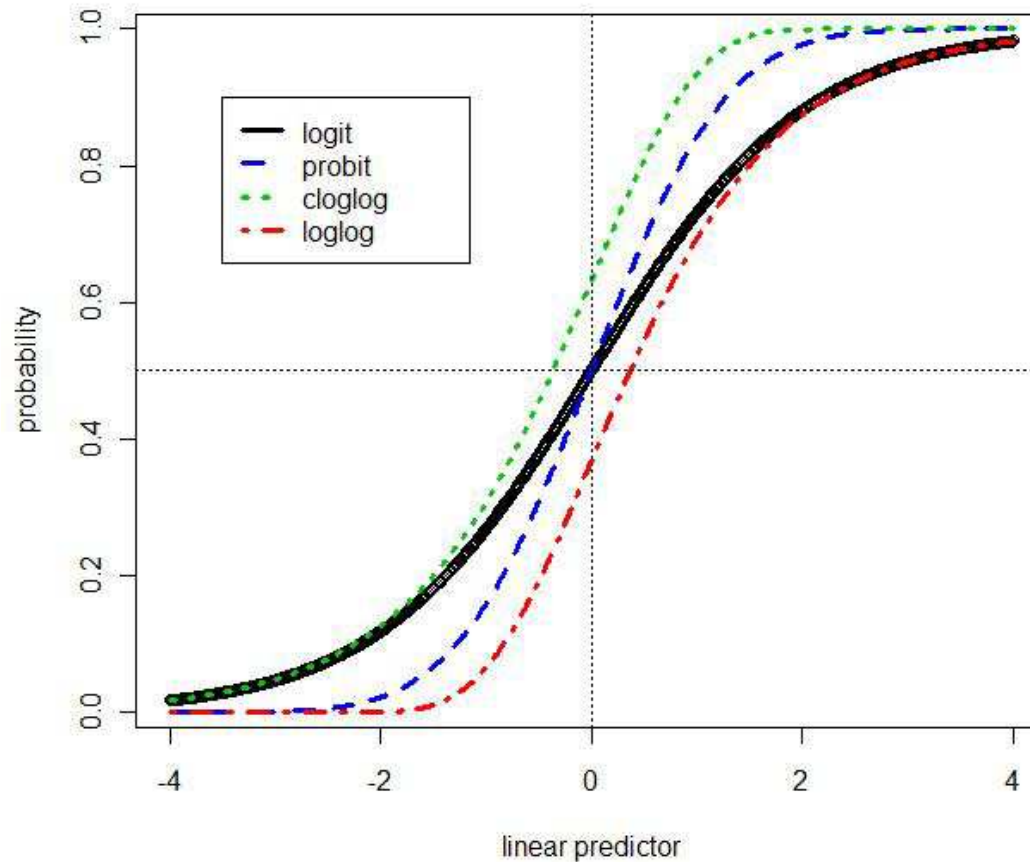


Figura 1. Funções de ligação em Regressão Binária. Logito e Probit são simétricas em torno de $p=0.5$ e o preditor linear = 0, o que é diferente para Cloglog e Loglog.

Em dados binários:

- Ligações simétricas (em torno de $p_i = 0.5$): Probit e Logito.
- Ligações simétricas podem ser inadequadas e mal especificadas (Figura 2 e 3) (Collet, 2003).
- Ligações assimétricas: Prentice (1976), Aranda-Ordaz (1981), Guerrero e Johnson (1982), Stukel (1988), Czado e Santner (1992a,b), Nagler (1994), Chen et al., (1999), Basu e Mukhopadhyay (2000), Haro-López, et al. (2000).
- Ligações Probit assimétricas: Czado (1994), Chen et al. (1999) e Bazán et al. (2005).

Maiores detalhes em:

Bazán, J. L., Bolfarine, H. e Branco, D. M. (2010) A framework for skew-probit links in Binary regression. *Communications in Statistics - Theory and Methods*, **39**, 678-697.

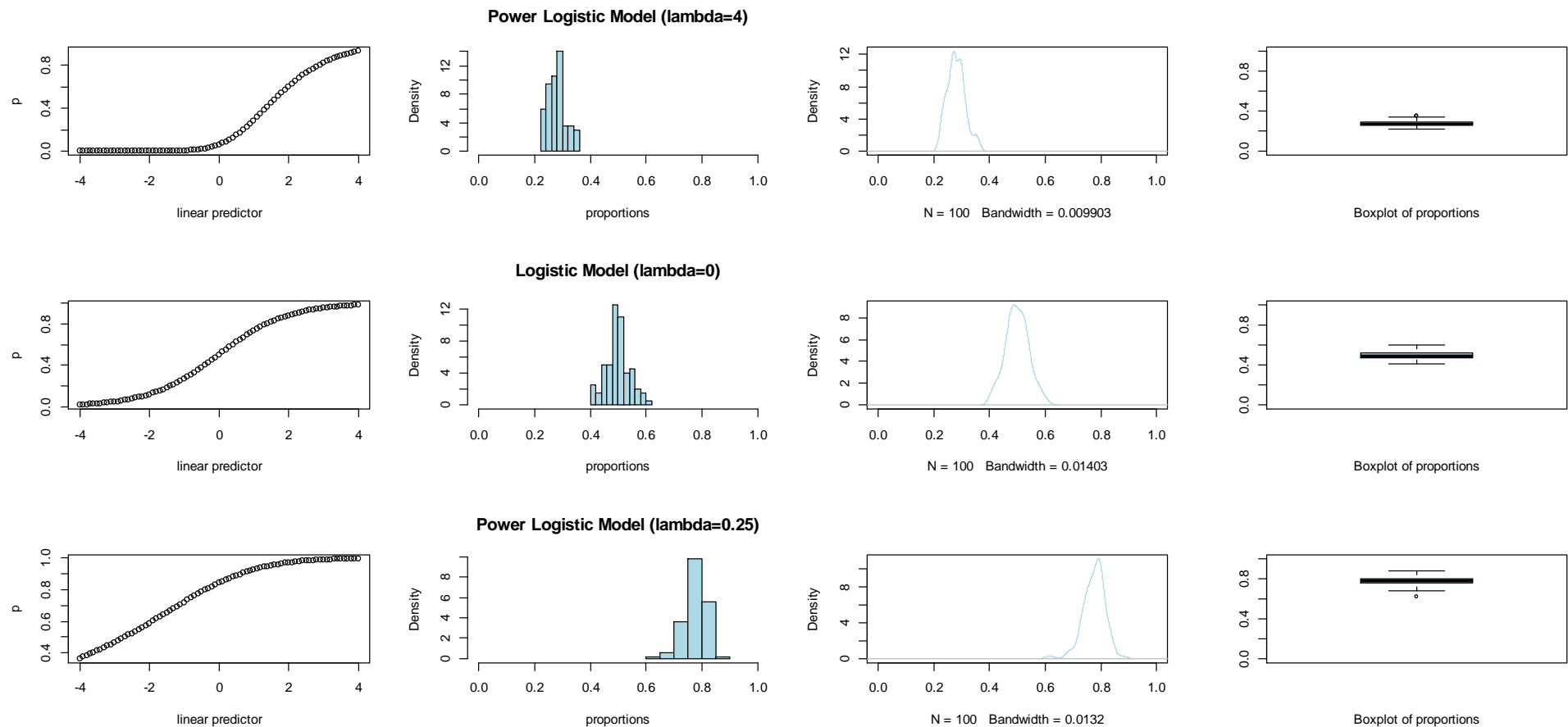


Figura 2. Distribuição de 1's e 0's ($n=81$) em 100 amostras simuladas considerando a ligação Power Logística (PL) com o preditor linear fixo. Note que PL é mais adequado para altos e baixos valores de proporção observada.

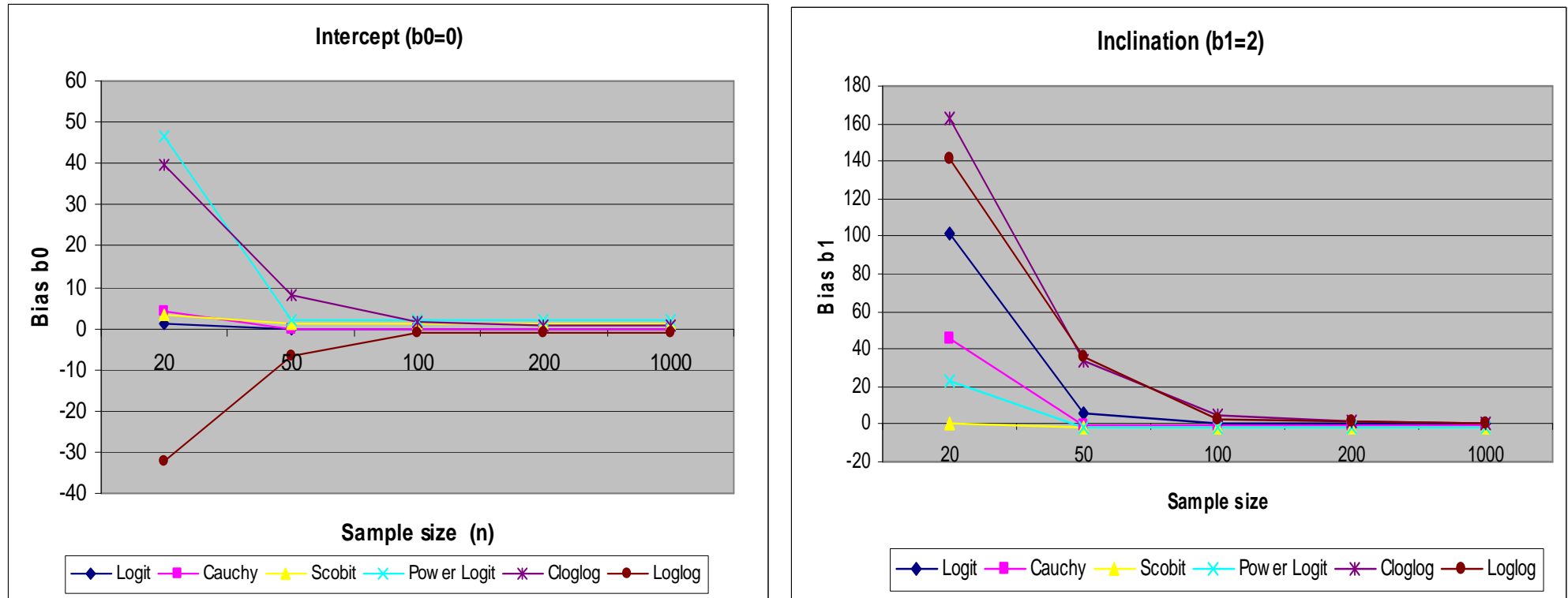
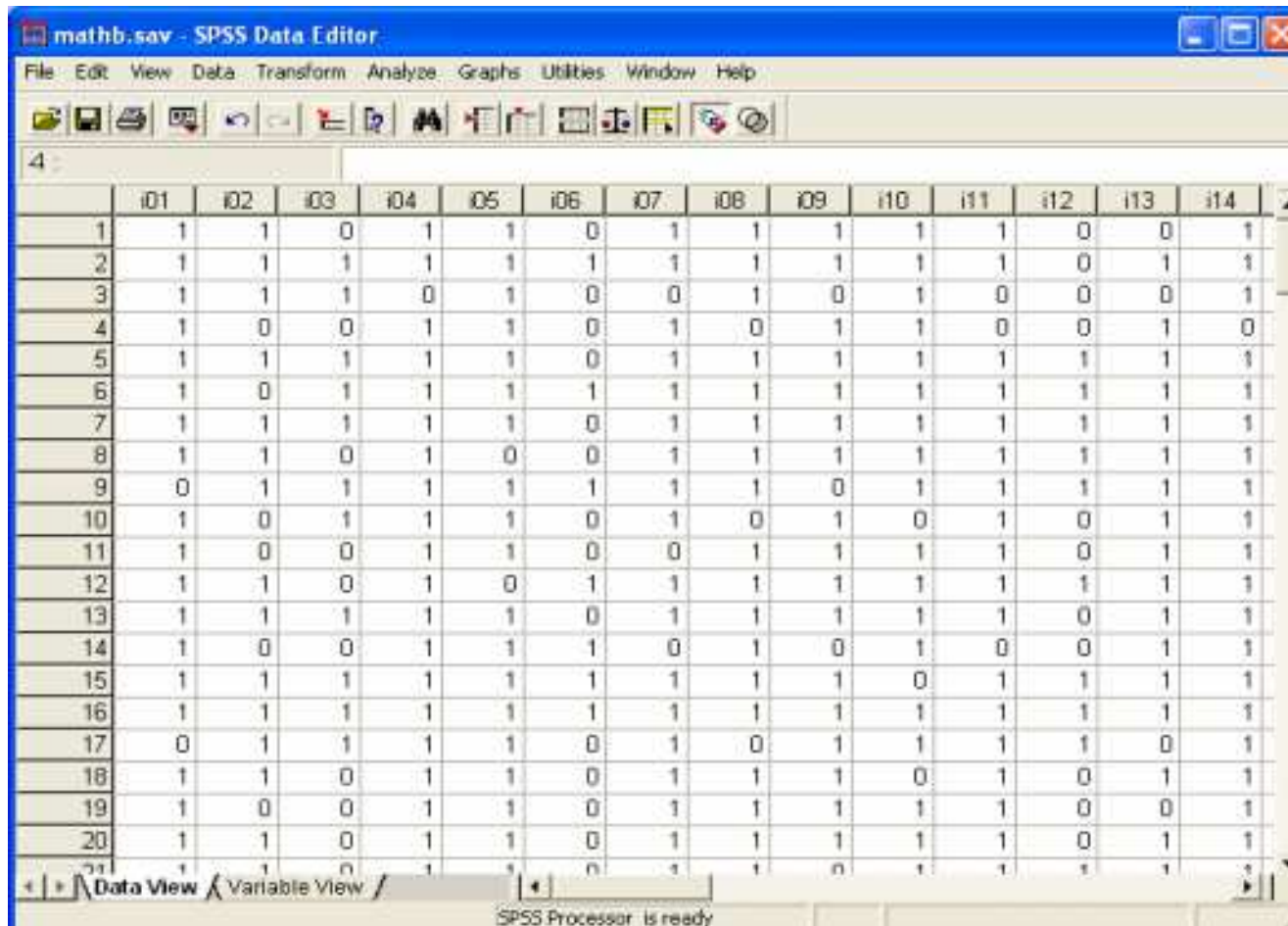


Figura 3. Viés na estimação de coeficientes de uma regressão binária simples com diferentes ligações quando a ligação Logito é usada como padrão (modelo mal especificado). Note que quando o tamanho da amostra se incrementa o viés decresce, no entanto ainda permanece importante.

2. MODELOS DE RESPOSTA DICOTÔMICA NA TRI

2.1	Definição e Observações	pág 16
2.2	Casos particulares: O modelo TRI com ligação Logito	pág 21
2.3	Verossimilhança na TRI	pág 25
2.4	Formulações como modelo de regressão latente	pág 26
2.5	Verossimilhança aumentada do modelo TRI	pág 29



4

	i01	i02	i03	i04	i05	i06	i07	i08	i09	i10	i11	i12	i13	i14
1	1	1	0	1	1	0	1	1	1	1	1	0	0	1
2	1	1	1	1	1	1	1	1	1	1	1	0	1	1
3	1	1	1	0	1	0	0	1	0	1	0	0	0	1
4	1	0	0	1	1	0	1	0	1	1	0	0	1	0
5	1	1	1	1	1	0	1	1	1	1	1	1	1	1
6	1	0	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	0	1	1	1	1	1	1	1	1
8	1	1	0	1	0	0	1	1	1	1	1	1	1	1
9	0	1	1	1	1	1	1	1	0	1	1	1	1	1
10	1	0	1	1	1	0	1	0	1	0	1	0	1	1
11	1	0	0	1	1	0	0	1	1	1	1	0	1	1
12	1	1	0	1	0	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	0	1	1	1	1	1	0	1	1
14	1	0	0	1	1	1	0	1	0	1	0	0	1	1
15	1	1	1	1	1	1	1	1	1	0	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	0	1	1	1	1	0	1	0	1	1	1	1	0	1
18	1	1	0	1	1	0	1	1	1	0	1	0	1	1
19	1	0	0	1	1	0	1	1	1	1	1	0	0	1
20	1	1	0	1	1	0	1	1	1	1	1	0	1	1

SPSS Processor is ready

Figure 4. Dados para a Teoria da Resposta ao item (itens: I01 até I18, indivíduos: 1 até n=978).

Os primeiros modelos da TRI se desenvolveram nas décadas dos anos 40 e 50 do século anterior e tem sido popularizados ao serem usados nos sistemas de avaliação educacional internacionais e nacionais.

No Brasil:

- Sistema Nacional de Avaliação do Ensino Básico – SAEB.
- Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo – SARESP.
- ENEM.

No mundo:

LLECE, TIMS e PISA.

- Andrade, D. F., Tavares, H. R. e Valle, R. C. (2000). *Teoria da Resposta ao Item. Conceitos e Aplicações*. Associação Brasileira de Estatística: São Paulo.
- Klein, R. (2003). *Utilização da Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB)*. Ensaio: Avaliação e Políticas Públicas em Educação. Rio de Janeiro, **11** (40), 283-296.
- Pasquali, L. (2007). *Teoria de resposta ao item: teoria, procedimentos e aplicações*. Brasília: LabPAM/UnB.
- Moreira, J. F. (2010). Aplicações da Teoria da Resposta ao item (TRI) no Brasil. *Rev. Bras. Biom.*, São Paulo, **28** (4) , 137-170. Disponível http://jaguar.fcav.unesp.br/RME/fasciculos/v28/v28_n4/A9_Fernando.pdf
- Azevedo, C. e Gamerman, D. (2009). *Introdução à Teoria de Resposta ao Item*. Minicurso. I Conbratri.

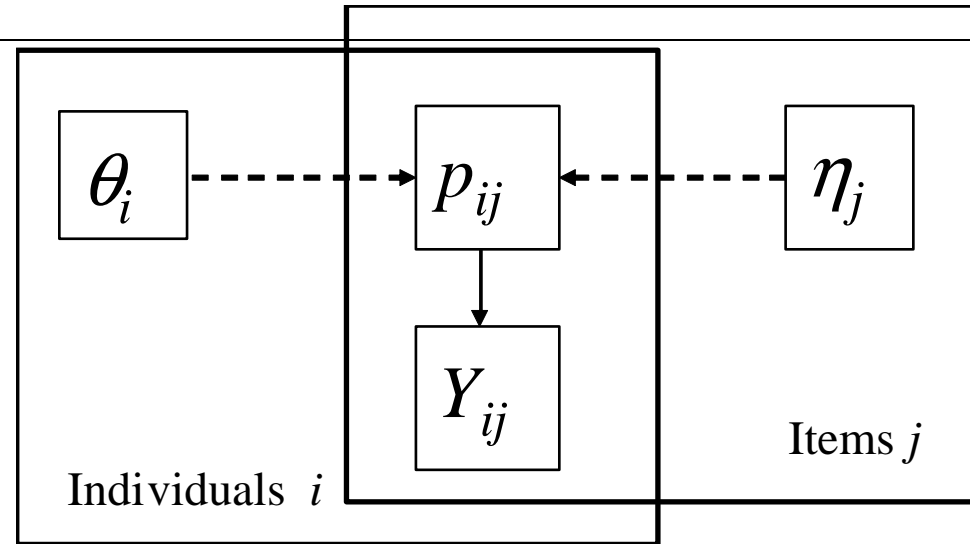


Figura 5. Diagrama da Teoria da Resposta ao Item

- \longrightarrow Variáveis manifestas, $- - \rightarrow$ Variáveis latentes.
- “Caixas” para indivíduos e itens (variável latente e parâmetros de item).
- Sob perspectiva bayesiana é possível adicionar estruturas distribucionais explícitas para os parâmetros de item η_j e para a variável latente θ_i .

2.1 Definição e Observações

Considere n indivíduos avaliados em um teste de k itens. Um modelo TRI de dois parâmetros binário e unidimensional é um sistema no qual para cada indivíduo i existe um modelo de variável latente monótono unidimensional (Y, Θ_i) , definido pelas seguintes expressões:

$$\begin{aligned} Y_{ij} | u_i, \eta_j &\sim \text{Bernoulli}(p_{ij}) \\ p_{ij} &= P(Y_{ij} = 1 | \theta_i, \eta_j) = F(m_{ij}) \\ m_{ij} &= a_j(\theta_i - b_j), \\ i &= 1 \dots, n, \quad j = 1, \dots, I \end{aligned}$$

em que:

- Y_{ij} é a variável manifesta a qual modela a resposta binária de cada indivíduo responder corretamente ou não ao item j .

- $\eta_j = (a_j, b_j)$ são dois parâmetros que representam a discriminação e a dificuldade do item j , respectivamente.
- θ_i é o valor da variável ou traço latente Θ_i do indivíduo i , em algumas situações a habilidade do indivíduo i .
- p_{ij} é a probabilidade condicional dado $\Theta_i = \theta_i$.
- F é chamada a curva característica do item (CCI) e
- m_{ij} é o preditor linear latente relacionada com o traço latente do indivíduo i que responde ao item j .

Observações

O modelo TRI de 2 parâmetros:

- Satisfaz a propriedade de independência condicional latente, isto é, para cada indivíduo i as respostas Y_{ij} a diferentes itens é condicionalmente independente dada a variável latente θ_i .
- Satisfaz a propriedade de monotonicidade latente, porque é uma função estritamente não decrescente de θ_i .
- É latente unidimensional.
- $F(m_{ij})$ é a mesma para cada caso e $F^{-1}(\cdot)$ é chamada a função de ligação.
- Também, assume-se independência entre as respostas de diferentes indivíduos.

- Os parâmetros de dificuldade b_j e de discriminação a_j representam a posição e inclinação do item, respectivamente, a_j é o valor proporcional à inclinação da CCI no ponto b_j .

- Valores de $a_j < 0$ não são esperados. O espaço paramétrico para o parâmetro b_j é arbitrário e corresponde ao mesmo que o traço latente que esta geralmente na reta real.
- Outra parametrização para preditor latente é $m_{ij} = a_j\theta_i - b_j$ a qual é preferida na inferência bayesiana, pois facilita o tempo de convergência computacional. Quando é usada esta parametrização, o parâmetro prévio de dificuldade pode ser obtido fazendo $\frac{b_j}{a_j}$.

O primeiro modelo TRI binário foi introduzido pelo Lord (1952) com uma CCI dado por $F(x) = \Phi(x)$, em que Φ corresponde a f.d.a. da Normal padrão. Este modelo é conhecido na literatura psicométrica como o modelo de Ogiva Normal o qual corresponde no contexto dos modelos lineares generalizados à ligação probito e pode ser denotada como modelo $2P$.

Também, Birbaum (1968) propôs uma CCI dada por $F(x) = L(x)$, em que $L(x) = \frac{e^x}{1 + e^x}$ corresponde a f.d.a. da distribuição Logística padrão. Na linguagem dos modelos lineares generalizados corresponde à ligação Logito. Este modelo é conhecido como modelo *Logístico* e pode ser chamado modelo $2L$.

2.2 Casos particulares: Os modelos TRI com ligação Logito

Os modelos TRI admitem diversas formulações, as quais dependem basicamente de como a CCI é considerada. Uma versão simples acontece quando $a_j = 1$:

$$P(Y_{ij} = 1 \mid \theta_i, b_j) = F(\theta_i - b_j).$$

Este é chamado de modelo TRI de um parâmetro e quando as ligações Probit ou Logito são consideradas, denotamos como os modelos TRI 1P ou TRI 1L, respectivamente.

Também pode ser considerada a CCI dada por:

$$P(Y_{ij} = 1 \mid u_i, a_j, b_j, c_j) = c_j + (1 - c_j)F(a_j(\theta_i - b_j)),$$

em que o parâmetro $c_j \in [0, 1]$ representa a probabilidade de que um indivíduo com mínimo traço latente responda apropriadamente ao item j e F é uma f.d.a.

Este é conhecido como modelo TRI de 3 parâmetros. Se $c_j = 0$, o modelo se reduz ao modelo de 2 parâmetros.

Quando as ligações Probit e Logito são consideradas temos os modelos TRI 3P e TRI 3L, respectivamente.

Os modelos TRI com ligação Logito

Este modelo, chamado também modelo Logístico é o modelo mais usado na TRI.

A versão do modelo 3L estabelece que a probabilidade do que o indivíduo i responda corretamente o item j é dada por:

$$P(Y_{ij} = 1|\theta_i, a_j, b_j, c_j) = c_j + (1 - c_j) \frac{1}{1 + e^{-Da_j(\theta_i - b_j)}},$$

em que usualmente é assumido $D = 1$ e outros autores consideram $D = 1.7$ para aproximar este modelo para o modelo de Ogiva Normal. Os casos particulares são:

- $P(Y_{ij} = 1|\theta_i, a_j, b_j) = \frac{1}{1 + e^{Da_j(\theta_i - b_j)}}$ (com $c_j = 0$) e
- $P(Y_{ij} = 1|\theta_i, b_j) = \frac{1}{1 + e^{D(\theta_i - b_j)}}$ (com $c_j = 0$ y $a_j = 1$).

O modelo de um parâmetro é conhecido também como o modelo de Rasch. No entanto, esse modelo tem uma outra justificativa (veja, por exemplo, Fischer e Molenaar, 1995).

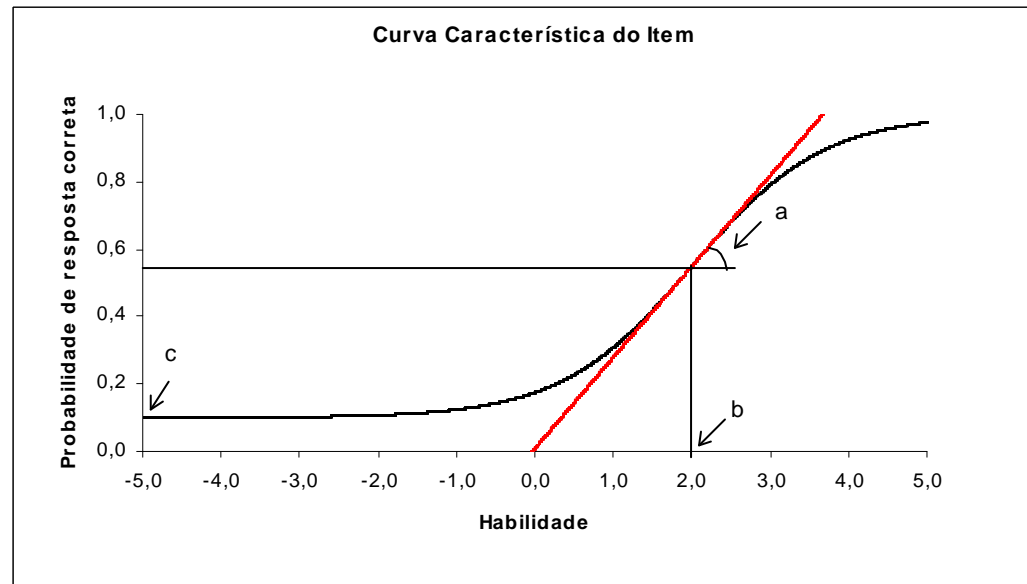


Figura 5. CCI do modelo TRI de 3 parâmetros

- A CCI fornece a probabilidade de indivíduos de vários níveis de traço latente responderem corretamente ao item.
- Condicionado ao traço latente, a probabilidade de um indivíduo responder corretamente ao item só depende do modelo adotado e dos parâmetros do item.
- Resume-se toda a informação com relação a desempenho do indivíduo no teste em θ .

2.3 Verossimilhança na TRI

No modelo TRI de 2 parâmetros, a densidade conjunta das respostas multivariadas $\mathbf{Y} = (Y_1, \dots, Y_n)'$, com $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{ik})$ dado o vetor de variáveis latentes $\Theta = (\Theta_1, \dots, \Theta_n)'$ e o vetor de parâmetros de item $\eta = (\eta_1, \dots, \eta_k)'$ considerando a independência condicional, pode ser escrita como:

$$f(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\eta}) = \prod_{i=1}^n \prod_{j=1}^k F(m_{ij})^{y_{ij}} (1 - F(m_{ij}))^{1-y_{ij}}$$

Porém, seja $D_{obs} = \mathbf{y}$ os dados observados. A função de verossimilhança para o modelo TRI de 2 parâmetros é dada por:

$$L(\boldsymbol{\theta}, \boldsymbol{\eta}|D_{obs}) = \prod_{i=1}^n \prod_{j=1}^I F(m_{ij})^{y_{ij}} (1 - F(m_{ij}))^{1-y_{ij}},$$

em que $F(\cdot)$ é a f.d.a. da Normal padrão ou Logística padrão.

2.4 Formulações como modelo de regressão latente

Uma formulação alternativa do modelo TRI de 2 parâmetros, inicialmente devido a Albert (1992), é obtida assumindo:

$$y_{ij} = \begin{cases} 1, & Z_{ij} > 0; \\ 0, & Z_{ij} \leq 0. \end{cases},$$

$$Z_{ij} = m_{ij} + e_{ij}, e_{ij} \sim F(\mu = 0, \sigma = 1),$$

em que:

- $F(\cdot)$ pode ser a f.d.a. de uma distribuição simétrica como a Normal ou Logística padrão em que $F(s) = 1 - F(-s)$, $\forall s$,
- Z_{ij} é uma variável latente auxiliar associada com a resposta binária y_{ij} a qual pode ser vista como uma variável manifesta indicadora de sim a variável latente é positiva ou não.

$$\left[\begin{aligned} P(y_{ij} = 1 \mid \theta_i) &= P(Z_{ij} > 0 \mid \theta_i) = P(e_{ij} > -m_{ij} \mid \theta_i) = 1 - P(e_{ij} < -m_{ij} \mid \theta_i) = 1 - F(-m_{ij}) = F(m_{ij}), \\ P(y_{ij} = 0 \mid \theta_i) &= P(Z_{ij} \leq 0 \mid \theta_i) = P(e_{ij} \leq -m_{ij} \mid \theta_i) = F(-m_{ij}) = 1 - F(m_{ij}). \end{aligned} \right]$$

Reescrevendo a formulação de a variável latente auxiliar para o item j temos:

$$Z_{ij} = \beta_{0j} + \beta_{1j}\theta_i + e_{ij}, \quad e_{ij} \sim F(\mu = 0, \sigma = 1), \quad i = 1, \dots, n.$$

O qual corresponde à regressão da variável latente associada Z_{ij} em relação à variável latente θ_i no item j , em que:

- $\beta_{0j} = -a_j \times b_j$ é o intercepto que expressa a facilidade do item pesado por sua discriminação, assim quando $\beta_{0j} > 0$ temos um item fácil e no caso contrário um item difícil.
- $\beta_{1j} = a_j$ a discriminação, é vista agora como a inclinação da reta regressora e é dito coeficiente de regressão. Quando $0 < a_j \leq 1$ a inclinação corresponde até 45° e uma mudança no traço latente produz uma mesma da probabilidade de acerto, se $a_j > 1$ a inclinação corresponde até 90° e a probabilidade de acerto aumenta.
- e_{ij} é um erro aleatório nesta regressão distribuído de acordo com a função de ligação.

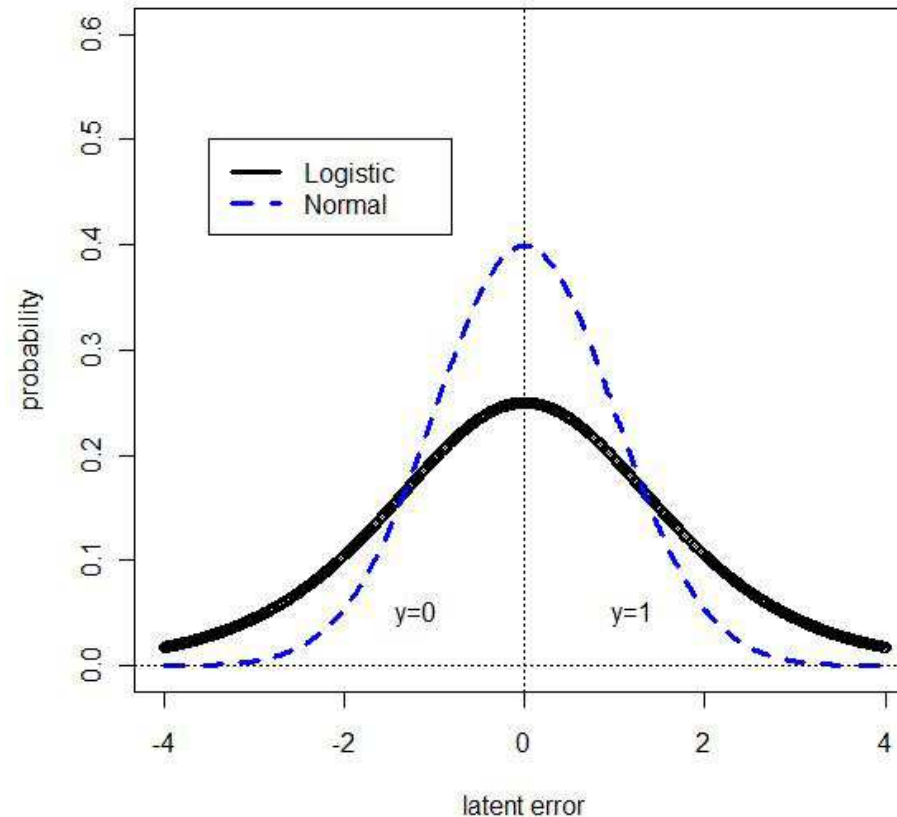


Figura 6. Erro latente da regressão entre a variável latente auxiliar, subjacente à resposta correta do item ($y=1$) ou à resposta incorreta ($y=0$), e o traço latente.

2.5 Verossimilhança aumentada do modelo TRI

Usando a nova formulação, a verossimilhança aumentada considerando a variável latente auxiliar Z_{ij} é:

$$L(\theta, \eta | D) = \prod_{i=1}^n \prod_{j=1}^I f(Z_{ij}; m_{ij}, 1) I(Z_{ij}, y_{ij}),$$

em que $f(\cdot)$ corresponde à função de densidade da variável latente auxiliar (ou do erro latente)

e $I(Z_{ij}, y_{ij})$ corresponde à função indicadora:

$$I(Z_{ij}, y_{ij}) = I(Z_{ij} > 0)I(y_{ij} = 1) + I(Z_{ij} \leq 0)I(y_{ij} = 0).$$

Os modelos TRI tradicionais binários, sob perspectiva Bayesiana, podem ser revisados em:

- Fox, J. -P. (2010). *Bayesian Item Response Modeling: Theory and Applications*. New York: Springer.
- Bazán, J., Valdivieso, L. e Calderón, A. (2010). *Enfoque bayesiano en modelos de Teoría de Respuesta al Ítem*. Reporte de Investigación. Serie B. Nro 25. Departamento de Ciências. PUCP. Disponível em:
<http://argos.pucp.edu.pe/~jlbazan/download/Reportef27.pdf>
- Bazán, J., Calderón, A. e Valdivieso, L. (2009). *Modelos de Teoría de Respuesta al Item Bajo enfoque Bayesiano*. Minicurso. XXVII COLOQUIO DE LA SOCIEDAD MATEMÁTICA PERUANA. Agosto 10 al 14. Puno, Perú. Disponível em:
http://www.guzlop-editoras.com/web_des/est01/Modelos_enfoque_bayesiano_JB.pdf

O modelo 1L (Rasch) pode ser revisado em:

- Bazán, J. (2007). *Modelos de respuestas al item usando Winbugs para Psicometría. Décimocuartas Jornadas en Estadística e Informática.* ECUADOR 2007. Octubre 17 al 19. Guayaquil. Ecuador. Disponível em:

http://www.icm.espol.edu.ec/jornadas/14/archivos/Diapositivas/BazánJorge/BazánJorge_Modelos de respuestas al item usando Winbugs para Psicometría.zip

O modelo de 2P parâmetros pode ser revisado em:

- Bazán, J. L., Bolfarine, H. e Leandro, A. R. (2006). Sensitivity analysis of prior specification for the probit-normal IRT model: an empirical study. *Estadística*. Journal of The Inter-American Statistical Institute, **58**, 17-42.

Disponível em: <http://argos.pucp.edu.pe/~jlbazan/download/bazanestadistica.pdf>

O modelo de três parâmetros pode ser revisado em:

- Sahu, S. K. (2002). Bayesian Estimation and Model Choice in Item Response Models. *Journal of Statistical Computation and Simulation*, **72**, 217-232.

3. NOVAS LIGAÇÕES NOS MODELOS TRI DICOTÔMICOS

- | | |
|--|--------|
| 3.1 O modelo de Ogiva Skew Normal da TRI | pág 34 |
| 3.2 Os modelos LPE e RLPE da TRI | pág 37 |
| 3.3 Interpretação do novo parâmetro de item | pág 40 |
| 3.4 Aplicação do modelo LPE a uma prova de Proficiência
em Matemática | pág 53 |
| 3.5 Outros desenvolvimentos e futuros trabalhos | pág 61 |

- Nos modelos TRI tradicionais, são consideradas CCI simétricas; este é o caso do modelo Logístico e o modelo de Ogiva Normal.
- No entanto, como tem sido observado por Samejima (2000), Bazán et al (2006) e Bolfarine e Bazán (2010), CCI assimétricas podem ser incorporadas considerando um novo parâmetro de item que controla a forma da curva.
- Como acontece na regressão binária, esta assimetria é necessária em situações em que é observada respostas com baixas ou altas proporções de acertos 1's.
- Apresentamos três modelos TRI com assimetria nas CCI recentemente formulados.

3.1 O modelo de Ogiva Skew Normal da TRI

Este modelo, denotado como modelo OSN, proposto por Bazán et al (2006), assume que:

$$Y_{ij} | \theta_i, \eta_j \sim \text{Bernouilli}(p_{ij})$$

$$p_{ij} = P(Y_{ij} = 1 | \theta_i, \eta_j) = \Phi_{SN}(m_{ij}; d_j), i = 1, \dots, n, j = 1, \dots, I$$

em que $-1 \leq d_j \leq 1$ é uma parâmetro de forma, m_{ij} o preditor linear latente e Φ_{SN} denota a f.d.a. da distribuição Skew Normal (Azzalini, 1985) com função de densidade:

$$\phi_{SN}(z; d) = 2\Phi(z)\Phi\left(\frac{d}{\sqrt{1-d^2}}z\right) \text{ em que } \phi \text{ e } \Phi \text{ correspondem à densidade e à acumulada da}$$

Normal padrão.

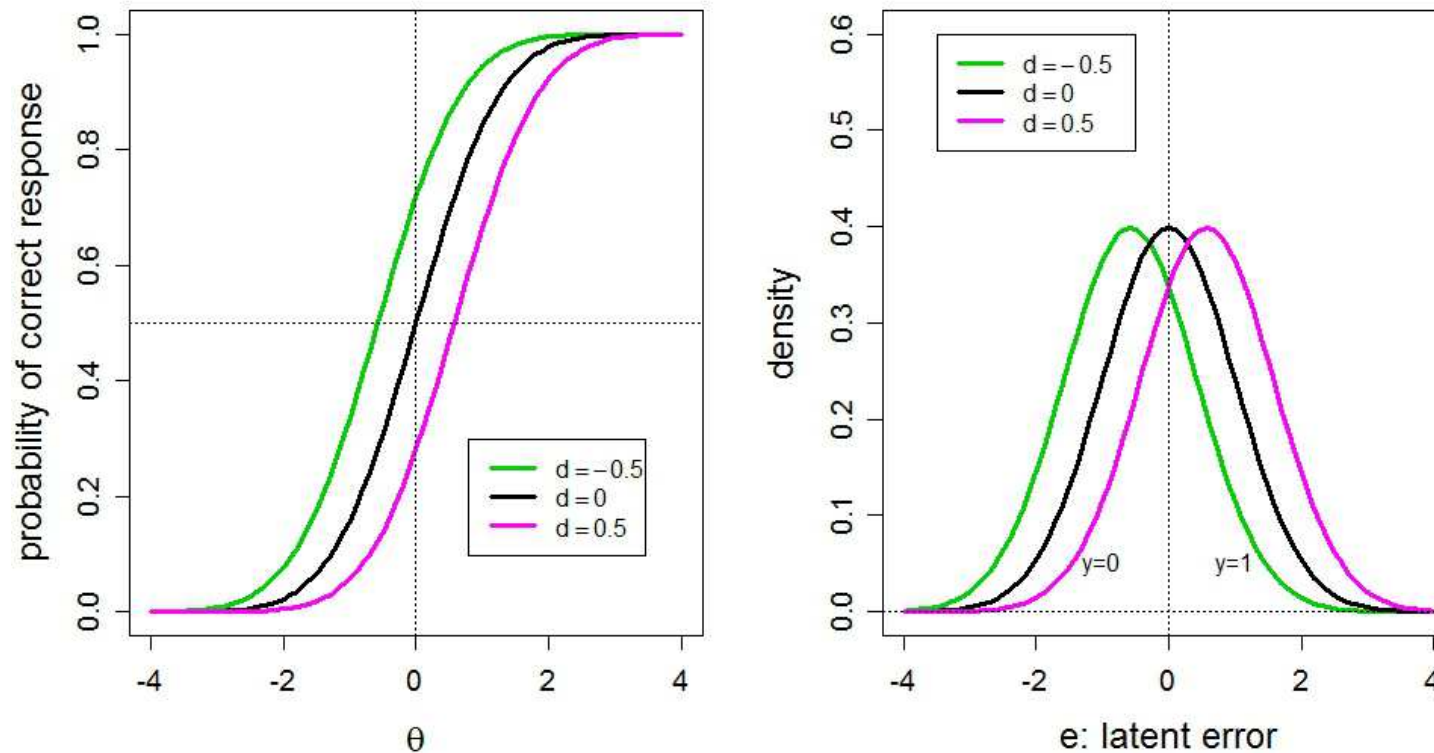


Figura 7. CCI para diferentes valores do traço latente θ e do erro latente da regressão entre a variável latente, subjacente ao acerto ($y=1$) ou fracasso ($y=0$) do item, e a habilidade no modelo de Ogiva Skew-Normal ($a = 1, b = 0$). Note que quando $d = 0$ o modelo 2P é obtido.

Para maiores detalhes ver:

Bazán, J. L., Branco, D. M. e Bolfarine, H. (2006). A skew item response model. *Bayesian Analysis*, **1**, 861- 892. Disponível em:

<http://ba.stat.cmu.edu/journal/2006/vol01/issue04/bazan.pdf>

3.2. Os Modelos LPE e RLPE da TRI

O Modelo Logístico de Expoente Positivo (LPE) foi proposto por Samejima (2000). A reflexão do modelo Logístico de Expoente Positivo ou Logístico de Expoente Positivo Reverso (RLPE) foi formulado por Bolfarine e Bazán (2010). Estes modelos assumem que:

$$Y_{ij} | \theta_i, \eta_j \sim \text{Bernouilli}(p_{ij})$$

$$p_{ij} = P[Y_{ij} = 1 | \theta_i, a_j, b_j] = F_{d_j}(m_{ij}), \quad i = 1, \dots, n, \quad j = 1, \dots, I$$

em que $-1 \leq d_j \leq 1$ é um parâmetro de forma, m_{ij} o preditor linear latente e F_{d_j} e uma função de distribuição acumulada indexada por d_j .

Para LPE $F_{\lambda_j} = L(m_{ij})^{\lambda_j}$ e para RLPE $F_{\lambda_j} = 1 - L(-m_{ij})^{\lambda_j}$ as quais corresponde as f.d.a.

$F_1(m_{ij}) = 1 - (1 + e^{m_{ij}})^{-\lambda_j}$ e $F_2(m_{ij}) = (1 + e^{-m_{ij}})^{-\lambda_j}$ respectivamente com $\lambda_j = \frac{1+d_j}{1-d_j}$, que correspondem às funções de ligação Scobit e Power Logito.

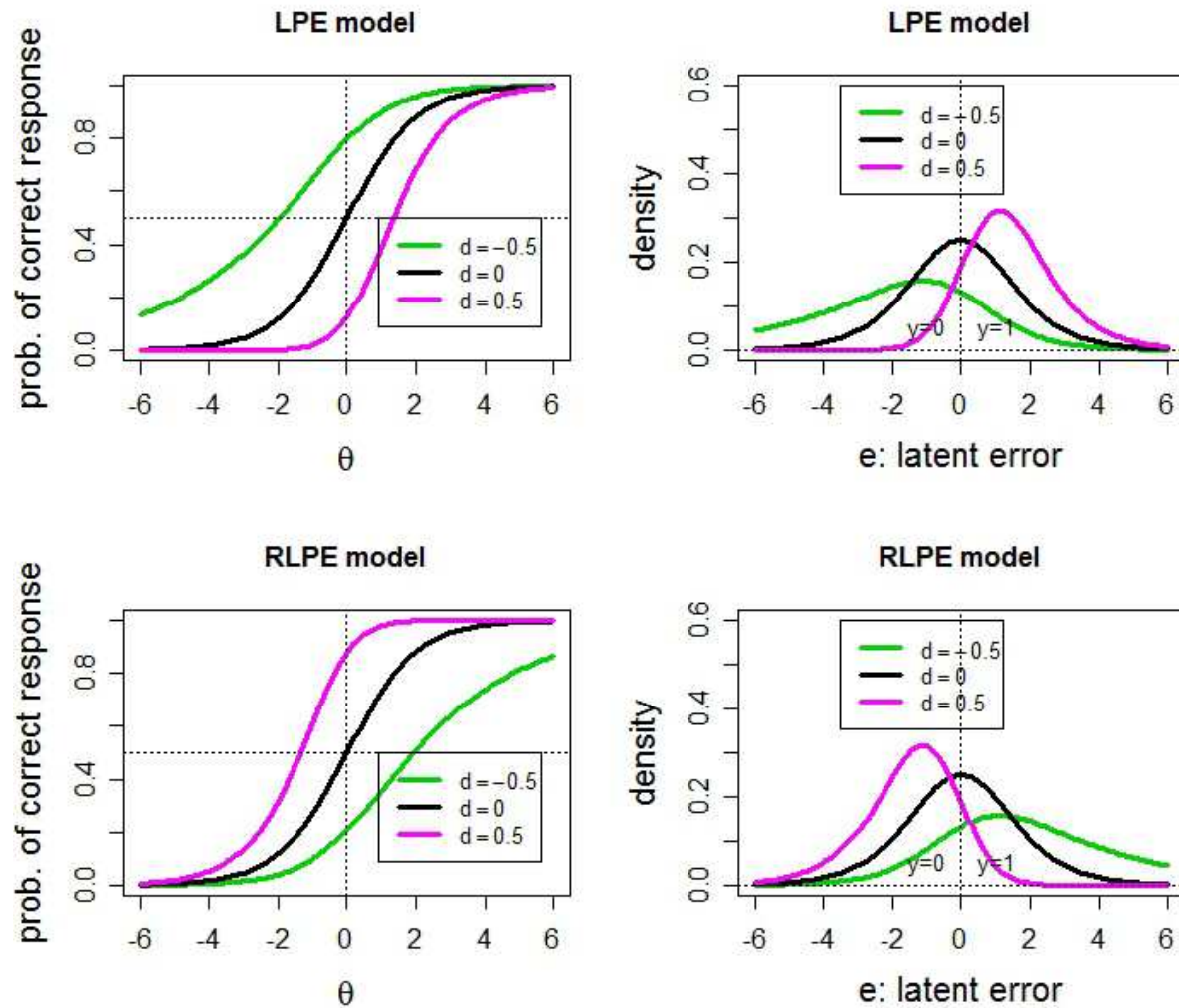


Figura 8. CCI e densidades do erro latente nos modelos LPE e RLPE ($a = 1, b = 0$). Modelo 2L ($d = 0$) é um modelo particular.

Veja detalhes em:

- Samejima, F. (2000). Logistic positive exponent family of models: Virtue of asymmetric item characteristics curves. *Psychometrika*, **65**, 319-335.

- Bolfarine, H. e Bazán, J. L. (2007). *Skewed Logit Item Response Models*. CLAPEM, Congreso Latinoamericano de Probabilidad y Estadística Matemática, Lima-Peru.

- Bolfarine, H. e Bazán, J. L. (2010) Bayesian Estimation of the Logistic Positive Exponent IRT model. *Journal of Educational Behavioral Statistics*, **35**, 693-713.

3.3 Interpretação do novo parâmetro do item

Novo parâmetro de item e assimetria do erro da regressão latente

O parâmetro d_j nos modelos LPE, RLPE e OSN encontra-se associado com a forma assimétrica da curva CCI ou de modo equivalente com a assimetria considerada no erro latente da regressão da variável auxiliar latente a respeito do traço latente que é subjacente a resposta correta ou incorreta do item.

A Figura 9 mostra a medida de assimetria definida por Arnold e Groeneveld (1995) em relação ao valor do parâmetro d correspondente para os modelos LPE, RLPE e OSN. Note que quando o valor de $d = 0$, temos ausência de assimetria.

Nos modelos LPE e OSN:

- Quando o valor $d > 0$ a CCI apresentam assimetria positiva e temos que pensar que o erro latente tem assimetria positiva. No caso contrário, $d < 0$ tem assimetria negativa.

Note também que para o modelo RLPE o efeito é reverso.

- Quando o valor $d > 0$ a CCI apresentam assimetria negativa e temos que pensar que o erro latente tem assimetria negativa. No caso contrário $d < 0$, o tem assimetria positiva.

Logo veremos que:

- Nos modelos LPE e OSN: d positivos penalizam as probabilidades de acerto (bonificam traço latente) e d negativos bonificam as probabilidades de acerto (penalizam traço latente). No modelo RLPE: d negativos penalizam as probabilidades de acerto (bonificam habilidades) e d positivos bonificam as probabilidades de acerto (penalizam traço latente).

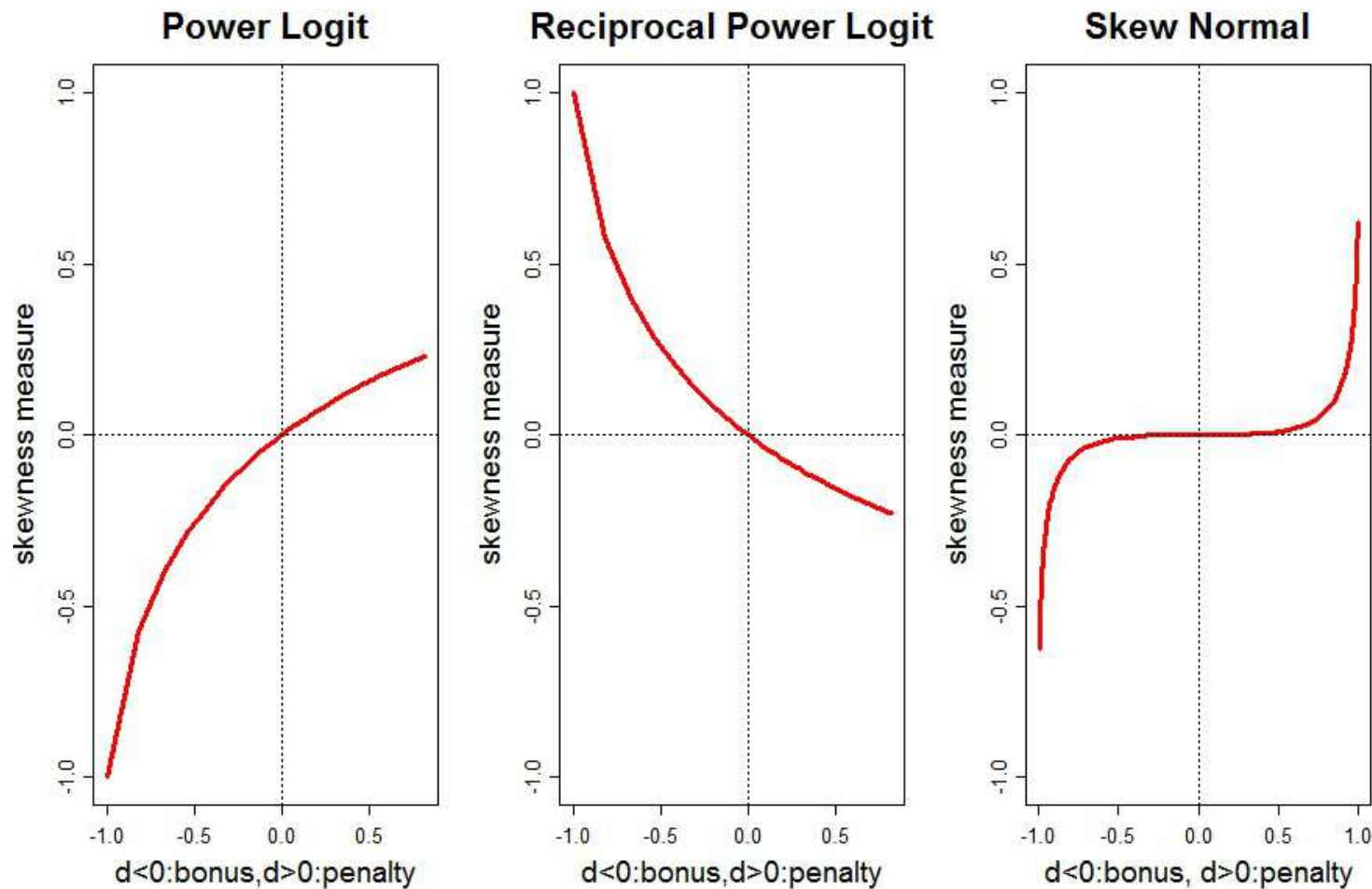


Figura 9. Assimetria como função do parâmetro d nos modelos LPE (distribuição Power Logito), RPLE (distribuição Power Logito Reversa) e OSN (distribuição Skew-Normal)

Probabilidade de acerto do item e parâmetro d

Nos modelos LPE e OSN, temos que:

- Quanto maior $d_j > 0$ então a probabilidade de acerto do item é penalizada a respeito do caso de ausência de assimetria.
- Quanto maior $d_j < 0$ então a probabilidade de acerto do item é bonificada a respeito do caso de ausência de assimetria.

No modelo RLPE temos que:

- Quando maior $d_j < 0$ então a probabilidade de acerto do item é penalizada a respeito do caso de ausência de assimetria.
- Quando maior $d_j > 0$ então a probabilidade de acerto do item é bonificada a respeito do caso de ausência de assimetria.

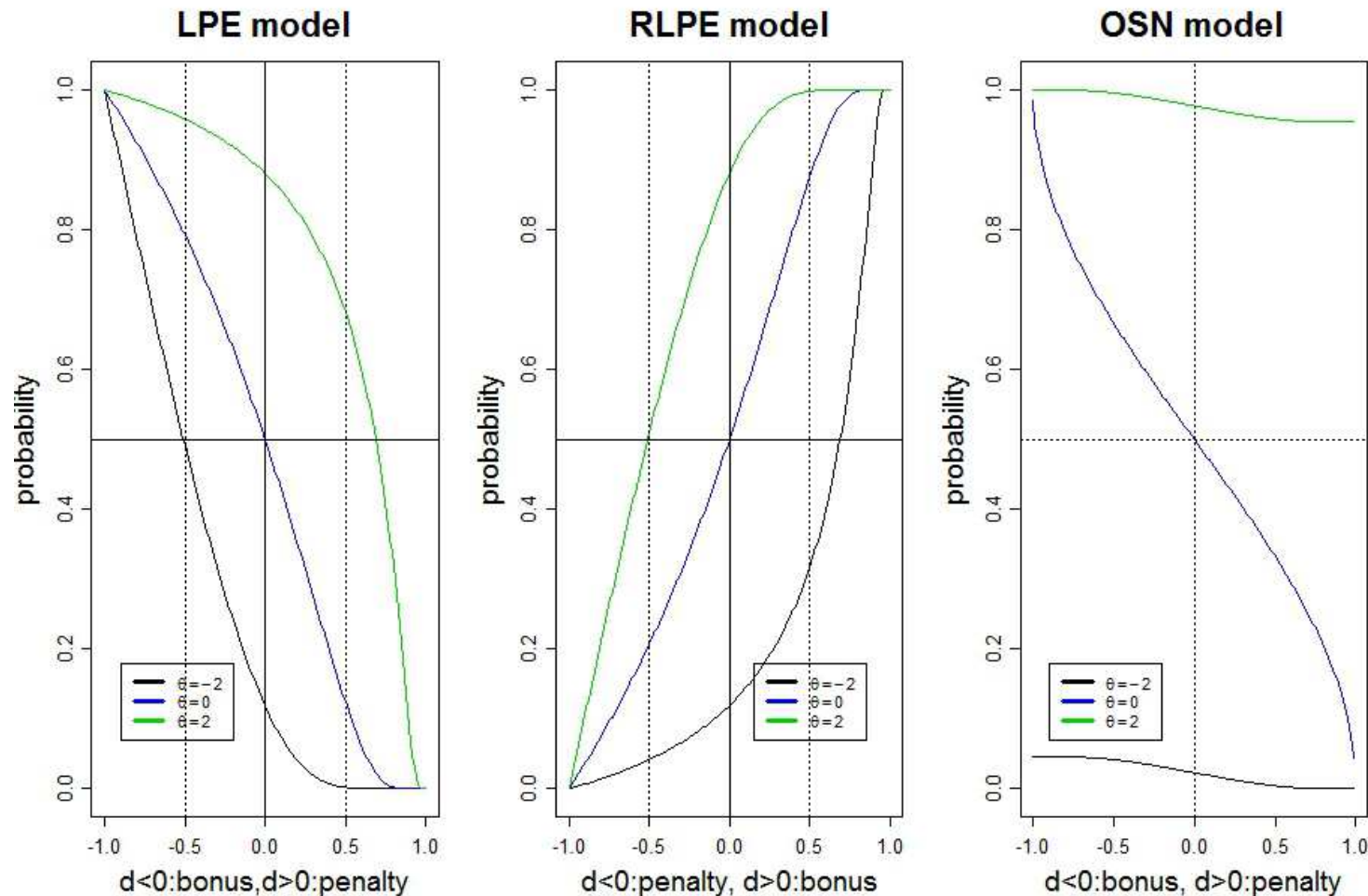


Figura 10. Probabilidade da resposta correta como função do parâmetro d nos modelos LPE, RPLE e OSN ($a = 1, b = 0$) para diferentes níveis de traços latentes. O modelo OSN tem baixa penalidade (bonificação para níveis extremos dos traços latentes).

Dificuldade e Parâmetro d

No modelo LPE (OSN) temos que:

- Quanto maior $d_j > 0$ então as curvas características dos itens são deslocadas à direita e embaixo a respeito do caso de ausência de assimetria. Quanto maior $d_j < 0$ então as curvas características dos itens são deslocadas a esquerda e acima a respeito do caso de ausência de assimetria.
- Para uma mesma probabilidade (proporção de acertos), conforme d cresce, corresponde a um maior nível de traço latente e, no caso contrário, conforme d decresce, corresponde a um menor nível de traço latente. Maior bonificação e maior penalidade nos traços latentes, respectivamente.

- Para um mesmo nível de traço latente conforme d cresce, corresponde a uma menor probabilidade (proporção de acertos) e no caso contrário, conforme d decresce corresponde a uma maior probabilidade. Maior penalização e maior bonificação nas proporções de acerto, respectivamente.
- O parâmetro de dificuldade é preservado no modelo, curvas características de itens fáceis estão à esquerda e curvas características de itens difíceis estão à direita.

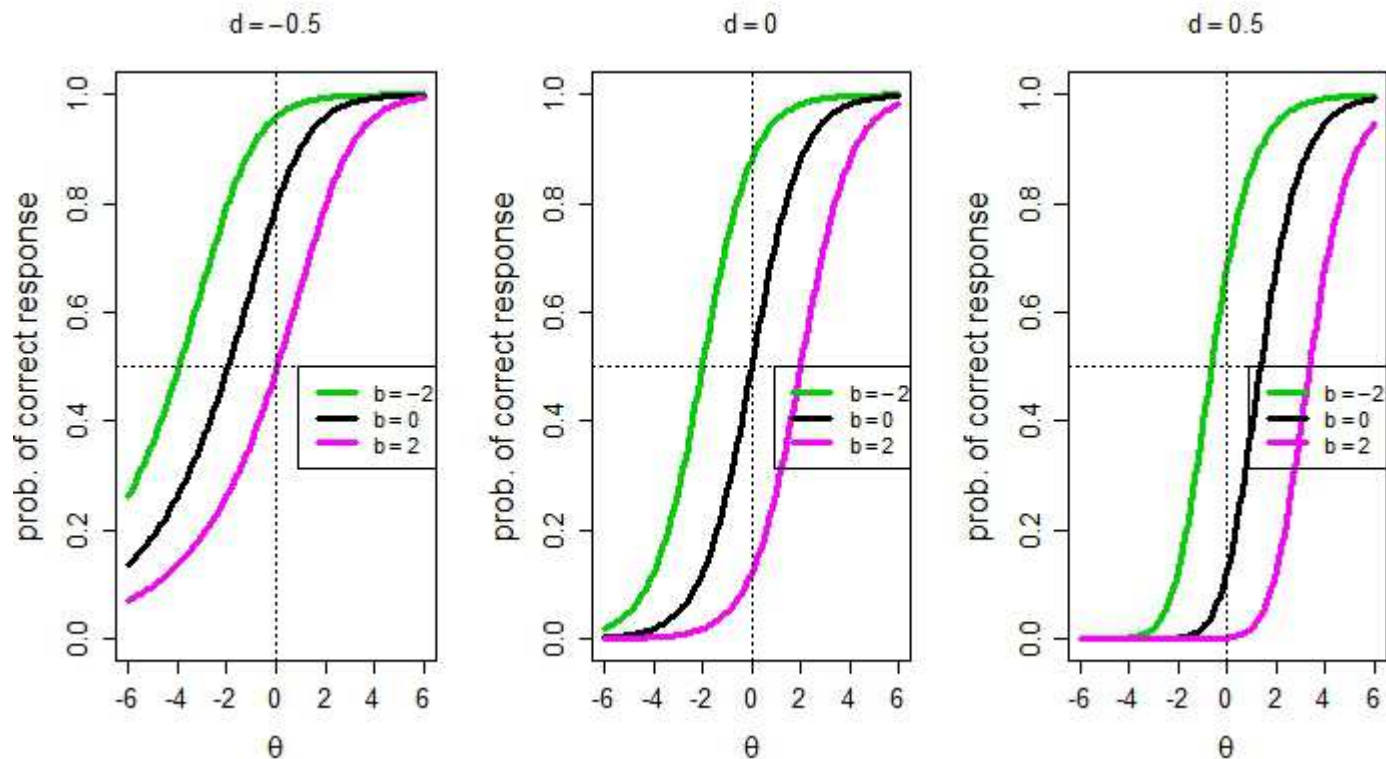


Figura 11. CCI para três valores do parâmetro d no modelo LPE ($a = 1, b = 0$) para itens com diferentes valores do parâmetro de dificuldade.

Efeito reverso acontece no modelo RLPE:

- Quanto maior $d_j < 0$ então as curvas características dos itens são deslocadas à direita e embaixo a respeito do caso de ausência de assimetria. Neste caso, tem-se maior bonificação do nível de traço latente para uma mesma proporção de acertos ou maior penalização das proporções de acerto para um mesmo nível de traço latente.
- Quanto maior $d_j > 0$ então as curvas características dos itens são deslocadas a esquerda e acima a respeito do caso de ausência de assimetria. Neste caso, tem-se maior penalização do nível de traço latente para uma mesma proporção de acertos ou maior bonificação das proporções de acerto para um mesmo nível de traço latente.

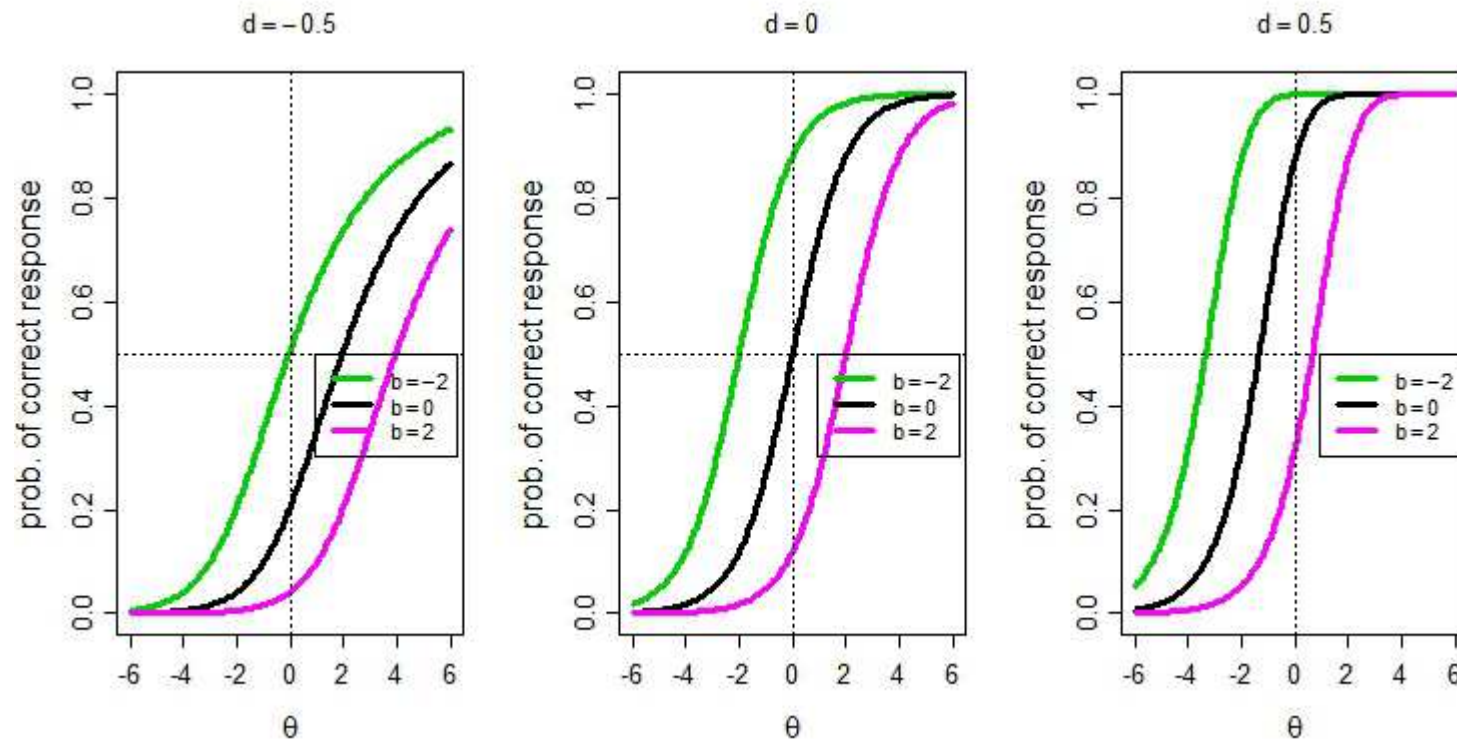


Figura 12. CCI para três valores do parâmetro d no modelo RLPE ($a = 1, b = 0$) para itens com diferentes valores do parâmetro de dificuldade.

Discriminação e parâmetro d

No modelo LPE (OSN) temos que:

- A probabilidade de acerto para uma habilidade média ($\theta = 0$) (ponto de inflexão da CCI).
- Quanto maior $d_j > 0$ então o ponto de inflexão da CCI (probabilidade de acerto para uma habilidade média $\theta = 0$) é menor a respeito do caso de ausência de assimetria.
- Quanto maior $d_j < 0$ então o ponto de inflexão da CCI (probabilidade de acerto para uma habilidade média $\theta = 0$) é maior a respeito do caso de ausência de assimetria.
- O parâmetro de discriminação é preservado no modelo, a inclinação das curvas características de itens mais discriminativos é maior do que aqueles menos discriminativos.

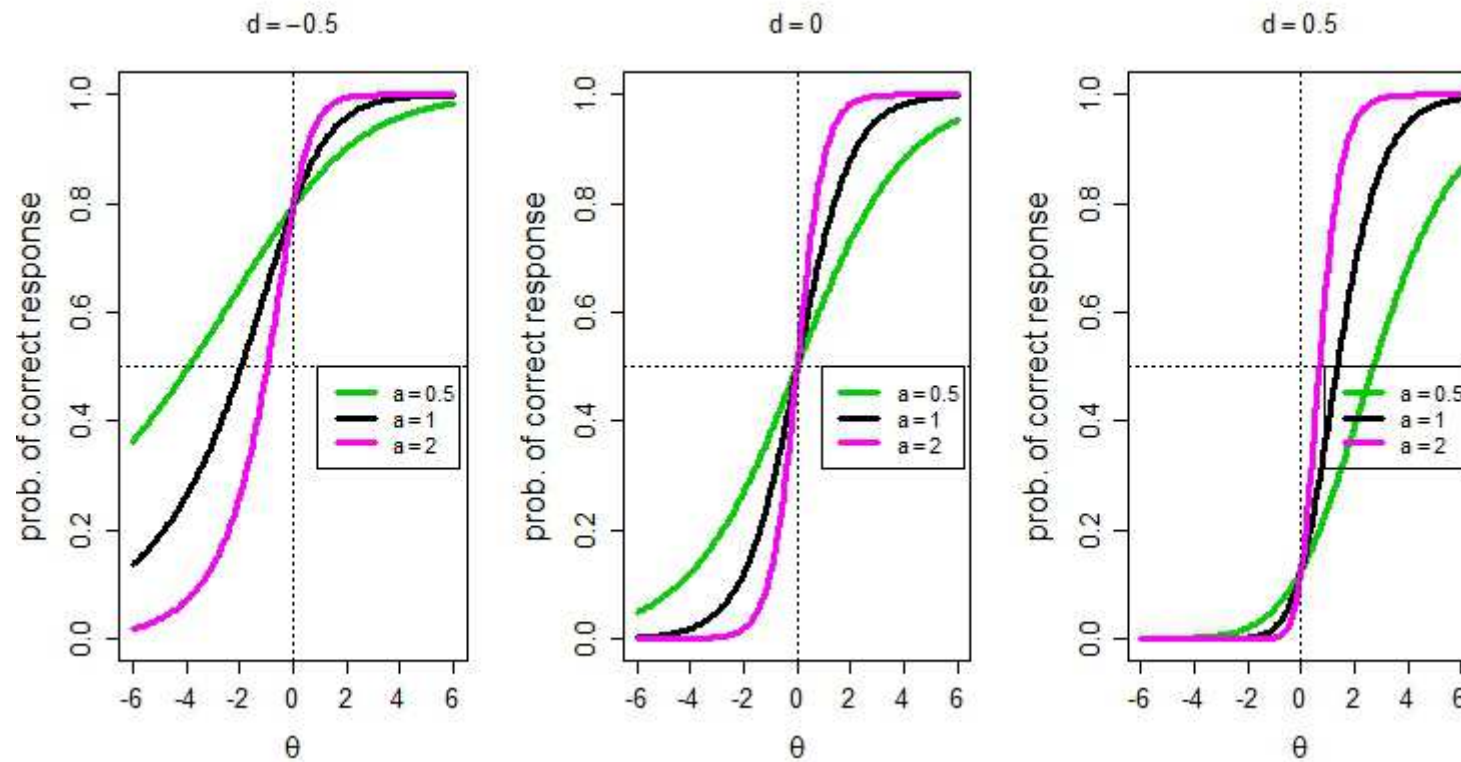


Figura 13. CCI para três valores do parâmetro d no modelo LPE ($a = 1, b = 0$) para itens com diferentes valores do parâmetro de discriminação.

Em resumo:

- O novo parâmetro tem a ver com a forma da CCI, não mexe com a dificuldade e a discriminação, e tem um efeito duplo.
- Quando positivo (negativo no RLPE), item mais complexo, penaliza as probabilidades ou bonifica os níveis do traço latente. Isto é ideal para modelagem de níveis baixos do traço latente, já que mudanças menores nesse intervalo se correspondem a pequenas mudanças da probabilidade de acerto.
- Quando negativo (positivo RLPE), item menos complexo, bonifica as probabilidades ou penaliza os níveis do traço latente. Isto é ideal para modelagem de níveis altos do traço latente, já que mudanças menores nesse intervalo se correspondem com pequenas mudanças da probabilidade de acerto.

3.4 Aplicação do modelo LPE a uma prova de Proficiência em Matemática

Consideramos os dados reportados por Bolfarine e Bazan (2010) que correspondem a uma prova de 18 questões de múltipla escolha aplicada a 974 alunos da 4ª série de escolas rurais e foi desenhada por uma unidade do Ministério de Educação do Peru.

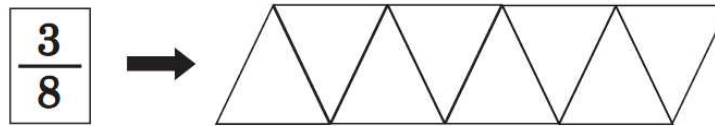
O escore médio do teste é de 8.27 e o desvio padrão é de 4.20.

Os índices de assimetria e curtosis são, respectivamente, iguais a -0.075 e -0.836.

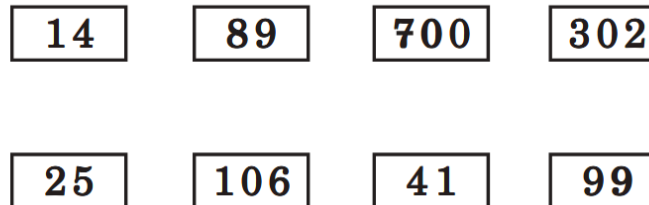
O alfa de cronbach é de 0.83 e a média das proporções de acerto é 0.449.

Exemplos destes itens são apresentados a seguir:

14. **PINTA** en la figura la fracción que se indica:



2. Fíjate bien en estos números:



Marca con un aspa (X) todos los números **MENORES QUE 100**.

11. Completa la suma:

$$300 + \square = 350$$

Figura 14. Item 14, 2 e 11 para os dados de Matemática (Bolfarine e Bazan, 2010).

TABLE 1
Results Comparing the Skewed Logistic IRT Models With Logistic IRT Models

Models		p	$Dbar$	$Dhat$	DIC	EAIC	EBIC
Symmetric	1L	992	17,227.3	16,548.1	17,906.6	19,211.3	26,920.9
logistic IRT	2L	1,010	16,886	16,104.3	17,667.8	18,906	26,755.5
models	3L	1,028	17,021.1	16,506.7	17,535.4	19,077.1	27,066.5
Asymmetric	LPE	1,028	16,885.4	16,728.3	17,042.5	18,941.4	26,930.8
skewed logistic	RLPE	1,028	16,832.3	16,708.7	16,955.8	18,888.3	26,877.6
IRT models							

Note. DIC = deviance information criterion; EAIC = expected Akaike's information criterion; EBIC = expected Bayesian information criterion; IRT = item response theory; LPE = logistic positive exponent; RLPE = reflection logistic positive exponent.

Considerando diferentes critérios de comparação de modelos, os modelos considerando CCI assimétricas tiveram melhor ajuste que os correspondentes modelos simétricos.

TABLE 2

Item Parameters for Alternative IRT Models for Item 14, Item 2, and Item 11 in Math Data

Items	Models	Item Parameters			
		Discrimination, <i>a</i>	Difficulty, <i>b</i>	Guessing, <i>c</i>	Acceleration (deceleration), λ
Item 14	2L	1.003	-0.358		
	3L	1.065	-0.105	0.108	
	LPE	0.934	-1.214		2.087
	RLPE	4.559	-1.603		0.208
Item 2	2L	1.2777	-0.453		
	3L	1.607	-0.088	0.165	
	LPE	1.768	0.253		0.559
	RLPE	1.060	0.380		2.107
Item 11	2L	1.826	0.092		
	3L	2.065	0.220	0.067	
	LPE	1.680	0.039		2.060
	RLPE	2.070	0.023		1.046

Note. IRT = item response theory; LPE = logistic positive exponent; RLPE = reflection logistic positive exponent.

Mostramos as estimativas dos diferentes parâmetros para quatro modelos TRI. Note que aqui é apresentado o parâmetro de aceleração $\lambda = \frac{d+1}{d-1}$.

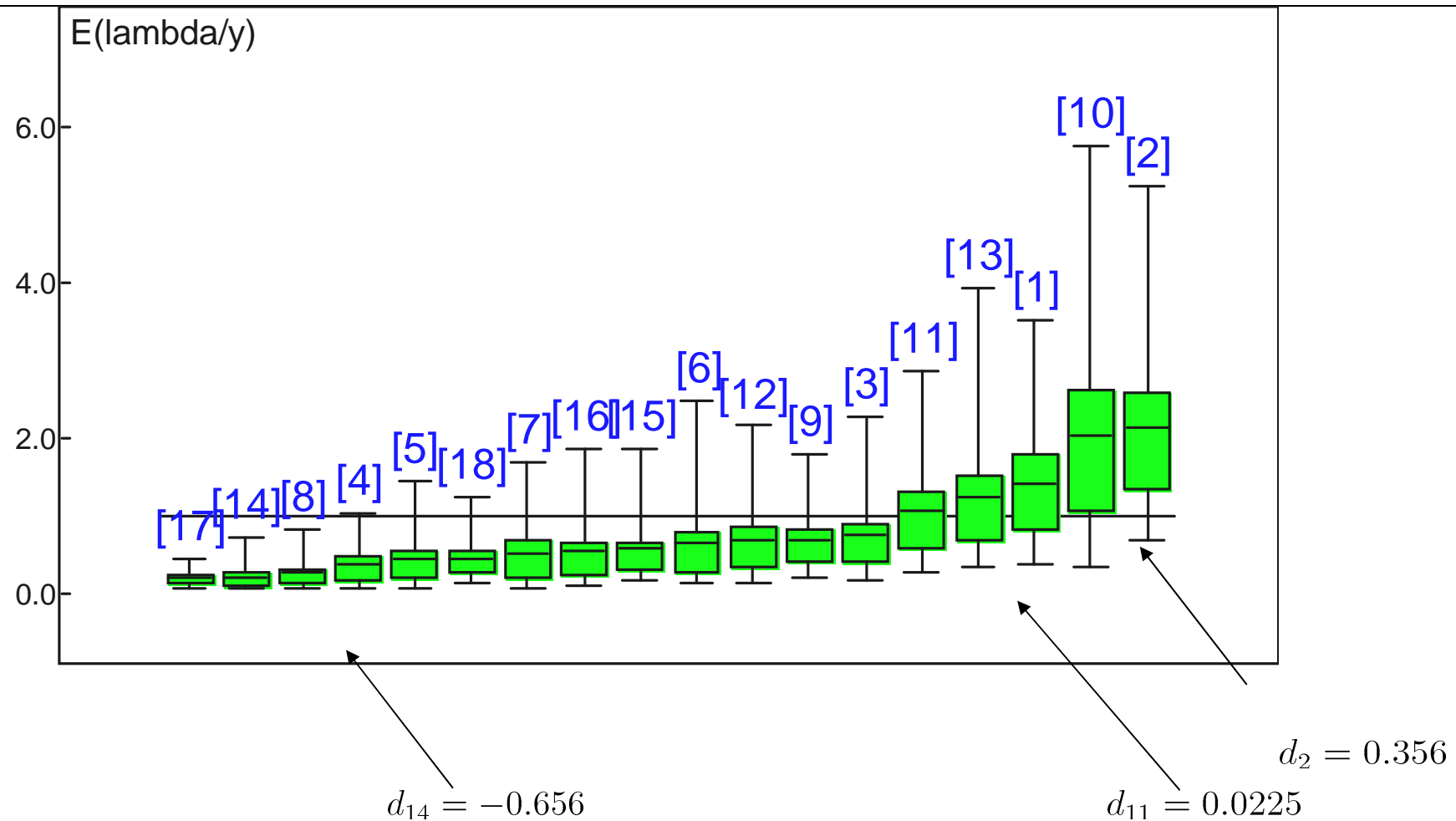


Figure 15: Box-Plots para os parâmetros λ dos 18 itens dos dados de Matemática sob o modelo RLPE (Bolfarine e Bazan, 2010).

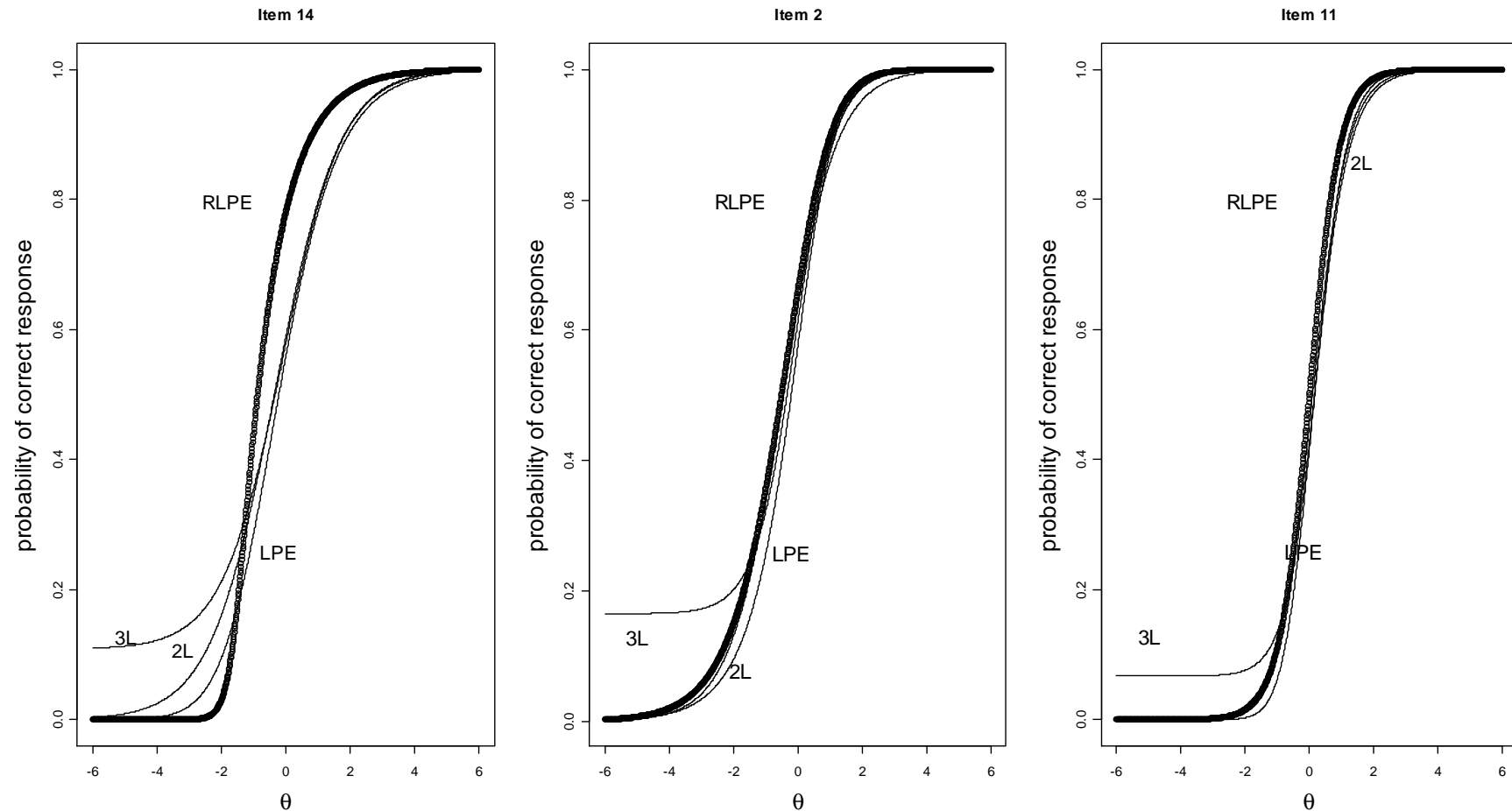


Figure 16: CCI para os itens 14, 2 e 11 sob os modelos 3L, 2L, RLP e LPE para os dados de Matemática (Bolfarine e Bazan, 2010).

No item 14:

- Comparando as probabilidades de resposta correta dos modelos L, LPE e RLP encontramos que o modelo RLPE apresenta menores valores para $\theta < -1.5$ e maiores probabilidades para valores $\theta > -1.5$.
- Embora fácil, a resposta ao item não é direta e deve ser considerado complexo, pois precisa que um aluno desenvolva passos sucessivos de processos sequenciais (como indicado pelo Samejima, 2000) para responder satisfatoriamente o problema completo.
- Como o processo é simples e o item é fácil, os alunos são penalizados em termos de probabilidade de resposta correta a respeito do caso de ausência de assimetria para um dado nível de habilidade. Ou para um mesmo nível de probabilidade (acertos) dado as habilidades estimadas são melhores do que obtidas a respeito do caso de ausência de assimetria.

- Corresponde a um caso limite e a uma CCI do tipo função de grau (heavyside) e foi obtida porque o item tem uma grande discriminação estimada ($a = 4.559$).
- Corresponde a uma CCI do tipo Guttman: para um baixo nível de habilidade ($\theta \leq -1.5$), a probabilidade de passar o item é praticamente zero, para níveis de habilidade intermediários, isto é ($-1.5 < \theta < 1.9$), uma pequena mudança na habilidade corresponde a uma grande mudança na probabilidade de responder corretamente o item. Finalmente para altos níveis de habilidade ($\theta \geq 1.9$), a probabilidade de responder corretamente o item é praticamente 1.

3.5 Outros desenvolvimentos e futuros trabalhos

- Detecção da assimetria e modelo de três parâmetros:

- Santos, V. L. F.. *Teoria de Resposta ao Item: uma abordagem generalizada das Curvas Características dos Itens*. Rio de Janeiro: UFRJ/ IM - DME, 2009. Orientadores Dani Gamerman e Tufi Machado Soares.

- Outras ligações:

- Goldstein, H. (1980). Dimensionality, bias, independence and measurement scale problems in latent trait test score models. *British Journal of Mathematical and Statistical Psychology*, **23**, 234-246.

- Pinheiro, A. N. C. *Modelos de resposta ao item com função e ligação t-assimétrica*. São Carlos: UFSCar, 2007. Orientador Josemar Rodrigues.

- Pinheiro, A. e Rodrigues, J. (2006). Modelos de resposta ao item com função de ligação t-assimétrica. *Rev. Bras. Biom.*, São Paulo, **26** (1) , 131-143.

http://jaguar.fcav.unesp.br/RME/fasciculos/v26/v26_n1/A8_Josemar.pdf

- Azevedo, Caio L. N. e Migon, H. S. (2012). *Bayesian inference in an item response theory model with a generalized student t link function*. In: XI Brazilian Meeting on Bayesian Statistics: EBEB 2012, Amparo-SP.

- Tavares H. R. e Serra, C. M. V. (2005). *Função de resposta assimétrica em modelos de resposta ao item*. In: Escola de Modelos de Regressão, 2005, São Pedro.

- Uso na construção de provas:

- Dodonov, Y. S. e Dodonova, Y. A. (2010). Assymmetric item characteristics curve in item response theory model construction and application. *Procedia Social and Behavioral Sciences*, **5**, 1592-1595.

- Samejima, F. (2000). Logistic positive exponent family of models: Virtue of asymmetric item characteristics curves. *Psychometrika*, **65**, 319-335.

- Extensões para a resposta ordinal.

- Molenaar, D. , Dolan C. V. de Boeck. P. (2012). The Heteroscedastic Graded Response Model with a Skewed Latent Trait: Testing Statistical and Substantive Hypotheses related to Skewed Item Category Functions. *Psychometrika*, **77**, 455-478.

- Extensões para o caso de resposta policotômica.
- Extensões para testlets.
- Inferência clássica dos modelos apresentados.
- Estudo de diagnósticos dos modelos apresentados.
- Relacionamento deste parâmetro com DIF e o parâmetro de acerto casual.
- Relacionamento com o modelo de Rasch.
- Novas propostas de CCI assimétricas na TRI.
- Muitas aplicações.
- Interpretação do novo parâmetro em questionários da área clínica e em aplicações como Marketing e análise política.

4. INFERÊNCIA BAYESIANA NA TRI

4.1 Verossimilhança da TRI	pág 67
4.2 Especificações das prioris	pág 70
4.3 Estimação Bayesiana usando MCMC	pág 74
4.4 Critérios de comparação de modelos usando saídas MCMC	pág 78
4.5 Uso de pacotes	pág 81
4.6 Tarefa	pág 83
4.7 Adicional	pág 84

- Na Inferência Bayesiana (Carlin e Louis, 2000), os parâmetros de interesse são assumidos como variáveis aleatórias e precisa-se estabelecer uma distribuição de probabilidade a priori que reflete nosso conhecimento prévio de seu comportamento.
- Na modelagem da TRI, as prioris são discutidas em Albert (1992), Rupp, Dey e Zumbo (2004), Bazán, Bolfarine e Leandro (2006), Sahu (2002) e Fox (2010).
- Combinando a função de verossimilhança e a distribuição a priori, podemos obter a distribuição a posteriori dos parâmetros de interesse. Quando esta distribuição não tem forma fechada ou conhecida, procedimentos MCMC podem ser usados para obter uma amostra simulada da posteriori.
- Usando uma programação mínima e formulando uma estrutura hierárquica do modelo e possível programar MCMC para a TRI.
- Inferência Bayesiana da TRI usando MCMC é demorada se o número de parâmetros é alto.

- Inferência Bayesiana é facilitada com o uso de diferentes métodos MCMC programados em vários pacotes como R e SAS, mas também no pacote WinBUGS e OpenBUGS.
- Para maiores detalhes a respeito do uso deste programa para a Inferência Bayesiana sugerimos os livros: Congdon (2005), Congdon (2010) e Ntzoufras (2009).
- Sintaxes no R (por exemplo, MCMC pack de Martin e Quinn, 2002) e Matlab (Johnson e Albert, 1999, Sheng, 2008) são também disponíveis na rede.

4.1 Verossimilhança da TRI

De modo geral, a função de verossimilhança do modelo TRI de 3 parâmetros é dada por:

$$L(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \theta, | \mathbf{y}) = \prod_{i=1}^n \prod_{j=1}^k \{F_{d_j} [a_j \theta_i - b_j; c_j]\}^{y_i} \{1 - F_{d_j} [a_j \theta_i - b_j; c_j]\}^{1-y_i}$$

em que $F_{d_j} [a_j \theta_i - b_j; c_j]$ é a f.d.a. da distribuição assimétrica indexada pelo parâmetro d_j associado com a ICC assimétrica.

Os modelos Logístico (ligação Logito), Ogiva Normal (ligação Probit), LPE e RLPE (ligações Power Logito e sua reversa) e seus casos particulares de 1 ou 2 parâmetros são escritos no *Bayesian Modeling* considerando esta versão de verossimilhança.

O modelo de Ogiva Skew Normal (de 1 e 2 parâmetros) é o único que considera uma versão aumentada ou ampliada baseada na regressão latente, análoga a mostrada nos modelos simétricos. A verossimilhança original é dada por:

$$L(\theta, \mathbf{a}, \mathbf{b}, \mathbf{d} | D_{obs}) = \prod_{i=1}^n \prod_{j=1}^k [\Phi_{SN}(m_{ij}; d_j)]^{y_{ij}} [1 - \Phi_{SN}(m_{ij}; d_j)]^{1-y_{ij}}.$$

Mais usando a formulação de regressão latente:

$$Y_{ij} = \begin{cases} 1, & Z_{ij}^* > 0; \\ 0, & Z_{ij}^* \leq 0, \end{cases} \quad Z_{ij}^* \sim N(-d_j v_{ij} + m_{ij}, 1 - d_j^2) \quad \text{and} \quad V_{ij} \sim HN(0, 1)$$

em que HN corresponde à distribuição Normal Positiva ou Half Normal.

Assim a verossimilhança aumentada dado $D_2 = (\mathbf{z}^*, \mathbf{v}, \mathbf{y})$ é dada por:

$$L(\theta, \mathbf{a}, \mathbf{b}, \mathbf{d} | D_2) \propto \prod_{i=1}^n \prod_{j=1}^k \phi(Z_{ij}^* ; -d_j v_{ij} + m_{ij}, 1 - d_j^2) \phi(V_{i,j} ; 0, 1) I(v_{ij} > 0) p(y_{ij} | z_{ij}^*),$$

em que $p(y_{ij} | z_{ij}^*) = I(z_{ij}^* > 0) I(y_{ij} = 1) + I(z_{ij}^* \leq 0) I(y_{ij} = 0)$.

Isto torna fácil sua escrita em código BUGS.

Maiores detalhes são dados em Bazán, Branco e Bolfarine (2006).

4.2 Especificações das prioris

É um aspecto importante da análise Bayesiana. É muito mais importante para tamanhos de amostra pequenos. No presente trabalho consideramos distribuições a priori próprias conhecidas com variância grande assim como prioris independentes (veja Nzoufras, 2009).

Na literatura da TRI existe consenso com respeito à especificação da priori para o traço latente:

$$\theta_i \sim N(0, 1) \text{ para } i = 1, \dots, n,$$

No entanto, para os parâmetros a_j e b_j dos itens tem sido propostas diferentes prioris (ver Rupp et al., 2004).

Embora exista evidência empírica da correlação das distribuições a posteriori desses

parâmetros (Patz e Junker, 1999), é complicado especificar a priori esta estrutura de dependência, porém são preferidas priors independentes e comuns para cada parâmetro dos itens incluindo o caso do novo parâmetro d ,

$$\pi(\theta, \mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) = \prod_i^n \phi(\theta_i) \prod_j^k \pi_1(a_j) \pi_2(b_j) \pi_3(c_j) \pi_4(d_j).$$

em que $\phi(\cdot)$ é a densidade da Normal padrão e $\pi_1(\cdot)$, $\pi_2(\cdot)$, $\pi_3(\cdot)$, $\pi_4(\cdot)$ são as distribuições a priori para os parâmetros a_j , b_j , c_j e d_j , respectivamente.

Estrutura Hierárquica:

$$y_{ij}|a_j, b_j, \theta_j, \lambda_j \sim Ber\left(F_\lambda[a_j\theta_i - b_j; c_j]\right)$$

$$\theta_i \sim N(0, 1)$$

$$a_j \sim \pi_1(\mu_a, \sigma_b)$$

$$b_j \sim \pi_2(\mu_b, \sigma_b^2)$$

$$c_j \sim \pi_3(r, s)$$

$$d_j \sim \pi_4(m, n)$$

$$i = 1, \dots, n; j = 1, \dots, I$$

Sugerimos a leitura de Bazán, Bolfarine e Leandro (2006); Bazán, Branco e Bolfarine (2006); Bolfarine e Bazán (2010); Patz e Junker (1999); Sinharay (2004) e Fu, Tao e Shi (2009) ao respeito da escolha das prioris.

Sugerimos, especialmente no caso de amostras pequenas, o uso de:

- $a_j \sim HN(\mu_a, \sigma_a^2)$ com $\mu_a = 1$ e $\sigma_a^2 = 0.5$ para a_j e então é esperado $E(a_j) = 1.1126$ e $V(a_j) = 0.3747$ com a $HN(.)$ a distribuição Normal Positiva.
- $b_j \sim N(\mu_b, \sigma_b^2)$ para b_j , $\mu_b = 0$ e $\sigma_b^2 = 2$, isto é, $E(b_j) = 0$ y $V(b_j) = 1$
- $c_j \sim Beta(5, 17)$ e então é esperado $E(c_j) = 0.227$, $V(c_j) = 0.0076$.
- $d_j \sim U(-1, 1)$ e então é esperado $E(d_j) = 0$, $V(c_j) = 0.33$. Isto é usado atualmente para os modelos Ogiva Skew Normal. Para os modelos LPE e RLPE podes se usar $\lambda_j \sim gamma(0.25, 0.25)$ o que implica que $E(\lambda_j) = 1$, $V(\lambda_j) = 4$, com $q(0.025) = 1.055 \times 10^{-6}$ e $q(0.975) = 6.86647$.

4.3 Estimação Bayesiana usando MCMC

Considerando as versões da verossimilhança e a especificação das prioris pode-se programar a estimação Bayesiana com métodos Markov Chain Monte Carlo os quais facilitam de modo simples a amostragem eficiente da distribuição marginal a posteriori dos parâmetros de interesse.

Quando considerada a primeira versão de verossimilhança um algoritmo tipo Metrópolis (Gilks e Wild, 1992) é programado WinBugs.

Quando considera a segunda versão de verossimilhança um algoritmo tipo Gibbs pode ser programado.

A seguir mostramos o caso do modelo de 2 parâmetros de ogiva Normal ou modelo 2P.

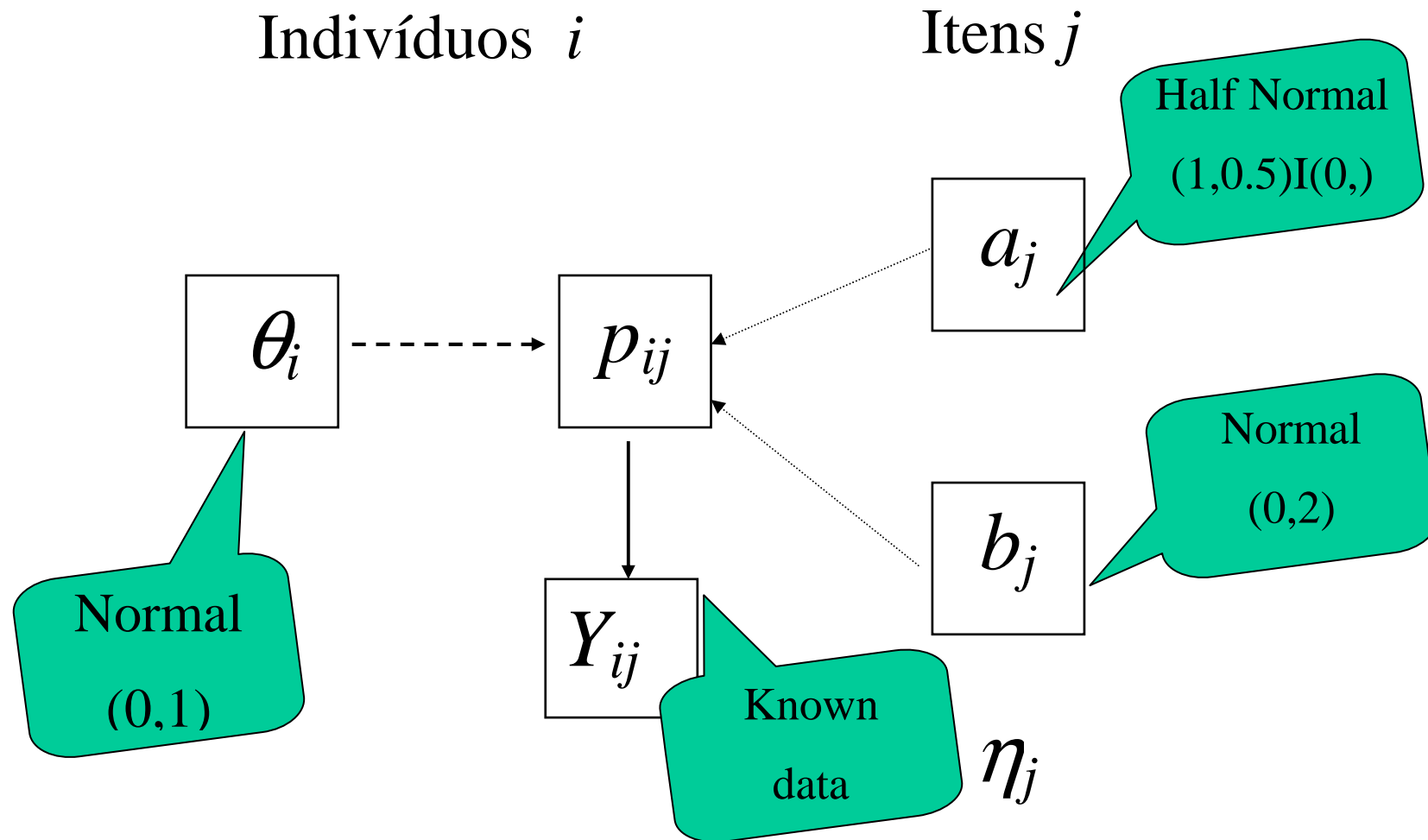


Figure 17: Esquema da Inferência Bayesiana do modelo TRI dicotômico usando MCMC

```
#2P MH TYPE
model{
  for (i in 1:n) {    for (j in 1:I) {
    m[i,j]<-a[j]*theta[i]-b[j]
    p[i,j] <- phi(m[i,j])
    y[i,j] ~ dbern(p[i,j])}    }
  #abilities priors
  for (i in 1:n) {    theta[i]~dnorm(0,1)    }
  #items priors
  for (j in 1:I) {
    # usual priors
    #Bazan et al (2006)
    b[j] ~ dnorm(0,0.5)
    a[j] ~ dnorm(1,2)I(0,)
    bc[j] <- b[j]/a[j]    }
}
```

Figure 18: Código Bugs da Inferência Bayesiana do modelo TRI dicotômico usando Metrópolis-Hasting

```
# 2P model type Gibbs
model{
  for (i in 1 : n) {
    for (j in 1 : I) {
      m[i,j] <- a[j]*theta[i] - b[j]
      z[i,j] ~ dnorm(m[i,j],1)I(lo[y[i,j]+1],up[y[i,j]+1])
    }
    theta[j] ~ dnorm(0, 1)
  }

# Priors
  for (j in 1:I) {
    b[j] ~ dnorm(0, 0.5)
    a[j]~ dnorm(1,2)I(0,)
    bc[j] <- b[j]/a[j] }
lo[1] <- -50; lo[2] <- 0;  ## i.e., z | y=0 ~ N(m,1)I(-50,0)
up[1] <- 0; up[2] <- 50;  ## i.e., z | y=1 ~ N(m,1)I(0,50)
}
```

Figure 19: Código Bugs da Inferência Bayesiana do modelo TRI dicotômico usando Gibbs Sampling

4.4 Critérios de comparação de modelos usando saídas MCMC

Existem uma variedade de metodologias para comparar o ajuste de modelos bayesianos alternativos, no entanto o principal critério é a Deviance information criterion (DIC) proposta por Spiegelhalter et al. (2002), e os Expected information criteria para Akaike (EAIC) e para Schwarz ou Bayesian (EBIC) reportados em Bolfarine e Bazan (2010). Estes critérios são baseados na *posterior mean of the deviance*: $E[D(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \theta)]$, em que:

$$D(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \theta) = -2\ln(p(\mathbf{y}|\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \theta)) = -2 \sum_{i=1}^n \ln P(Y_{ij} = y_{ij}|\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \theta),$$

A qual também é uma medida de ajuste que pode ser aproximada usando a saída MCMC considerando:

$$D_{bar} = \frac{1}{G} \sum_{i=1}^G D(\mathbf{a}^g, \mathbf{b}^g, \mathbf{c}^g, \mathbf{d}^g, \theta^g),$$

em que o índice g representa a g -ésima realização do total de G realizações e D é a *Bayesian deviance*.

EAIC, EBIC e DIC podem ser estimados usando a saída MCMC considerando:

$$\widehat{EAIC} = Dbar + 2p,$$

$$\widehat{EBIC} = Dbar + p \log N,$$

$$\widehat{DIC} = Dbar + \widehat{\rho_D} = 2Dbar - Dhat,$$

em que p é o número de parâmetros no modelo ($n + 3k$, por exemplo, nos modelos 3P ou 3L), N é o número total de observações ou respostas, e neste caso corresponde a $N = k \times n$ e ρ_D , é o *effective number of parameters* proposto por Spiegelhalter et al. (2002).

Dada a comparação de dois modelos alternativos, o modelo que ajusta melhor aos dados é aquele com menor $Dbar$, DIC , $EBIC$ e $EAIC$.

EAIC e EBIC, $2p$ e $p \log N$ são fixos para penalizar a média a posteriori da deviance, fato que não acontece no DIC.

Como não existe consenso no uso do DIC (veja discussão em Spiegelhalter et al., 2002), o uso de mais de um critério é mais apropriado para fazer a comparação de modelos.

4.5 Uso de pacotes

- Sob perspectiva Bayesiana os modelos apresentados podem ser implementados nos pacotes WinBUGS ou OpenBUGS. No R usando os pacotes R2WinBUGS ou BRugs ou no SAS com o proc mcmc.
- Algumas referências específicas do uso de WinBUGS na TRI são:
 - Bazán, J. L. (2012). Item Response Theory in WinBUGS using Bayesian Modeling Software. *Software Highlight in ISBA Bolletín*, **19** (1), 8-10.
 - Kim, J-S. e Bolt, D. M. (2007). Markov chain Monte Carlo estimation of item response models. *Educational Measurement: Issues and Practice*, **26**, 38-51.
 - González, J. (2010). Bayesian methods in psychological research. The case of IRT. *International Journal of Psychological Research*, **3** (1), 164-176.
 - Curtis, M. S. (2010). BUGS Code for Item Response Theory. *Journal of Statistical Software*, **36**: Snippet 1. 1-34.

- *Bayesian Modeling* é um pequeno pacote (Bazan, 2011), livre, que gera a sintaxis para a inferência bayesiana do tipo BUGS dos seguintes 14 modelos.

Simétricos:

- Logísticos: 1L, 2L, 3L.
- Ogiva Normal: 1P, 2P, 3P.

Assimétricos:

- Logísticos de Expoente Positivo: LPE1, LPE2, LPE3.
- Reverso do Logístico de Expoente Positivo: RLPE1, RLPE2, RLPE3.
- Ogiva Skew Normal: 1SP, 2SP.

Para maiores detalhes ver:

- Bazan, J. (2011). *Bayesian Modeling User's Guide*. Sciences Department. Pontificia Universidad Catolica del Peru. Disponível em: <http://www.icmc.usp.br/pessoas/jlbazan/software.html>

4.6 Tarefa

1. Implemente o modelo LPE para os dados do teste de Matemática usando WinBUGS.
2. Proponha outras ligações alternativas. Use os o programa Bayesian Modeling para gerar códigos dos modelos alternativos.
3. Considerando o critério DIC, qual modelo é o mais apropriado para os dados?
4. Quais itens apresentam um parâmetro de penalidade significativo? Interprete estes resultados em termos da aplicação.
5. Recentemente foi proposta uma ligação usando a acumulada da distribuição Power Normal (Bazán, Romeo e Rodrigues, 2014). Desenvolva um código para um modelo TRI considerando esta nova ligação.

4.7 Adicional

A modo de comparação, replique os resultados usando os seguintes pacotes para a inferência clássica:

- Pacotes comerciais: IRTPRO, BILOG-MG, XCALIBRE

Maiores detalhes em <http://www.ssicentral.com/irt/> <https://assess.com/xcart/home.php?cat=22>

- Pacotes no R: ltm, TAM, mirt, irtoys,

Maiores detalhes em <http://cran.r-project.org/web/views/Psychometrics.html>

- Para outros modelos, ver os pacotes disponíveis em:

http://en.wikipedia.org/wiki/Psychometric_software

Observação. Na aproximação clássica a estimação é dividida em duas etapas. Na primeira, chamada de calibração, são estimados os parâmetros dos itens e na segunda, chamada de estimação, são estimadas as proficiências.

5. EXTENSÕES DOS MODELOS DA TRI

- 5.1 CONSIDERANDO TRAÇOS LATENTES ASSIMÉTRICOS** pág 86
- 5.2 CONSIDERANDO RESPOSTAS POLITÔMICAS** pág 94
- 5.3 CONSIDERANDO VÁRIOS TRAÇOS LATENTES MEDIDOS** pág 111

5.1 CONSIDERANDO TRAÇOS LATENTES ASSIMÉTRICOS

- Grande evidência empírica de traços latentes é normal na psicologia e em outras áreas.
- Vários trabalhos de teses de doutorado. Quatro teses de doutorado em Estatística no IME-USP: Bazan (2005), Matos (2008), Azevedo (2008) e Targino (2010), envolvendo pesquisadores como Heleno Bolfarine, Márcia Branco e Dalton Andrade.
- Uso da distribuição Skew Normal como distribuição para os traços latentes considerando diferentes parametrizações da Skew Normal (Azzalini, Sahu, centrada).
- Ênfases na inferência Bayesiana.
- Ênfases no caso multidimensional.

- Uma área com pesquisa em andamento a nível teórico.
- A programação dos modelos não encontra-se disponível para os usuários.
- Alguns estudos de simulação são prometedores mais ainda existem poucas aplicações.
- Esta faltando uma organização dos trabalhos e pode ser motivo de outro minicurso com ênfase em traços latentes não gaussianos.

2004-2005

Bazán, J. L, Bolfarine, H. e Branco, D. M. (2004). *A new family of asymmetric models for item response theory: A Skew-Normal IRT Family*. Technical report (RT-MAE-2004-17). Department of Statistics. University of São Paulo.

Bazán, J. L. (2005). *Uma familia de modelos de resposta ao item normal assimétrica*. Teses de Doutorado. Instituto de Matemática e Estatística, USP.

$$\theta_i \sim SN(\mu, \sigma^2, \omega), \quad i = 1, \dots, n$$
$$\mu \sim N(0, 1), \quad 1/\sigma^2 \sim Gamma(0.01, 0.01), \quad \omega \sim U(-1, 1).$$

- Modelo hierárquico para traços latentes considerando hiperparâmetros.
- A mesma priori para os indivíduos: locação, escala e formas comuns.
- É possível estimar os traços latentes a posteriori e, a média, a escala e a forma da população de indivíduos a posteriori.

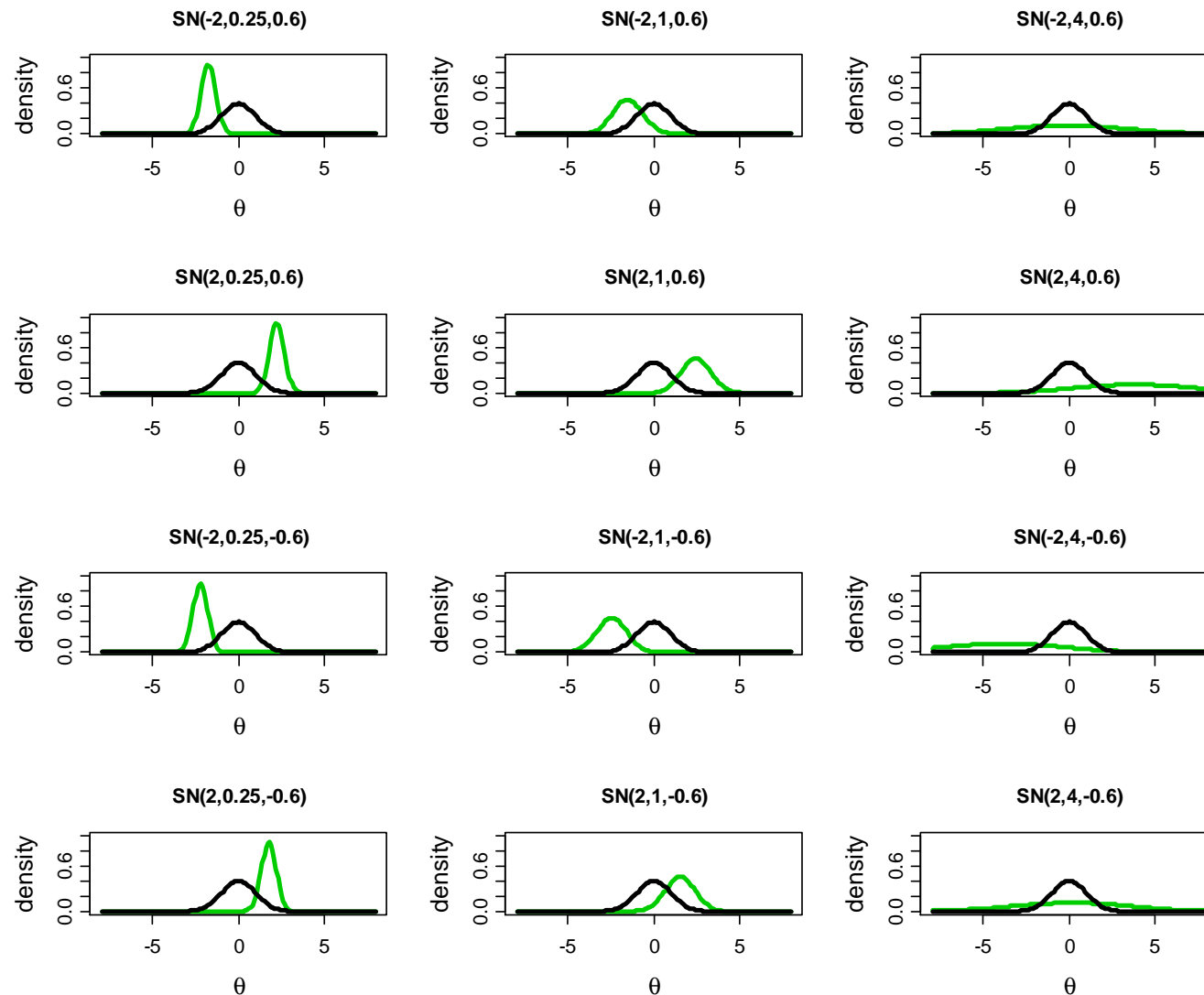


Figura 20: Traços latentes considerando a distribuição Skew-Normal de Azzalini.

2007

- Azevedo, C. L. N. e Andrade, D. F. (2007). *Latent trait estimation in nominal response model: latent trait asymmetry and hierarchical and empirical framework*. IME-USP.
- Molenaar, D. (2007). Accounting for non-normality in latent regression models using a cumulative normal selection function. *Measurement and Research Department Reports*, **3**, Arnhem: Cito.

2008

- Sass, D. A., Schmitt, T. A. e Walker, C. M. (2008). Estimating Non-Normal Latent Trait Distributions within Item Response Theory Using True and Estimated Item Parameters. *Applied Measurement in Education*, **21**, 65-88.
- Matos, G. S. (2008). *Modelos multidimensionais da TRI com distribuições assimétricas para os traços latentes*. Tese de Doutorado, IME-USP, São Paulo-SP.
- Azevedo, C. L. N. (2008). *Modelos longitudinais de grupos múltiplos multiníveis na Teoria da Resposta ao Item: métodos de estimação e seleção estrutural sob uma perspectiva bayesiana*. Tese de Doutorado, IME-USP, São Paulo-SP.

2010

- Broccoli S. e Cavrini G. (2010). *A Bayesian IRT model for mixed responses and skew latent variable*. Scientific Meetings of SIS, 45th Scientific Meeting of the Italian Statistical Societ.

- Nojosa, T. R. (2010). *Inferência Bayesiana em Modelos Multidimensionais de Resposta ao Item*. Tese de Doutorado, IME-USP, São Paulo-SP.
- Molenaar, D., Dolan, C. e Verhelst, N. (2010). Testing and modeling non-normality within the one factor model. *British Journal of Mathematical and Statistical Psychology*, **63**, 293-317.
- Moleenar, D., Dolan, C., Wicherts, J. e van der Mass, H. (2010). Modeling Differentiation of Cognitive Abilities within the Higher-Order Factor Model using Moderated Factor Analysis. *Intelligence*, **38**, 611–624.
- Montenegro, A. (2010). Multidimensional Item Response Theory Models where the Ability has a Latent Linear Structure. Teses de Doutorado. Universidad Nacional de Colombia.

2011-2014

- Azevedo, C. L. N., Bolfarine, H. e Andrade, D. F. (2011). Parameter recovery for a skew-normal IRT model under a Bayesian approach: hierarchical framework, prior and kernel sensitivity and sample size. *Journal of Statistical Computation and Simulation*, **82** (11), 1679-1699.
- Azevedo, C. L. N., Bolfarine, H. e Andrade, D. F. (2011). Bayesian inference for a skew-normal IRT model under the centred parametrization, *Computational Statistics and Data Analysis*, **55**, 353-365.

- Xu, X. e Jia, Y (2011). The Sensitivity of Parameter Estimates to the Latent Ability Distribution. Research Report. ETS RR–11-40.
- Stingo, F. C., Stanghellini, E. e Capobianco, R. (2011). On the estimation of a binary response model in a selected population. *Journal of Statistical Planning and Inference*, **141**, 3293-3303.
- Azevedo, C. L. N., Bolfarine, H. e Andrade, D. F. (2012). Parameter recovery for a skew-normal IRT model under a Bayesian approach: hierarchical framework, prior and kernel sensitivity and sample size. *Journal of Statistical Computation and Simulation*, **82**, 1679-1699.
- Santos, J. R. S., Azevedo, C. L. N. e Bolfarine, H. (2013). A multiple group item response theory model with centered skew-normal latent trait distributions under a Bayesian framework. *Journal of Applied Statistics*, **40**, 2129-2149.
- Bazán, J. L., Branco, M. e Bolfarine, H. (2014) Extensions of the Skew-Normal Ogive Item Response Model. *Brazilian Journal of Probability and Statistics*. **28** (1), 1-23.

5.2 CONSIDERANDO RESPOSTAS POLITÔMICAS

- Muitos formatos de resposta de instrumentos de avaliação (escalas, testes, inventários, questionários, checklist, etc.) contêm mais de duas opções de respostas igualmente válidas sendo denominados neste caso respostas politômicas.
- As respostas podem ser nominais ou ordinais, isto é, podem ser simplesmente respostas diferentes ou podem seguir uma ordem pre determinada como é o caso, por exemplo, de escalas do tipo Likert variando de nenhuma importância a total concordância, nenhuma diferenciação a total diferenciação ou de total discordância a total concordância.

- O uso destes modelos ainda é menor comparado com o uso de procedimentos de testes estatísticos clássicos em que as escalas são assumidas como intervalares, mas que na realidade são ordinais.
- Se as respostas são dicotomizadas podem ser usados os modelos TRI dicotômicos, no entanto, informação a respeito das diferentes opções de resposta são perdidas.
- Estes modelos podem ser classificados de modo similar ao que acontece no caso dos modelos Logísticos politômicos definindo como passar de uma categoria de resposta (etapa) a outra (a acumulativa, a de crédito parcial e a condicional).
- A ênfase atual é na inferência clássica existindo poucos trabalhos na inferência Bayesiana.
- Uma área com pesquisa tanto prática quanto teórica e existem poucos programas para implementar estes modelos.

Algumas dissertações que podemos citar são:

- Eduardo, V. F. (2014). *Modelos da Teoria de Resposta ao Item assimétricos de grupos múltiplos para respostas politômicas nominais e ordinais sob um enfoque bayesiano*. 2014. Mestrado em Estatística. Unicamp.
- Pereira, S. R. S. (2012). *Contribuições ao estudo do modelo de resposta nominal*. Mestrado em Estatística. Unicamp.
- Dos Santos, M. A. (2011). *Análise bayesiana da tri com resposta gradual*. Mestrado em Estatística. Mestrado em Estatística. USP.
- Vargas, E. G. T. (2013). *Modelos alternativos de respuesta graduada con aplicaciones en calidad de servicios*. Maestría en Estadística. Pontificia Universidad Católica del Peru.

- Bazán J., Merino, M. H. e Mazzon, J. A. (2011). Classificação de modelos de resposta ao item policotômicos com aplicação ao marketing. *Revista Brasileira de Estatística*, **72**, 7-39.

Modelos de itens

a) *Item de resposta gradual*

Exemplo 1.

Item: Usar o *mobile banking* para fazer transferências bancárias é uma boa idéia.

Categorias de resposta:

- a) Discorda muito
- b) Discorda pouco
- c) Nem concorda nem discorda
- d) Concorda pouco
- e) Concorda muito

Assume-se uma ordem estrita nas etapas do item. É claro que o avaliado precisa analisar todas as categorias antes de escolher uma, e se, por exemplo, ele escolhe a categoria “Nem concorda

nem discorda” significa que ele não optou por nenhuma das categorias prévias nem por nenhuma das categorias seguintes; portanto, seu escore deve ser 3.

b) Item de resposta parcial

Exemplo 2.

Item: Escreva e explique três atributos que caracterizam o *mobile banking* para um consumidor.

Categorias de resposta:

- a) Não escreveu nenhum
- b) Escreveu um atributo
- c) Escreveu dois atributos
- d) Escreveu três atributos

O avaliado constrói sua solução independentemente de conhecer todas as etapas do problema.

É o avaliador que qualifica sua resposta de acordo com o critério definido nas categorias do item, que neste caso pode ser de 0 a 3. O avaliador atribui os escores dependendo se cada etapa

é apresentada na solução do avaliado. Para cada etapa da solução presente, o avaliador atribui um "crédito" ou score igual a 1. Se o avaliador não encontra a etapa (falha), então o avaliado fica com os "créditos acumulados" ou o score até a etapa prévia. Observe que aqui o avaliado pode tentar todas as etapas, mas o score que ele consegue depende de quantas etapas ele completa corretamente.

c) Item sequencial

Exemplo 4.

Item: Qualificação do uso pelo consumidor do
mobile banking.

Categorias de resposta:

- a) Insuficiente
- b) Regular
- c) Boa
- d) Excelente

Assume-se uma ordem estrita nas etapas do item. No caso do item 4, semelhante ao item 1, as etapas do problema são conhecidas pelo avaliado e, também semelhante ao item 3, é atribuído um escore com o qual se consegue calcular a etapa correspondente.

É o avaliador que qualifica sua resposta de acordo com o critério definido nas categorias do item que, neste caso, foi de 0 a 3. O avaliador atribui os escores dependendo se cada etapa é apresentada na solução do avaliado. Observe que a diferença do item 3 em relação ao item 4 é que neste o avaliador não precisa observar a solução das etapas seguintes, pois claramente essas não estão presentes. Se acontecer uma falha numa etapa o processo cessa e as etapas seguintes não são consideradas.

Modelos politômicos

Tabela 2. Função de resposta e logitos de passo de uma etapa de um item policotômico (com respostas ordinais)

Funções de resposta de uma etapa do item (FREI) (Hemker, 2001)	Logitos de passo de uma etapa do item (L_{ijk}) (van der Ark, 2001)
<i>Acumulativo:</i> P_{ijk}^+	<i>Acumulativo:</i> $L_{ijk} = \text{logito}[P_{ijk}^+] = \log\left(\frac{P_{ijk}^+}{1 - P_{ijk}^+}\right)$
<i>De crédito parcial:</i> $\frac{P_{ijk}}{P_{ijk-1} + P_{ijk}}$	<i>De categorias adjacentes:</i> $L_{ijk} = \text{logito}\left[\frac{P_{ijk}}{P_{ijk-1} + P_{ijk}}\right] = \log\left(\frac{P_{ijk}}{P_{ijk-1}}\right)$
<i>Condicional:</i> $\frac{P_{ijk}^+}{P_{ijk-1}^+}$	<i>De razão-continuação:</i> $L_{ijk} = \text{logito}\left[\frac{P_{ijk}^+}{P_{ijk-1}^+}\right] = \log\left(\frac{P_{ijk}^+}{P_{ijk-1}^+ - P_{ijk}^+}\right)$

$\sum_{h=1}^m P_{ijh} = 1$ em que P_{ijh} é a probabilidade do indivíduo i optar pela categoria h das m no item j e

$P_{ijk}^+ = P(y_{ij} \geq k) = \sum_k^m P_{ijh}$ é a correspondente probabilidade de atingir o passo k (entre categorias) dos $(m-1)$ passos.

a) Modelo generalizado de resposta gradual:

$$\text{logito}[P_{ijk}^+] = \log\left(\frac{P_{ijk}^+}{1 - P_{ijk}^+}\right) = a_{jk}(\theta_i - b_{jk}) \Rightarrow P_{ijk}^+ = \frac{\exp[a_{jk}(\theta_i - b_{jk})]}{1 + \exp[a_{jk}(\theta_i - b_{jk})]} = \frac{1}{1 + \exp[-a_{jk}(\theta_i - b_{jk})]}$$

para $k=1, \dots, m-1, , ; i=1, \dots, n$ e $j=1, \dots, I$.

$b_{j1} < b_{j2} < \dots < b_{jm-1}$ quando $a_{jk} = a_j$ temos o modelo de resposta gradual (Samejima, 1969).

b) Modelo doblemente generalizado de Crédito parcial:

$$L_{ijk} = \text{logito}\left[\frac{P_{ijk}}{P_{ijk-1} + P_{ijk}}\right] = \log\left(\frac{P_{ijk}}{P_{ijk-1}}\right) = a_{jk}(\theta_i - b_{jk}) \Rightarrow P_{ijk} = \frac{\exp\left[\sum_{l=1}^k a_{jl}(\theta_i - b_{jl})\right]}{\sum_{h=1}^m \exp\left[\sum_{l=1}^h a_{jl}(\theta_i - b_{jl})\right]} \quad \text{para } k=1, \dots, m \quad ; \quad i=1, \dots, I \quad \text{e}$$

$j=1, \dots, n$, considerando-se por conveniência de notação $b_{iI} \equiv 0$.

Quando $a_{jk} = a_j$ temos o modelo generalizado de credito parcial (Muraki, 1992); quando $a_{jk} = 1$ temos o modelo de Máster, 1982).

c) Modelo Sequencial (Tutz, 1990):

$$L_{ijk} = \text{logito} \left[\frac{P_{ijk}^+}{P_{ijk-1}^+} \right] = \log \left(\frac{P_{ijk}^+}{P_{ijk-1}^+ - P_{ijk}^+} \right) = a_{jk} (\theta_i - b_{jk}) \text{ para } k=1, \dots, m-1 ; i=1, \dots, n \text{ e } j=1, \dots, l$$

$$P_{ijk} = \prod_{r=1}^k \frac{P_{ijr}^+}{P_{ijr-1}^+} - \prod_{r=1}^{k+1} \frac{P_{ijr}^+}{P_{ijr-1}^+} = \left\{ \prod_{r=1}^k \frac{\exp[a_{jk}(\theta_i - b_{jk})]}{1 + \exp[a_{jk}(\theta_i - b_{jk})]} \right\} \left[\frac{1}{\exp[a_{jk+1}(\theta_i - b_{jk+1})]} \right], k=1, \dots, m.$$

considerando-se por conveniência de notação:

$$\prod_{r=1}^0 \frac{P_{ijr}^+}{P_{ijr-1}^+} \equiv 1 \text{ e } \prod_{r=1}^{m+1} \frac{P_{ijr}^+}{P_{ijr-1}^+} \equiv 0$$

A partir destes três tipos de modelos, diversos modelos particulares podem ser obtidos como discutidos em Bazán, Merino e Mazzon (2011).

Modelo de resposta graduada

$$Y_{ij} \mid \theta_i, \xi_j \sim \text{categorica}(p_{ij1}, \dots, p_{ijm})$$

em que as probabilidades de responder a categoria k no item j é dada por:

$$p_{ijk} = P(Y_{ij} = k \mid \theta_i, \xi_j) = \begin{cases} 1 - P_{y_{ij1}}^+ = 1 - F_L(\eta_{ij1}) & \text{si } k = 1 \\ P_{y_{ijk}}^+ - P_{y_{ijk+1}}^+ = F_L(\eta_{ijk}) - F_L(\eta_{ij(k+1)}) & \text{si } 2 < k < m \\ P_{y_{ijm}}^+ = F_L(\eta_{ijm}) & \text{si } k = m \end{cases}$$

em que $F_L(\cdot)$ é a função de distribuição acumulada da Logística e $\eta_{ijk} = a_j(\theta_i - b_{jk})$ é o preditor latente com $\xi_j = (a_j, b_{j1}, \dots, b_{j(m-1)})$ o vetor de parâmetros do item j : a_j é a discriminação ou inclinação e b_{jk} é a posição do passo k no item j .

As probabilidades acumuladas são dadas por:

$$P_{y_{ijk}}^+(\theta_i) = P(Y_{ij} \geq k \mid \theta_i, \xi_j) = F_L(\eta_{ijk}) = \frac{1}{1 + e^{-\eta_{ijk}}}$$

```
# Modelo de resposta graduada usando probabilidades acumuladas a direita (similar ltm e IRTPRO)
model {
for ( i in 1: n ){
  for ( j in 1: I){ y[i , j ] ~ dcat ( prob [ i , j , 1: m]) }
  theta [ i ] ~ dnorm (0.0 , 1.0)
#cumulatives probabilities
  for ( j in 1: I){
    for ( k in 1:m-1 ){Pmas[i,j,k] <- 1/(1+exp(-a[j]*theta[i]+b[j,k] )) }
    Pmas[i,j,m]<- 0}
#probabilities of categories
  for ( j in 1: I){
    for ( k in 1: m-1){ p[i , j , k ] <-Pmas[i,j,k]-Pmas[i,j,k+1] }
    prob[i,j,1]<-1-Pmas[i,j,1]
    for ( k in 2: m) { prob[i,j,k] <- p[i,j,k-1] }
  }
#priors to item parameters
  for ( j in 1: I){ a[j ] ~ dnorm (1, 2 )I(0 , ) }
for ( j in 1: I) {
  for ( k in 1:( m -1) ){ b.star [ j , k ] ~ dnorm (0 , 0.5 )
    b[j,k]<- ranked(b.star [ j , 1:( m -1) ] , k)
    bc[j,k]<-b[j,k]/a[j] }
  }}
}
```

Figura 21: Código BUGS para o modelo de Resposta graduada.

Aplicação do modelo de resposta graduada a uma escala de atitudes em Marketing

Para ilustração consideramos a escala lickert de 6 opções (1 discordo totalmente até 6 concordo totalmente) que mede as atitudes frente ao mobile banking de Puschel, Mazzon e Hernandez (2010) a qual consta de 4 itens. Os dados correspondem a 333 não usuários de mobile banking, mas que possuem ao menos uma conta bancária e serviço de celular, obtidos de uma amostra intencional.

Atitudes frente ao mobile banking

1. Usar o <i>mobile banking</i> para fazer transferências bancárias é uma boa ideia
2. Utilizar o <i>mobile banking</i> é uma ideia inteligente
3. Gosto da ideia de utilizar o <i>mobile banking</i>
4. É/Seria muito divertido utilizar o <i>mobile banking</i>

Os testes foram aplicados em um processo de duas etapas: primeiro responderam se concordam ou não com o item da escala e, em seguida, indicaram o seu grau de concordância ou discordância.

Tabela 4. Itens, estatísticas de itens e estimativas dos parâmetros dos itens usado o modelo GRM da TRI.

Itens	Média dos escores	Desvio Padrão dos scores	b1	b2	b3	b4	b5	a
1. Usar o <i>mobile banking</i> para fazer transferências bancárias é uma boa ideia	4.1	1.5	-1.625	-1.024	-0.553	0.188	1.084	2.746
2. Utilizar o <i>mobile banking</i> é uma ideia inteligente	4.7	1.3	-2.154	-1.746	-1.065	-0.269	0.671	3.042
3. Gosto da ideia de utilizar o <i>mobile banking</i>	4.1	1.5	-1.411	-0.944	-0.565	0.210	0.955	4.251
4. É/Seria muito divertido utilizar o <i>mobile banking</i>	3.6	1.6	-1.370	-0.860	-0.163	0.708	1.561	2.002

b_j = parâmetro de concordância no nível j (dificuldade); quando negativo expressa a facilidade em concordar nesse nível. Quando positivo, expressa o umbral até onde é possível concordar e o valor expressa o grau da concordância.

a = parâmetro de discriminação, quanto mais alto indica que o item é apropriado para identificar as atitudes.

Considerando o modelo GRM da TRI foi obtida a dificuldade e discriminação dos itens e a função de informação dos itens e da escala, a qual proporciona a quantidade de informação em cada ponto do traço latente como mostrado na Tabela 4.

Como as discriminações (a) são maiores de 2 em todos os itens confirmamos a qualidade do teste, ainda observamos que as dificuldades de passo entre as categorias de resposta são negativas até o terceiro passo indicando que foi relativamente fácil atingir o acordo de nível 4.

O item 4 (É/Seria muito divertido utilizar o *mobile banking*) é o menos discriminativo em relação a atitude frente ao *móbile banking*. Além disso, apresenta uma maior dificuldade para concordar no nível 5.

Considerando o sinal dos valores dos parâmetros b identificou-se as “dificuldades” de passo entre as categorias de resposta. Assim, no caso do item 2 (Utilizar o *mobile banking* é uma ideia inteligente) é mais fácil atingir a última categoria (concordância total) do que nos outros itens. O contrário acontece no item 4.

Tarefa

1. Implemente o modelo de resposta graduada de Samejima para os dados do teste de Atitudes frente ao Mobile Banking usando WinBUGS.
2. Proponha outras ligações alternativas para as probabilidades acumuladas a direita. Por exemplo, usando a ligação Probit, Cloglog e Power Logito.
3. Considerando o critério DIC, qual modelo é o mais apropriado para os dados?
4. Quais itens apresentam um parâmetro de penalidade significativo? Interprete estes resultados em termos da aplicação.
5. Recentemente foi proposta uma ligação usando a acumulada da distribuição Power Normal (Bazán, Romeo e Rodrigues, 2014). Desenvolva um código para um modelo TRI de resposta graduada considerando esta nova ligação.

Adicional

A modo de comparação, replique os resultados usando os seguintes pacotes para a inferência clássica.

- Pacotes comerciais: IRTPRO, PARSCALE, MULTILOG

Maiores detalhes em <http://www.ssicentral.com/irt/>

- Pacotes no R: ltm e mirt

Maiores detalhes em <http://cran.r-project.org/web/views/Psychometrics.html>

- Para outros modelos, ver os pacotes disponíveis em:

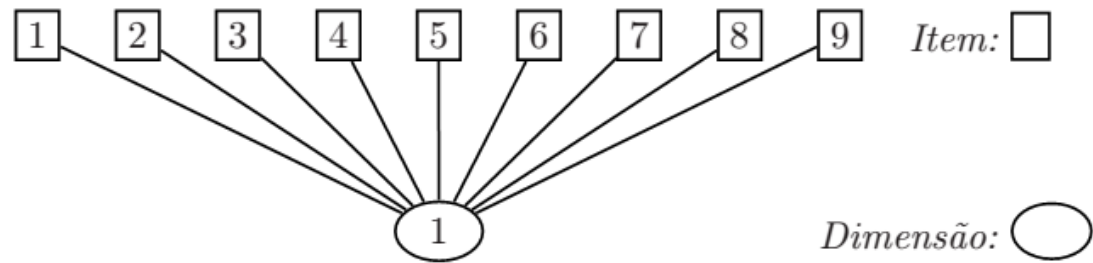
http://en.wikipedia.org/wiki/Psychometric_software

Observação. Na aproximação clássica a estimação é dividida em duas etapas. Na primeira, chamada de calibração, são estimados os parâmetros dos itens e na segunda, chamada de estimação, são estimadas as proficiências.

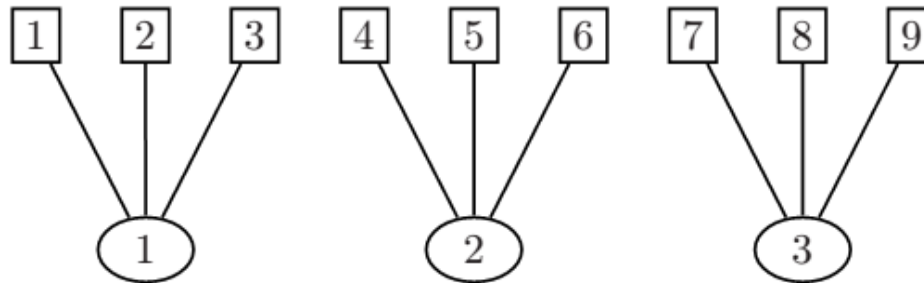
5.3 CONSIDERANDO VÁRIOS TRAÇOS LATENTES MEDIDOS

- Comumente, as avaliações envolvem mais de um traço latente sendo medido, seja em formatos de resposta dicotômica ou politômica.
- Um item é dito multidimensional quando ele avalia mais de uma proficiência. Caso contrário, o item é unidimensional.
- Devemos distinguir entre uma avaliação multivariada e uma avaliação multidimensional. Na avaliação multivariada vários traços latentes são medidos em diferentes provas com diferentes itens e conseqüentemente os itens avaliam um traço só. Na avaliação multidimensional, diferentes traços latentes são os responsáveis da resposta observada no item.

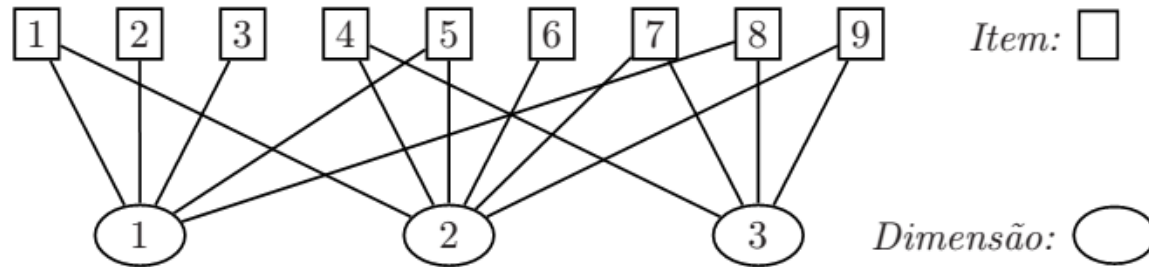
Modelo Unidimensional:



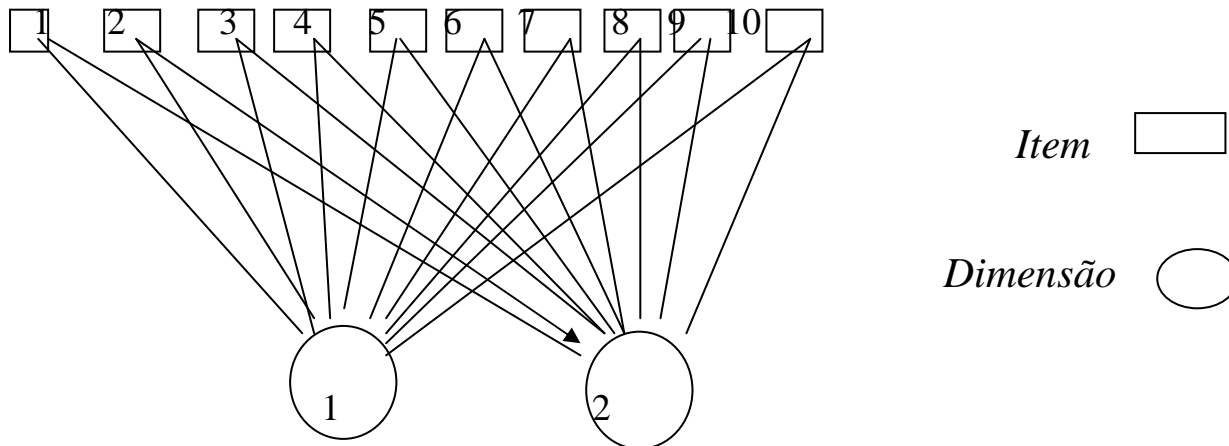
Modelo multivariado:



Modelo Multidimensional tipo A (itens multidimensionais e unidimensionais)



Modelo Multidimensional tipo B (itens multidimensionais)



- O ENEM é multivariado pois são assumidas diferentes provas independentes, mas ao mesmo tempo cada uma delas pode ser considerada provas multidimensionais desde que várias habilidades podem ser as responsáveis das respostas aos itens. Por exemplo, na prova de Matemática, exige mais de uma proficiência do aluno para que ele responda corretamente a um item (multidimensional). Por exemplo, habilidade matemática e habilidade de Compreensão e Leitura de textos.
- Modelos multidimensionais são interessantes mas apresentam problemas de identificabilidade. Precisa-se informação para definir o número de dimensões além de alocar os itens nas diferentes dimensões.
- A avaliação da dimensionalidade de um teste é por se mesmo um tema de alto interesse na psicometria. Diversas metodologias continuam a serem propostas para isto.

- Os modelos multidimensionais da TRI podem ser separados em duas classes: os compensatórios e os não compensatórios. Um modelo é dito compensatório quando a mesma probabilidade de acerto do item é mantida quando a diminuição do valor de um traço latente é compensada pelo aumento no valor de outro traço latente.
- Em modelos não compensatórios um examinado necessita de uma certa quantidade de habilidade em cada dimensão avaliada para que atinge uma probabilidade de sucesso de um item.
- Muitos modelos compensatórios e não compensatórios foram e continuam sendo propostos.
- A ênfase atual é na inferência clássica, existindo poucos trabalhos na inferência bayesiana.
- É uma área com pesquisa tanto prática quanto teórica mas existem poucos programas para implementar estes modelos.

Algumas dissertações que podemos citar são:

- Matos, G. S. (2008). *Modelos multidimensionais da TRI com distribuições assimétricas para os traços latentes*. Tese de Doutorado, IME-USP, São Paulo-SP.
- Nojosa, R. T. (2001). *Modelos multidimensionais para a Teoria da resposta ao item*. Dissertação de Mestrado, DE-UFPE, Recife-PE.
- Fragoso T. d. M. (2010). *Modelos multidimensionais da Teoria da Resposta ao Item*. Dissertação de mestrado, ICMC-USP, São Paulo-SP.
- Nojosa, R. T. (2011). *Inferência Bayesiana em Modelos Multidimensionais de Resposta ao Item*. Doutorado em Estatística. USP.

- Fragoso, T. M. e Cúri, M. (2013). Improving psychometric assessment of the Beck Depression Inventory using Multidimensional Item Response Theory. *Biometrical Journal*, **55**, 527-540.

Modelo compensatório multidimensional com itens multidimensionais dicotômicos

Considere n indivíduos avaliados em um teste de I itens. Um modelo TRI de dois parâmetros binário multidimensional compensatório é um sistema no qual para cada indivíduo i existe um modelo de variável latente monótono multidimensional (Y, Θ_i) , definido pelas seguintes expressões:

$$Y_{ij}|u_i, \eta_j \sim \text{Bernoulli}(p_{ij})$$
$$p_{ij} = P(Y_{ij} = 1 | \boldsymbol{\theta}_i, \eta_j) = F(m_{ij})$$
$$m_{ij} = \mathbf{a}_j^T \boldsymbol{\theta}_i - b_j = \sum_{k=1}^d a_{jk}(\theta_{ki} - d_{jk}),$$
$$i = 1, \dots, n, \quad j = 1, \dots, I$$

em que:

- Y_{ij} é a variável manifesta a qual modela a resposta binária de cada indivíduo responder corretamente ou não ao item j .

- $\eta_j = (a_{j1}, \dots, a_{jd}, b_j)$ o vetor relacionado ao parâmetro de discriminação do item j e o parâmetro relacionado a dificuldade do item j respectivamente.
- θ_i é o vetor ou traço latente multidimensional Θ_i do indivíduo i , p_{ij} é a probabilidade condicional dado $\Theta_i = \theta_i$.
- F é uma função de distribuição acumulada. Quando utilizamos a distribuição Logística temos o modelo logístico multidimensional de dois parâmetros. É possível estender o modelo para incluir um parâmetro de adivinação e ter o modelo multidimensional de três parâmetros.
- m_{ij} é o preditor linear latente relacionada com o traço latente multidimensional do indivíduo i que responde ao item j .
- Quando desenhamos a probabilidade como função dos traços latentes avaliados obtemos a Superfície característica para um item (SCI).

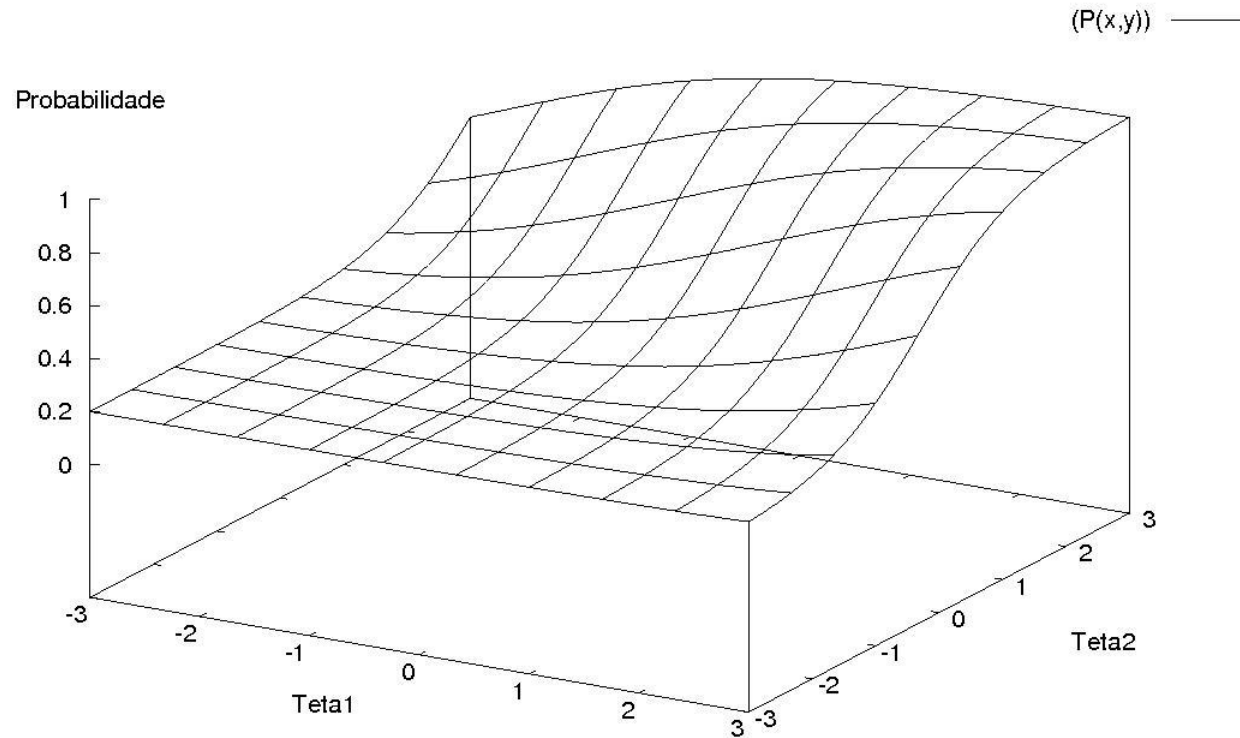


Figura 22: Superfície característica de um Item de um modelo logístico de três parâmetros com duas dimensões

São propostos alguns índices resumos para itens multidimensionais:

$MDIFICM_j = \frac{b_j}{MDISCM_j}$ que é o parâmetro de dificuldade resumo.

$MDISCM_j = (\sum_{k=1}^d a_{jk}^2)^{1/2}$ que é o correspondente parâmetro de discriminação resumo.

Fragoso e Cúri (2013) propõem usar: $a_j^* = \frac{a_j}{MDISCM_j}$ o qual permite interpretar os parâmetros de discriminação como um número entre 0 e 1 análogo as cargas fatoriais de um Análise fatorial.

Modelos politômicos multidimensionais são possíveis. Por exemplo os modelos multidimensionais de crédito parcial generalizado e de resposta graduada.

Maiores detalhes em:

Reckase MD (2009). *Multidimensional Item Response Theory*. Springer-Verlag, New York.

Os programas em que são implementados alguns modelos TRI multidimensionais:

- Martin A. D., Quinn, K. M. e Park, J. H. (2011). MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software*, **42** (9), 1-21. URL <http://www.jstatsoft.org/v42/i09/>
- Sheng, Y. (2010). Bayesian Estimation of MIRT Models with General and Specific Latent Traits in MATLAB . *Journal of Statistical Software*, **34** (3), 1-27. URL <http://www.jstatsoft.org/v34/i03/>
- Yao, L. (2003). *BMIRT: Bayesian multivariate item response theory. [Computer software]*. Monterey, CA: Defense Manpower Data Center.
- Simon, J. (2012). *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, Stanford University*. Department of Political Science, Stanford University. Stanford, California. R package version 1.04.4. URL <http://pscl.stanford.edu/>

```
#Modelo de Bidimensional Logístico de dois parâmetros
model{
  for (i in 1 : n) {
    for (j in 1 : I) {
      logit(p[i,j]) <- a1[j]*theta1[ i ]+a2[j]*theta2[ i ] - b[j ]
      y[i, j ] ~ dbern( p[i, j] )      }
    }
#priors para os parâmetros de item e medidas derivada
  for (j in 1:I) {
    b[j] ~ dnorm(0, 1)
    a1[j] ~ dlnorm(1, 2)
    a2[j] ~ dlnorm(1, 2)
    mdific[j] <- b[j]/mdisc[j]
    a1quad[j] <- pow(a1[j],2); a2quad[j] <- pow(a2[j],2)
    mdisc[j] <- pow(a1quad[j] + a2quad[j] ,1/2)
    a1star[j] <- a1[j]/mdisc[j]; a2star[j] <- a2[j]/mdisc[j]
  }
#priors para os tracos latentes
  for (i in 1:n) {
    theta1[i] ~ dnorm(0,1)
    theta2[i] ~ dnorm(0,1)      }
}
```

Figura 23: Código BUGS para um modelo logístico de dois parâmetros com duas dimensões

Aplicação da TRI a um inventário de depressão em Psiquiatria

Para ilustração consideramos um inventário de depressão de Beck (BDI) o qual consta de 21 itens. Os dados correspondem a 1111 universitários e foram cedidos pelo professor Chei-Tung Teng, do Departamento de Psiquiatria do Hospital das Clínicas (HCFMUSP, Hospital das Clínicas da Faculdade de Medicina da USP) e pela professora Mariana Cúri.

As questões do BDI abrangem diversos aspectos da vida do indivíduo possivelmente afetados pela presença de depressão, as quais são respondidas em função de uma escala de intensidade de 0 (Baixa intensidade) a 3 (Alta intensidade) referente ao quanto a depressão afeta o indivíduo na atividade em questão.

Exemplos de itens do inventário de depressão de Beck

<http://www.hcnet.usp.br/ipq/revista/vol26/vol25/n5/depre255b.htm>

Sentimento de fracasso

item 3. 0 Não me sinto um fracasso.

- 1 Acho que fracassei mais do que uma pessoa comum.
- 2 Quando olho para trás, na minha vida, tudo o que posso ver é um monte de fracassos.
- 3 Acho que, como pessoa, sou um completo fracasso.

Fatigabilidade

item 17. 0 Não fico mais cansado que de hábito.

- 1 Fico cansado com mais facilidade do que costumava.
- 2 Sinto-me cansado ao fazer quase qualquer coisa.
- 3 Estou cansado demais para fazer qualquer coisa.

Os dados foram dicotomizados adotando-se um valor de 0 (fracasso) para respostas iguais a 0 ou faltantes e 1 (sucesso) para respostas de valor 1, 2 ou 3.

Para uma análise diferente usando o modelo de resposta graduada unidimensional ver:

- Castro, S. M. J., Trentini, C. e Riboldi, J. (2010). Teoria da resposta ao item aplicada ao inventário de depressão de Beck. *Revista Brasileira de Epidemiologia*, **13**, 487–501.

Para uma análise usando o modelo Logístico de três parâmetros ver:

- Cúri M. (2006). *Análise de questionários com itens constrangedores* Tese de doutorado. São Paulo: Instituto de Matemática e Estatística da USP.

Análise clássica:

- Gorenstein, C., Andrade, L., Vieira Filho, A. H. G., Tung, C. T. e Artes, R. (1999). Psychometric properties of the Portuguese version of the beck depression inventory on brazilian college students. *Journal of Clinical Psychology*, **55**, 553–562.

SINAPE 2014. Minicurso Enfoque Bayesiano nos modelos de Teoria da Resposta ao Item.

Item	Descrição	a1	a2	b	mdific	mdisc	a1*	a2*
BDI1	Tristeza	0.84	1.15	-0.75	-0.53	1.43	0.59	0.8
BDI2	Preocupação pelo Futuro	1.16	1.09	0.72	0.45	1.6	0.73	0.68
BDI3	Fracasso	1	2.63	1.55	0.55	2.82	0.36	0.93
BDI4	Ausência do Prazer	1.68	1.06	-0.09	-0.04	1.99	0.84	0.53
BDI5	Culpa	0.69	2.6	1.46	0.54	2.69	0.26	0.96
BDI6	Preocupação pela Punição	0.44	1.47	0.9	0.59	1.54	0.29	0.95
BDI7	Decepção	0.69	3.01	-0.49	-0.16	3.1	0.23	0.97
BDI8	Inferioridade	0.43	1.52	-0.84	-0.54	1.58	0.27	0.96
BDI9	Suicídio	1.22	1.22	2.51	1.46	1.73	0.7	0.7
BDI10	Choro	1.02	0.83	1.14	0.87	1.32	0.77	0.63
BDI11	Irritação	1.13	0.59	-0.06	-0.05	1.28	0.88	0.46
BDI12	Perda do Interesse	1.37	0.68	0.74	0.49	1.53	0.89	0.45
BDI13	Dificuldade nas Decisões	1.26	1	0.63	0.39	1.61	0.78	0.62
BDI14	Aparência não atrativa	0.75	0.68	0.72	0.71	1.02	0.74	0.67
BDI15	Dificuldade para Trabalhar	1.32	1.15	0.16	0.09	1.75	0.75	0.66
BDI16	Falta de Sono	1.13	0.43	-0.01	-0.01	1.21	0.93	0.36
BDI17	Cansaço	1.35	0.59	-0.56	-0.39	1.47	0.91	0.4
BDI18	Ausência de Appetite	0.81	0.5	1.11	1.16	0.96	0.84	0.53
BDI19	Perda de Peso	0.69	0.29	2.35	3.19	0.75	0.91	0.39
BDI20	Preocupação pelo aspecto Físico	0.66	0.58	0.82	0.93	0.89	0.74	0.66
BDI21	Perdida de interesse no Sexo	0.84	0.44	1.29	1.38	0.95	0.88	0.46

Considerando a discriminação, encontramos que os sintomas 6 e 7 são menos discriminativos em uma dimensão, enquanto que os sintomas 11, 16, 17, 18, 19, 20 e 21 são menos discriminativos na segunda dimensão. No caso dos outros itens eles se associam a ambas dimensões.

Em geral, considerando a medida resumo de discriminação, os sintomas 3, 5 e 7 (fracasso, culpa e decepção) são os sintomas mais discriminativos.

Adicionalmente os sintomas 1, 4, 7, 9, 11, 16 e 17 são os sintomas inicialmente mais fáceis de identificar (menos severos) e pelo contrario os sintomas 9 e 19 (de suicídio e perda de peso) são mais difíceis de identificar (mais graves)

Considerando as medidas $a1^*$ e $a2^*$ podemos identificar os sintomas que caracterizam a cada dimensão. Notamos sintomas que estão em ambas dimensões mas também sintomas que são próprios de cada dimensão.

Maiores detalhes em:

Fragoso, T. M. e Cúri, M. (2013). Improving psychometric assessment of the Beck Depression Inventory using Multidimensional Item Response Theory. *Biometrical Journal*, **55**, 527-540.

Tarefa

1. Implemente o modelo multidimensional Logístico de três parâmetros para os dados do inventário de Beck usando WinBUGS.
2. Compare seus resultados com os obtidos usando outros pacotes usando a aproximação Bayesiana.
3. Proponha outras ligações alternativas. Por exemplo, usando a ligação Probit, Cloglog e Power Logito.
4. Considerando o critério DIC, qual modelo é o mais apropriado para os dados?
5. Quais itens apresentam um parâmetro de penalidade significativo? Interprete estes resultados em termos da aplicação.
6. Recentemente foi proposto uma ligação usando a acumulada da distribuição Power Normal (Bazán, Romeo e Rodrigues, 2014). Desenvolva um código para um modelo TRI multidimensional considerando esta nova ligação.

Adicional

A modo de comparação, replique os resultados usando os seguintes pacotes para a inferência clássica:

- Pacotes comerciais: IRTPRO, TESTFACT

Maiores detalhes em <http://www.ssicentral.com/irt/>

- Pacotes no R: ltm, TAM, mirt e mcIRT

Maiores detalhes em <http://cran.r-project.org/web/views/Psychometrics.html>

- Para outros modelos, ver os pacotes disponíveis em:

http://en.wikipedia.org/wiki/Psychometric_software

Observação. Na aproximação clássica a estimação é dividida em duas etapas. Na primeira, chamada de calibração, são estimados os parâmetros dos itens e na segunda, chamada de estimação, são estimadas as proficiências.

6. COMENTÁRIOS FINAIS

- Foram apresentados os principais modelos da TRI para itens dicotômicos e politômicos unidimensionais assim como o modelo dicotômico multidimensional.
- A ênfase foi a estimação Bayesiana usando MCMC a qual estima simultaneamente parâmetros de itens e proficiências.
- Foram apresentados três exemplos em diferentes contextos, uma prova de Matemática em estudantes de escolas de ensino fundamental, uma escala de atitudes frente ao móbile banking em não usuários do serviço, e um inventário para medir a depressão aplicado em estudantes universitários. A ideia foi mostrar a amplitude de usos dos modelos da TRI assim como as diferentes interpretações de seus parâmetros.
- Em todos os casos, códigos BUGS, bases de dados e códigos em R são disponibilizados junto ao site do autor: <http://www.icmc.usp.br/pessoas/jlbazan/>

- A interface entre R e códigos bugs poder ser realizadas usando R2WinBUGS, R2Openbugs, Rjags e rbugs. Nosso resultados foram obtidos com R2WinBUGS usando WinBUGS.
- Recomendamos a leitura do Manual do programa Bayesian Modeling.

Bazán. J. (2011). Bayesian Modeling User Manual. Software. Departamento de Ciências. PUCP. [Download manual pdf](#)

- Recomendamos alguns tutoriais, disponíveis em:
<http://www.icmc.usp.br/pessoas/jlbazan/software.html>
- Recomendamos fazer estudos de simulação na TRI.

Para gerar diferentes modelos de TRI, recomendamos desenvolver seus próprios códigos ou usar WinGen disponível em: <http://www.umass.edu/remf/software/simcata/wingen/homeF.html>

7. OUTRA REFERENCIAS CITADAS

- Albert, J. (2009). *Bayesian Computation with R*. Springer Verlag.
- Aranda-Ordaz (1981). On two families of transformations to additivity for binary response data. *Biometrika*, **68**, 357-363.
- Arnold B, Groeneveld R. (1995). Measuring skewness with respect to the mode. *The American Statistician*, **49**, 34–38.
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statistical*, **12**, 171-178.
- Basu, S. e Mukhopadhyay, S. (2000). *Binary response regression with normal scale Mixtures links, in Generalized Linear Models: A Bayesian Perspective*, Eds. D. K. Dey, S. K. Ghosh, and B. K: Mallick, New York: Marcel Dekker.
- Bazan, J. L., Romeo, J. S. e Rodrigues, J. (2014). Bayesian skew-probit regression for binary response data. *Brazilian Journal of Probability and Statistics*, In Press.
- Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In Lord, F.M. y Novick, M.R. (Eds) *Statistical Theories of Mental Test Scores*. New York: Addison-Wesley.
- Carlin, B. P. e Louis, T. A., (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall, CRC, London, Boca Raton, FL.

- Czado, C., and Santner, T. J. (1992). The effect of link misspecification on binary regression inference. *Journal of Statistical Planning and Inference*, **33**, 213-231.
- Czado, C. (1994). Bayesian inference of binary regression models with parametric link. *Journal of Statistical Planning and Inference*, **41**, 121-140.
- Collet, D. (2003). *Modelling binary data*. Chapman & Hall/CRC, Second Edition, Boca Raton, USA.
- Congdon, P. (2010). *Applied Bayesian Hierarchical Methods*, Chapman & Hall / CRC.
- Congdon, P. (2005). *Bayesian Models for Categorical Data*, Wiley.
- Chen, M. H., Dey, D. e Shao, Q-M. (2001). Bayesian analysis of binary data using Skewed logit models. *Calcutta Statistical Association Bulletin*, **51**, 201-202.
- Fischer, G. e Molenaar, I. (1995). *Rasch Models Foundations, recent development, and applications*. The Netherlands: Springer-Verlag.
- Fu, Z. H., Tao, J. e Shi, N. Z. (2009). Bayesian estimation in the multidimensional three-parameter logistic model. *Journal Journal of Statistical Computation and Simulation*, **79**, 819 - 835.
- Guerrero, V. M., and Johnson, R. A. (1982). Use of the Box-Cox transformation With binary response models. *Biometrika*, **69**, 309-314.
- Gilks, W. R. e Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics*, **41**: 337-348

- Haro-López, R. A., Mallick, B. K. e Smith, A. F. M. (2000). Binary regression using data adaptive robust link functions, in *Generalized Linear Models: A Bayesian Perspective*, eds. D.K. Dey, S.K. Ghosh, and B.K. Mallick, New York: Marcel Dekker.
- Hemker, B. T. (2001). *Reversibility revisited and other comparison of tree types of polytomous IRT models*. In: A. Boomsma, M. A. J. Van Duijn, & T. A. B. Snijders (Ed.). *Essays on item response theory* (pp. 277-96). New York: Springer-Verlag.
- Johnson, V. e Albert, J. (1999). *Ordinal Data Modeling*. New York: Springer-Verlag.
- Lord, F. M. (1952). *A theory of test scores*. New York: Psychometric Society.
- Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, **47**, 149-74.
- Martin, A. D. e Quinn, K. M. (2002). *MCMC pack*. scythe.wustl.edu/mcmcpack.html.
- Nagler, J. (1994). Scobit: An alternative estimator to Logit and Probit. *American Journal of Political Science*, **38** (1), 230-255.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, **16**, 159-77.
- Ntzoufras, I. (2009). *Bayesian Modeling Using WinBugs*. Wiley Series in Computational Statistics, Hoboken, USA.
- Patz, R. J. e Junker, B. W., 1999, A straightforward approach to Markov chain Monte Carlo methods for item response models, *Journal of Educational and Behavioral Statistics*, **24**, 146–178.

- Prentice, R. L. (1976). A Generalization of the probit and logit methods for Dose response curves. *Biometrika*, **32** (4), 761-768.
- Puschel, J., Mazzon, J. A. e Hernandez, J. M. (2010). Mobile banking: proposition of an integrated adoption intention framework. *International Journal of Bank Marketing*, **28** (5), 389-409.
- Rupp, A., Dey, D. K. e Zumbo, B. (2004). To Bayes or Not to Bayes, from Whether to When: Applications of Bayesian Methodology To Item Response Modeling, Structural Equations Modeling, **11**, 424-451.
- Sahu, S. K. (2002). Bayesian estimation and model choice in item response models. *Journal Statistical Computing Simulation*, **72**, 217-232.
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, **17**.
- Sheng, Y., 2008, Markov chain Monte Carlo estimation of normal ogive IRT models in MATLAB, *Journal of Statistical Software*, **25** (8), 1–15.
- Sinharay, S. (2004). Experiences with Markov Chain Monte Carlo convergence assessment in two psychometric examples. *Journal of Educational and Behavioral Statistics*, **29** (4), 461-488
- Spiegelhalter, D. J., Thomas, A., Best, N. G. e Gilks, W. R. (1996). *BUGS 0.5 examples (Vol. 1 Version i)*. Cambridge, UK: University of Cambridge.

- Spiegelhalter, D. J., Thomas, A., Best, N. G., Lunn, D (2007) *OpenBUGS User Manual version 3.0.2*. MRC Biostatistics Unit, Cambridge.
- Stukel, T. (1988). Generalized logistic models. *Journal of the American Statistical Association*, **83**, 426-431
- Tutz, G. (1990). Sequential item response models with an ordered response. *British Journal of Mathematical and Statistical Psychology*, **43** (1), 39-55.
- Van der Ark, L. A. (1999). A reference card for the relationships between IRT models for ordered polytomous items and some relevant properties. *The Netherlands: Tilburg University*. Working Paper No. 99.10.02.

Agradecimentos

A todos meus coautores em diferentes artigos apresentados neste trabalho. Professores Heleno Bolfarine, Márcia D. Branco, Luis Valdivieso, José A. Mazzon e Martin Hernani.

A meus colegas no ICMC, especialmente os professores Adriano K. Suzuki e Mariana Cúri por sua disposição e ajuda na revisão do documento e disponibilidade no uso do exemplo Multidimensional.

A Diretoria da ABE, em especial aos professores Francisco Louzada Neto e Vera L. D. Tomazella.