Curvas de crescimento com estrutura de dependência em modelos de resposta ao item

Adriana Moraes de Carvalho e Héliton Ribeiro Tavares

Programa de Pós-Graduação em Matemática e Estatística - PPGME, ICEN, UFPA,
Campus do Guamá, 66075110, Belém, PA
E-mail: adri_para@yahoo.com.br e heliton@ufpa.br

Introdução

A avaliação educacional é um ponto muito importante no ensino, e vem sendo cada vez mais discutida por profissionais que buscam receber mais informações dos alunos do que simplesmente o número de itens (*questões*) respondidos corretamente em um teste. A Teoria da Resposta ao Item (TRI) entra nesse contexto com uma abordagem inovadora, propor modelos que relacionam os traços latentes (*habilidades*) do aluno, ou seja, características do aluno não observadas diretamente, tal como capacidade de leitura, por exemplo, com os itens apresentados à ele em um teste.Isso porque essa teoria modela a relação entre a probabilidade de um aluno responder corretamente a um determinado item e sua habilidade na área de conhecimento avaliada. Esta relação é geralmente expressa de tal forma que quanto maior a habilidade, maior a probabilidade de acerto ao item (Andrade, Tavares & Valle (2000)).

A Teoria da Resposta ao Item (TRI) vem tornando-se a técnica predominante no contexto de avaliações educacionais no Brasil. Foi usada pela primeira vez em 1995 na análise de dados do Sistema Nacional de Ensino Básico - SAEB. A partir dos resultados obtidos no SAEB, outras avaliações em larga escala também foram planejadas e implementadas de modo a serem analisadas através da TRI. Dentre as quais podemos destacar: Avaliação Nacional do Rendimento Escolar (Prova Brasil), Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (SARESP), Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja), Programa Internacional de Avaliação de Alunos (Pisa), Exame Nacional do Ensino Médio (Enem).

Objetivo

Neste trabalho são propostos modelos para representar o comportamento das habilidades médias de um mesmo grupo de indivíduos acompanhados ao longo do tempo (estudo longitudinal), através da estimação dos parâmetros de Curvas de Crescimento, tais como: Linear, Quadrática, Logística e Gompertz. Como a estrutura longitudinal induz uma dependência entre os resultados nas várias condições de avaliação, propomos estruturas de covariância para modelar tal dependência. Consideramos conhecidos os parâmetros dos itens envolvidos no estudo.

A aplicação é feita a dados resultantes da Pesquisa "Avaliação do desempenho: Fatores Associados"do projeto Fundescola, aplicada a alunos de 4a. até a 8a. série do Ensino Fundamental, no período de 1999 a 2003.

Metodologia

A proposta da TRI é modelar a probabilidade de acerto ao item i. A Função de Resposta ao Item (FRI), do modelos considerado, é dada por:

$$P(U_{jit} = 1|\theta_{jt}, \boldsymbol{\zeta}_i) = c_i + (1 - c_i)\frac{1}{1 + e^{-Da_i(\theta_{jt} - b_i)}},$$

onde, θ_{jt} é a habilidade do j-ésimo respondente ao teste t, $\boldsymbol{\zeta}_i$ é o conjunto de características associadas ao item i, com $\boldsymbol{\zeta}_i = (a_i, b_i, c_i)'$, b_i é o parâmetro de dificuldade, a_i é o parâmetro de discriminação, c_i é o parâmetro de discriminação, e D é um fator de escala, constante e igual a 1.

Para a situação em que o mesmo grupo de indivíduos é acompanhado ao longo do tempo, espera-se uma estrutura de dependência entre as habilidades, uma vez que os mesmos indivíduos são submetidos a tais testes. De forma geral, consideramos que o vetor de habilidades dos respondentes nos T testes, $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_T)'$, tem distribuição contínua multivariada com vetor de parâmetros $\boldsymbol{\eta}$ de componentes finitas, e função densidade de probabilidade $g(\boldsymbol{\theta}|\boldsymbol{\eta})$.

No caso Normal T-variado temos que $\boldsymbol{\eta} = (\boldsymbol{\mu}, \boldsymbol{\Sigma})$, onde $\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_T)'$ é o vetor de médias e $\boldsymbol{\Sigma} = (\sigma_{tj})_{t,j}$, é a matriz de covariâncias.

A equação de estimação para $\boldsymbol{\alpha}$ é

$$\boldsymbol{\alpha} \quad : \quad \sum_{j=1}^{N} \int_{I\!R} (\boldsymbol{\Sigma}^{-1})(\boldsymbol{\theta} - \boldsymbol{\mu}) \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\alpha}}\right) g_j^*(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathbf{0}.$$

onde,

$$g_j^*(\boldsymbol{\theta}) \quad = \quad \frac{P(\boldsymbol{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta})g(\boldsymbol{\theta}|\boldsymbol{\eta})}{P(\boldsymbol{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta})}.$$

Estruturas de Covariância

Quando o interesse está em acompanhar o desempenho de um mesmo grupo de alunos ao longo do tempo, como é o caso deste trabalho, tem-se a necessidade de incorporar possíveis estruturas de covariância entre as habilidades desses alunos. Algumas dessas estruturas apresentadas na literatura serão mostradas neste trabalho.

Matriz de Covariância Diagonal

$$\mathbf{\Sigma}^{-1} = \mathbf{diag}\{1/\sigma_1^2, 1/\sigma_2^2, \ldots, 1/\sigma_T^2\}.$$

Matriz de Covariância Uniforme

$$\mathbf{\Sigma}^{-1} = \frac{1}{(1-\rho)\sigma^2}\left(\boldsymbol{I}_T - \frac{\rho}{1+\rho(T-1)}\boldsymbol{J}_T\right).$$

Matriz de Covariância Bandas

$$\mathbf{\Sigma}^{-1} = (v_{ij})_{i,j \leq T}.$$

Com,

$$v_{ij} = \frac{(1-\beta^{2(T-m_{ij}+1)})(\beta^{j+i+1} - \beta^{|j-i|+1})}{\sigma^2\rho(1-\beta^2)(1-\beta^{2(T+1)})},$$

Onde, $m_{ij} = max\{i,j\}, i,j = 1,\ldots,T$, e $\beta = \frac{1}{2\rho}(\sqrt{1-4\rho^2}-1)$.

Matriz de Covariância AR(1)

$$\mathbf{\Sigma}^{-1} = (\sigma^{-2}v_{ij}).$$

Com,

$$v_{ij} = \begin{cases} 1/\alpha & \text{se } i = j \text{ e } i \in \{1, T\} \\ \beta/\alpha & \text{se } i = j \text{ e } i \in \{2,\ldots,T-1\} \\ -\rho/\alpha & \text{se } j \in \{i-1, i+1\},\ i \in \{1,\ldots,T\} \text{ e } j \in \{2,\ldots,T-1\} \\ 0 & \text{c.c.} \end{cases}$$

Curvas de Crescimento

Para obter as equações de estimação dos respectivos parâmetros, temos que encontrar as derivadas em relação a cada parâmetro, para as seguintes curvas

Curva de Crescimento Linear

$$\mu_t = \alpha_0 + \alpha_1 s_t.$$

Curva de Crescimento Quadrática

$$\mu_t = \alpha_0 + \alpha_1 s_t + \alpha_2 s_t^2.$$

Curva de Crescimento Logística

$$\mu_t = \frac{\alpha_0}{1 + e^{-\alpha_1(s_t - \alpha_2)}}.$$

Curva de Crescimento Gompertz

$$\mu_t \quad = \quad \alpha_0 + \alpha_1 e^{-e^{-\alpha_2(s_t - \alpha_3)}}.$$

Resultados

Usamos o software Ox (Doornik, 2000) para construir uma macro de estimação denominada **LongIRT** e com esta obter os resultados. Este software pode ser obtido em www.ufpa.br/heliton/longirt.

Foram selecionamos apenas os alunos que compareceram a todas os testes, o que corresponde a 1.987. Foram considerados apenas os itens da disciplina de Matemática.

Os momentos de avaliação, como já informado, foram Abril/1999, Novembro/1999, Novembro/2000, Novembro/2001, Novembro/2002 e Novembro/2003. Como os tempos entre as avaliações não são regulares, resolveu-se adotar como instantes de avaliação os tempos 4, 11, 23, 35, 47 e 59.

Dos 40 itens aplicados em cada teste foram usados apenas 34 itens do Teste 1, 38 do Teste 2, 36 do teste 3, 34 do teste 4, 40 do Teste 5 e 34 do Teste 6, o que totalizou 167 itens.

Nas Tabelas 1 a 4 estão mostradas as estimativas das médias e dos parâmetros das curvas de crescimento, de acordo com a estrutura de covariância adotada. Cabe informar que as estimativas das proficiências médias quando realizadas as análises com a base de dados completa, ou seja, com todos os alunos envolvidos no estudo, utilizando-se o software BilogMG (Mislevy & Bock, 1990), foram, respectivamente, 0, 0.25, 0.74, 1.31, 1.54 e 1.66, onde vemos que o baseline foi a população referência. Ao considerar apenas o conjunto de 1987 considerados no presente trabalho, as proficiências médias foram 0.37, 0.53, 0.95, 1.37, 1.46 e 1.65. Estes últimos resultados foram obtidos no próprio LongIRT considerando a estrutura diagonal, sem curvas de crescimento, fixados os parâmetros dos itens.

Tabela 1: Estimativas dos Parâmetros - Dados reais - Matriz Diagonal

	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	α_0	α_1	α_2	α_3
Linear	0.3598	0.6454	0.9310	1.2166	1.5022	1.7878	0.0742	0.2856		
Quadrática	0.2651	0.6525	0.9876	1.2702	1.5004	1.6782	-0.1748	0.4661	-0.0262	
Logística	0.3420	0.5438	0.9756	1.3336	1.5224	1.5990	1.6411	0.0904	18.7673	
Gompertz	0.3680	0.5277	0.9726	1.3401	1.5210	1.5940	0.3637	1.2734	1.0218	2.7023

Tabela 2: Estimativas dos Parâmetros - Dados reais - Matriz Uniforme

	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	α_0	α_1	α_2	α_3
Linear	0.3549	0.6387	0.9224	1.2062	1.4900	1.7737	0.0711	0.2838		
Quadrática	0.3284	0.6439	0.9431	1.2257	1.4919	1.7417	-0.0037	0.3403	-0.0082	
Logística	0.3948	0.5702	0.9395	1.2933	1.5380	1.6716	1.7851	0.0718	21.5344	
Gompertz	0.3752	0.5663	0.9323	1.2867	1.5281	1.6660	0.3462	1.4631	0.7278	2.8778

Tabela 3: Estimativas dos Parâmetros - Dados reais - Matriz Bandas

	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	α_0	α_1	α_2	α_3
Linear	-0.0124	0.3075	0.6274	0.9473	1.2672	1.5871	-0.3323	0.3199		
Quadrática	-0.0772	0.3205	0.6793	0.9992	1.2802	1.5223	-0.5138	0.4561	-0.0195	
Logística	0.0417	0.0995	0.3813	0.9418	1.3627	1.5039	1.5466	0.1300	31.5933	
Gompertz	0.0184	0.2183	0.6593	1.0803	1.3469	1.4863	0.0000	1.6092	0.8061	2.8586

Tabela 4: Estimativas dos Parâmetros - Dados reais - Matriz AR(1)

	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	α_0	α_1	α_2	α_3
Linear	0.3726	0.6521	0.9316	1.2111	1.4907	1.7702	0.0930	0.2795		
Quadrática	0.3599	0.6607	0.9504	1.2291	1.4968	1.7534	0.0481	0.3173	-0.0055	
Logística	0.4180	0.5947	0.9605	1.3092	1.5534	1.6899	1.8115	0.0697	21.2647	
Gompertz	0.3966	0.5872	0.9522	1.3059	1.5468	1.6845	0.3678	1.4597	0.7278	2.8784

Para identificarmos, dentre estas duas estruturas de covariância, qual curva de crescimento melhor se ajusta aos dados, utilizamos o critério AIC (*Akaike's Infromation Criterion*), obtido por $AIC = -2\log L(\boldsymbol{\zeta}, \widehat{\boldsymbol{\eta}}) + 2n_p$, onde $L(\boldsymbol{\zeta}, \widehat{\boldsymbol{\eta}})$ representa o valor da verossimilhança marginal, e n_p o número de parâmetros envolvidos no modelo (Jones, 1993). A estrutura de covariância, em conjunto com a curva de crescimento, que gerar o menor valor para AIC comporá o modelo escolhido.

Com base no valores dos AIC's foi observado que a estrutura AR(1) foi a que melhor se ajustou aos dados, com valores AIC predominantemente menores. E para esta, verifica-se que a curva de crescimento logística foi claramente a que melhor modela o comportamento das proficiências médias do grupo ao longo dos seis instantes de avaliação, com estimativas dos parâmetros da curva de crescimento dadas por $\widehat{\alpha}_0 = 1.8115$, $\widehat{\alpha}_1 = 0.0697$ e $\widehat{\alpha}_2 = 21.2647$, e da estrutura de covariância $\widehat{\rho} = 0.9000$ e $\widehat{\sigma}^2 = 1.1690$. As estimativas médias geradas com esta curva de crescimento para os tempos adotados são dadas por $\widehat{\mu}_4 = 0.4180$, $\widehat{\mu}_{11} = 0.5947$, $\widehat{\mu}_{23} = 0.9605$, $\widehat{\mu}_{35} = 1.3092$, $\widehat{\mu}_{47} = 1.5534$ e $\widehat{\mu}_{59} = 1.6899$.

Referências

[1] Andrade, D.F., Tavares, H.R., Valle, R.C. (2000). *Teoria da Resposta ao Item: Conceitos e Aplicações.* Associação Brasileira de Estatística: São Paulo.

[2] Doornik, J.A. (2000). *Object-Oriented Matrix Programming using Ox 2.0.* London: Timberlake Consultants Ltd and Oxford: www.nuff.ox.ac.uk/Users/Doornik.

[3] Jones, R. H. (1993). *Longitudinal Data with Serial Correlation: A State-Space Approach.* Chapman and Hall: London.

[4] Mislevy, R. J. and Bock, R. D. (1990). *BILOG 3 : Item Analysis and Test Scoring with Binary Logistic Models.* Chicago : Scientific Software, Inc.