



UNIVERSIDADE FEDERAL DO PARÁ

INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA

CURVAS DE CRESCIMENTO COM ESTRUTURA DE DEPENDÊNCIA
EM MODELOS DE RESPOSTA AO ITEM

Adriana Moraes de Carvalho

Orientador: Prof. Dr. Héilton Ribeiro Tavares

Belém
2009

Adriana Moraes de Carvalho

**CURVAS DE CRESCIMENTO COM ESTRUTURA DE DEPENDÊNCIA
EM MODELOS DE RESPOSTA AO ITEM**

Dissertação de Conclusão de Curso apresentado ao Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará para a obtenção do Título de Mestre em Estatística.

Área de Concentração: **Teoria da Resposta ao Item**

Orientador: **Prof. Dr. Héilton Ribeiro Tavares**

**Belém
2009**

Adriana Moraes de Carvalho

**CURVAS DE CRESCIMENTO COM ESTRUTURA DE DEPENDÊNCIA
EM MODELOS DE RESPOSTA AO ITEM**

Dissertação apresentada para a obtenção do Título de Mestre em Estatística no Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará.

Belém, 21 de Agosto de 2009

Prof. Dr. Mauro de Lima Santos
Coordenador do Programa de Pós-Graduação em Matemática e Estatística

Banca Examinadora

Prof. Dr. Héilton Ribeiro Tavares
Universidade Federal do Pará
Inep/MEC
Orientador

Prof. Dr. Dalton Francisco de Andrade
Universidade Federal de Santa Catarina
Examinador

Prof. Dr. Caio Lucidius Naberezny Azevedo
Universidade Estadual de Campinas
Examinador

*Aos meus Pais,
Claudio e Andrelina.*

Agradecimentos

A Deus por ter me dado a graça de atingir mais esse objetivo em minha vida.

A minha família, Claudio e Andrelina (meus pais), Cláudia, Tamara e Thayná (minhas irmãs), por todo amor, carinho e torcida.

Aos meus amigos, Leandro, Natália, Suzy e Helen.

Ao professor Dr. Héilton Ribeiro Tavares, pela orientação.

Aos colegas do Departamento de Estatística da UFPA e do PPGME/UFPA (Programa de Pós-Graduação em Matemática e Estatística da UFPA).

Ao CNPq pela apoio financeiro.

E a todos que contribuíram para esse trabalho e para a minha formação pessoal e profissional.

RESUMO

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. São feitas simulações e uma aplicação 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. Para este estudo, no qual a dependência entre as proficiências médias diminui com o tempo, a estrutura Auto-Regressiva de Ordem 1, em conjunto com a curva de crescimento logística, foi a que melhor ajustou os dados, tendo como definidor o Critério AIC (*Akaike's Information Criterion*).

Palavras Chaves: Estudo longitudinal, proficiências médias e curva de crescimento.

ABSTRACT

In this work are propose models to the behavior of means abilities of a same group of people accompanied along the time (longitudinal study), by the estimation of parameter of growth curves like: Linear,quadratic,logistical and Gompertz. The longitudinal characteristic puts a dependence among the results in the many conditions of evaluation that 's why we propose covariance structures to model this dependence. We considerate known the parameters of the item involved in the study. Are made simulations and an applications to the results of the research/evaluation of the abilities: Factors associates with the Fundescola project, applied to students in first series of school,in the period from 1999 to 2003. For this study, in which the dependence among the medium proficiencies decreases with the time, the Solemnity-regressive structure of Order 1, together with the curve of growth logistics, was what best adjusted the data, tends as molder the Criterion AIC (Akaike's Infromation Criterion).

Words Keys: Longitudinal study, medium proficiencies and curve of growth.

Conteúdo

Agradecimentos	vi
RESUMO	vii
ABSTRACT	viii
Lista de Tabelas	xi
1 Introdução	1
1.1 Conceitos Iniciais	1
1.1.1 Conteúdo histórico - TRI	2
1.2 Estimação	3
1.3 O contexto das avaliações educacionais no Brasil	4
1.4 Objetivos e Organização da dissertação	6
2 Descrição dos dados	7
2.1 Programa de Desenvolvimento da Escola	7
2.2 Etapas de Preparação para a Pesquisa	8
2.3 Amostra	9
3 Estimação	10
3.1 Modelo dicotômico para várias populações	10
3.1.1 Suposições do Modelo	11
3.2 Máxima Verossimilhança Marginal	12
3.2.1 Construção da Verossimilhança	12
3.3 Estimação dos Parâmetros Populacionais	13
3.4 Estruturas de Covariância	16
3.4.1 Matriz de Covariância Diagonal	16

3.4.2 Matriz de Covariância Uniforme	17
3.4.3 Matriz de Covariância Bandas	17
3.4.4 Matriz de Covariância AR(1)	18
4 Curvas de Crescimento com Estrutura de Dependência	20
4.1 Curva de Crescimento Linear	20
4.2 Curva de Crescimento Quadrática	21
4.3 Curva de Crescimento Logística	22
4.4 Curva de Crescimento Gompertz	22
4.5 Chutes iniciais	22
5 Aplicações	24
5.1 Aplicação a dados simulados	24
5.1.1 Resultados das simulações	25
5.2 Aplicação em Dados Reais	26
6 Conclusões e Sugestões	31
6.1 Conclusões	31
6.2 Sugestões para trabalhos futuros	32
A Parâmetros e Estimativas	33
A.1 Parâmetros de Itens - Simulações	33
A.2 Parâmetros de Itens - Dados Reais	35
Referências Bibliográficas	36

Lista de Tabelas

5.1	Valores dos parâmetros populacionais, Vetor de médias, e dos parâmetros das Curvas de Crescimento usados na simulação.	25
5.2	Resultados da simulação, usando a matriz Diagonal	25
5.3	Resultados da simulação, usando a matriz Uniforme	26
5.4	Resultados da simulação, usando a matriz Bandas	26
5.5	Resultados da simulação, usando a matriz AR(1)	26
5.6	Tabela de Itens Comuns	27
5.7	Matriz de correlações observada	28
5.8	Estimativas dos Parâmetros - Dados reais - Matriz Diagonal	28
5.9	Estimativas dos Parâmetros - Dados reais - Matriz Uniforme	28
5.10	Estimativas dos Parâmetros - Dados reais - Matriz Bandas	29
5.11	Estimativas dos Parâmetros - Dados reais - Matriz AR(1)	29
5.12	Log-verossimilhança associada a cada estrutura de covariância e curva . . .	30
A.1	Valores para os parâmetros dos itens usados na simulação	33
A.2	Estimativas dos parâmetros da matriz Diagonal obtidos na simulação . . .	34
A.3	Estimativas dos parâmetros das matrizes Uniforme, Bandas e AR(1), obtidos na simulação	34
A.4	Valores adotados para os parâmetros dos itens	35

Introdução

1.1 Conceitos Iniciais

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. Com a TRI temos informações dos níveis de dificuldade de cada item presente no teste, pois eles passam a ser analisados separadamente. É possível comparar alunos de séries distintas (por exemplo 8^a série do ensino fundamental e 3^a série do ensino médio), através de suas habilidades médias, desde que submetidas à provas que tenham alguns itens comuns, ou ainda, acompanhar o mesmo grupo de alunos ao longo do tempo, quando submetidos a um teste em cada condição temporal.

A proposta dessa teoria é modelar 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. Modelos que envolvem mais de duas categorias de resposta também são comuns na prática e serão comentados mais adiante.

Os modelos de resposta ao item se fundamentam em três fatores: a natureza do item (dicotômicos (0,1), entendidos por certo ou errado, ou policotômicos, recebendo escores classificados em 0, 1, 2 ou 3, por exemplo), o número de populações envolvidas e a quantidade de traços latentes (habilidades) que está sendo medida (podendo ser unidimensionais, medindo somente uma habilidade, ou multidimensionais, medindo mais de uma habilidade).

Maiores detalhes sobre essa metodologia, incluindo sua fundamentação matemática, podem ser encontradas em Andrade, Tavares & Valle (2000), entre outros.

1.1.1 Conteúdo histórico - TRI

A aplicação de Teoria da Resposta ao Item em testes educacionais teve início nos anos 50, quando Lord (1952) desenvolveu o modelo unidimensional de 2 parâmetros, os quais medem o nível de dificuldade do item e seu poder de discriminação. O parâmetro de discriminação refere-se ao quanto o item consegue separar alunos pertencentes a níveis de habilidades diferentes. Em seguida, Lord incluiu no modelo um terceiro parâmetro, chamado de acerto casual ou chute, este considera que alunos com baixa habilidade têm alguma probabilidade de acertar ao item.

Birnbaum, em 1968, substituiu a ogiva (normal acumulada), usadas por Lord, pela curva logística, tornando o tratamento matemático dos dados bem mais fácil, pois é uma função explícita dos parâmetros do item e de habilidade e não envolve integração. Em 1960 o matemático dinamarquês Georg Rasch reforçou os estudos da TRI criando o modelo unidimensional de 1 parâmetro, para muitos seu trabalho é tão pioneiro quanto o de Lord. Em 1968, Wright descreveu o modelo de 1 parâmetro de Rasch, desenvolvido com a ogiva normal, por um modelo logístico.

A própria necessidade de introduzir respostas que não fossem consideradas dicotômicas encarregou-se de motivar a ciência a desenvolver modelos para TRI que tratassem de tais necessidades como: Modelo de Resposta Gradual, Modelo de Crédito Parcial, Modelo de Escala Gradual, Modelos de Desdobramento, entre outros. Para maiores detalhes destes modelos recomenda-se Linden e Hambleton (1997).

Recentemente, Bock & Zimowski (1997) introduziram a abordagem para duas ou mais populações de respondentes. A introdução desta modelagem trouxe novas possibilidades para as comparações de rendimentos de duas ou mais populações submetidas a diferentes testes com itens comuns, conforme discutido em Hedges & Vevea (1997) e Andrade (1999), por exemplo.

Uma boa introdução sobre o assunto completando a teoria e prática pode ser encontrado em Andrade, Tavares & Valle (2000).

1.2 Estimação

Um passo importante na aplicação da TRI a dados de testes é a estimação dos parâmetros que caracterizam o modelo de resposta ao item. Assim, nos Modelos de Resposta ao Item (MRI) temos um problema de estimação que envolve os parâmetros dos itens e das habilidades dos indivíduos, ou ainda, a estimação dos parâmetros populacionais (habilidades médias, por exemplo) para modelos que envolvam mais de uma população. O problema pode ser caracterizado em três situações:

- Quando já conhecemos os parâmetros dos itens, temos apenas que estimar as habilidades;
- Quando já conhecemos as habilidades dos respondentes, estamos interessados apenas na estimação dos parâmetros dos itens (processo conhecido como *calibração*);
- Quando os parâmetros dos itens e das habilidades são desconhecidos, situação mais comum, desejamos estimar os parâmetros dos itens e das habilidades dos indivíduos simultaneamente.

Em qualquer uma das situações citadas acima, geralmente a estimação é feita pelo método da máxima verossimilhança através da aplicação de algum processo iterativo, como o algoritmo de *Newton-Raphson* (ver Issac & Keller, 1966, por exemplo) ou “*Scoring*” de Fisher (ver Rao, 1973, por exemplo). Alguns procedimentos Bayesianos também são aplicados com alguma frequência (ver Mislevy, 1986).

Na situação em que se deseja estimar tanto os parâmetros dos itens, quanto as habilidades, existem algumas formas de se abordar o problema de estimação: estimação *conjunta* (parâmetros dos itens e habilidades), *Método de Máxima Verossimilhança Marginal* (em duas partes, primeiro a estimação dos parâmetros dos itens e, posteriormente, das habilidades), *métodos bayesianos completos*, entre outros.

O Método de Máxima Verossimilhança Marginal, proposto por Bock & Lieberman (1970), baseia-se na existência uma distribuição de probabilidade latente associada à habilidade dos indivíduos da população em estudo Π , cuja função densidade de probabilidade (f_{dp}) é $g(\theta|\boldsymbol{\eta})$, onde $\boldsymbol{\eta}$ é o conjunto de parâmetros associados à Π . Apesar desse método produzir estimadores de máxima verossimilhança com as propriedades assintóticas desejáveis, ele exige a estimação conjunta dos parâmetros de todos os itens, o que torna inaplicável para testes com mais de 12 itens. Para contornar esse problema, Bock & Aitkin (1981) fizeram uma modificação no modelo de Bock & Lieberman adicionando a suposição

de independência entre os itens, e o uso do algoritmo **EM** (proposto por Dempster, Laird & Rubin, 1977), de tal modo que a matriz Hessiana, com relação aos parâmetros dos itens, pudesse se tornar bloco-diagonal, permitindo assim que os parâmetros para cada item pudessem ser estimados em separado.

1.3 O contexto das avaliações educacionais no Brasil

Atualmente a Teoria da Resposta ao Item (TRI) vem tornando-se a técnica predominante no contexto de avaliações educacionais no Brasil. Apesar dessa teoria ter mais de 50 anos, somente nos últimos 15 é que ela vem sendo aplicada em larga escala nas principais avaliações educacionais. Atribui-se este fato à complexidade matemática dos métodos envolvidos, praticamente inviáveis sem o auxílio do computador.

No Brasil ela foi usada pela primeira vez em 1995 na análise de dados do Sistema Nacional de Ensino Básico - SAEB. Uma amostra de cerca de 100 mil alunos brasileiros de 4ª e 8ª séries do Ensino Fundamental e 3ª série do Ensino Médio, de escolas públicas e particulares, localizadas nas áreas urbana e rural, representativas de todos os estados da Federação. Para a avaliação das proficiências dos alunos brasileiros, foram aplicados 1.245 itens de teste de matemática e de leitura. Os desempenhos desses alunos foram comparados e colocados em uma única escala de conhecimento. Desde então, o SAEB vem sendo aplicado a cada dois anos em uma amostra de alunos de cada Estado, permitindo acompanhar o desempenho médio desses alunos.

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

A prova Brasil foi criada em 2005 e tem o objetivo de avaliar alunos das 4ª e 8ª séries do ensino fundamental de todas as escolas públicas urbanas do país. A Prova Brasil avalia as habilidades em Língua Portuguesa, com foco em leitura, e Matemática, com foco na resolução de problemas. A avaliação é quase universal: todos os estudantes das séries avaliadas, de todas as escolas públicas urbanas do Brasil com pelo menos 20 alunos na série, devem fazer a prova. Por ser universal, expande o alcance dos resultados oferecidos pelo SAEB. Como resultado, fornece as médias de desempenho para o Brasil, regiões e unidades da Federação, para cada um dos municípios e escolas participantes. Em 2009

deverão ser avaliados cerca de 5,5 milhões de estudantes, e a análise dos dados é feita via TRI.

Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo - SA-RESP

O SARESP é um sistema de avaliação censitário dos alunos do Ensino Fundamental e Médio da Rede Estadual de São Paulo e de escolas municipais e particulares que aderiram ao Sistema de Avaliação. Visa alcançar dois objetivos. O primeiro seria ampliar o conhecimento do perfil de realização dos alunos, fornecendo aos professores informações sobre o desempenho dos alunos de modo a subsidiar o trabalho a ser desenvolvido em sala de aula. O segundo seria fornecer informações essenciais para a melhoria da gestão do sistema educacional, na medida em que identifica os pontos críticos do ensino e possibilita à SEE (Secretaria Estadual de Educação), por meio de seus órgãos centrais e das Delegacias de Ensino, apoiar as escolas e os educadores com recursos, serviços e orientações. A análise dos dados é hoje feita via TRI.

A título informativo, no município de São Paulo há outra avaliação chamada *Prova São Paulo*, em que a análise também é feita via TRI.

Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja)

O Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja) tem como principal objetivo avaliar as habilidades e competências básicas de jovens e adultos que não tiveram oportunidade de acesso à escolaridade regular na idade apropriada. O participante se submete a uma prova e, alcançando a nota mínima exigida, obtém a certificação de conclusão daquela etapa educacional. Tem sido feita em dois níveis, Fundamental e Médio. A análise do Encceja também feita via TRI.

Programa Internacional de Avaliação de Alunos (Pisa)

O PISA é um programa internacional de avaliação educacional organizado pela OCDE (Organização para Cooperação e Desenvolvimento Econômico), e propõe fazer uma avaliação do desempenho dos estudantes de 15 anos de idade para obter indicadores relativos aos sistemas educacionais dos países participantes. O Brasil participa, de forma voluntária, através do INEP. São avaliadas as capacidades de analisar, raciocinar e refletir sobre conhecimentos e experiências em três áreas: Leitura, Matemática e Ciências. A cada uma de suas edições, é dado foco a uma das áreas. O foco em 2000 foi Leitura; em 2003,

Matemática. Em 2006, a ênfase foi para a área de Ciências. Em 2009 reiniciou o ciclo, na mesma ordem das ênfases. Cabe ressaltar que a análise do Pisa é feita via TRI adotando-se o modelo logístico de um parâmetro, que é matematicamente o modelo de Rasch.

Exame Nacional do Ensino Médio (Enem)

Apartir da edição 2009 do Enem, a análise dos itens e as proficiências também serão realizadas com base na TRI. Milhares de itens foram e serão pré-testados continuamente para que haja mais de uma edição do Enem a cada ano a partir de 2010. A avaliação será constituída de quatro áreas: Ciências da natureza e suas tecnologias, Ciências humanas e suas tecnologias, Linguagens e códigos e suas tecnologias e Matemática e suas tecnologias. Para cada área haverá uma nota via TRI baseada em 45 itens.

Em todas essas avaliações, e outros estudos não citados, há ênfase na obtenção das proficiências médias dos grupos estudados, que podem apresentar uma determinada evolução ao longo do tempo. Neste caso, será de grande interesse a modelagem dessas proficiências médias, através de algumas funções de interesse. Ainda, no caso em que um mesmo grupo de indivíduos é avaliado ao longo do tempo, apresentando uma possível dependência entre as habilidades, é desejável estudar possíveis estruturas de covariância que acomodem tal dependência.

1.4 Objetivos e Organização da dissertação

O objetivo desta dissertação é o de estimar as habilidades médias de um mesmo grupo de indivíduos, ao longo de T condições de avaliação, através da estimação dos parâmetros de Curvas de Crescimento. Apresentamos as equações de estimação para os modelos Linear, Quadrático, Logístico e Gompertz.

No Capítulo 2 descrevemos os dados usados no trabalho, o planejamento longitudinal e os procedimentos de amostragem. O Capítulo 3 é dedicado a estimação dos parâmetros populacionais, com os parâmetros dos itens conhecidos, ainda serão apresentadas as diversas estruturas de covariância usadas no trabalho. No Capítulo 4 são apresentadas as Curvas de Crescimento associadas aos vetores de médias. No Capítulo 5 são feitas aplicações a dados simulados e a dados reais. E no Capítulo 6 são feitas conclusões e sugestões para trabalhos futuros.

Descrição dos dados

2.1 Programa de Desenvolvimento da Escola

O Programa de Desenvolvimento da Escola (PDE - Escola) foi um programa conduzido pelo MEC e coordenado pelo FUNDESCOLA (Fundo de Fortalecimento da Escola), em parceria com governos estaduais e municipais das regiões Norte, Nordeste e Centro-Oeste, com o objetivo de melhorar a qualidade do ensino, para garantir o melhor desempenho dos alunos e assumir o controle dos processos que desenvolve.

Embora a metodologia do PDE possa ser utilizada por escolas menores, as que participam do programa deveriam ter acima de 200 alunos no ensino fundamental de acordo com o censo escolar, condições mínimas de funcionamento (como banheiro, luz e água encanada), um diretor que exerça liderança na escola e, por fim, devem ser uma Unidade Executora constituída junto ao Fundo Nacional para o Desenvolvimento da Educação (FNDE).

Cada escola elaborou seu próprio PDE a partir de um diagnóstico de sua situação, que é a base para a definição de objetivos, estratégias, metas e planos de ação a serem alcançados a longo, médio e curto prazo. Assim, o sucesso do processo de elaboração e implantação do PDE depende da operacionalização de todos esses princípios.

Pesquisas foram desenvolvidas pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais) para avaliar os resultados obtidos pelas escolas que implementaram o PDE. Entre estas está a pesquisa “Avaliação do desempenho: Fatores Associados”, aplicada a alunos de 4a. até a 8a. série do Ensino Fundamental, no período de 1999 a 2003. O INEP propôs a realização desta pesquisa para medir o impacto do PDE sobre o desempenho dos alunos em provas de conhecimento, o fluxo escolar, redução na distorção idade/série, permanência ao longo das séries e frequência às aulas, entre outros.

2.2 Etapas de Preparação para a Pesquisa

O INEP selecionou seis Estados para a realização da pesquisa, dois em cada uma das regiões atendidas pelo Fundescola (Norte, Nordeste, Centro-Oeste). Os Estados escolhidos foram: Pará e Rondônia da Região Norte, Pernambuco e Sergipe da Região Nordeste e Mato Grosso do Sul e Goiás da Região Centro-oeste.

As Secretarias de Educação dos Estados e Municípios envolvidos com a pesquisas receberam contatos do INEP, e cada diretor das escolas incluídas na amostra recebeu um comunicado do INEP solicitando a colaboração para participar da pesquisa.

Foram utilizados procedimentos usuais de uma pesquisa de campo, ou seja, treinamento de pesquisadores de campo, elaboração de material para aplicação dos instrumentos, controle e crítica dos materiais aplicados. Todos esses instrumentos foram digitalizados, facilitando revisitá-los sem a necessidade de retornar aos instrumentos originais em papel.

Foi elaborado um Roteiro da Escola com questões referentes à estrutura e ao estado físico da unidade escolar. Tal questionário, a ser respondido pelos aplicadores dos testes, tinha por objetivo apresentar os aspectos materiais e físicos vividos pela escola, tais como características do aluno, do professor, do diretor e da escola.

Os itens das provas foram todos pré-testados antes da elaboração da pesquisa. Em Abril de 1999, foram aplicados testes de Português e de Matemática aos alunos da 4a. série, com itens do Banco Nacional de Itens (INEP/DAEB) compatíveis com o grau de dificuldade da população pesquisada. O objetivo era de construir a linha de base (*baseline*) para as futuras medidas de desempenho. Os professores foram convidados a responder um pequeno questionário sobre a prova, para isso, deveriam também responder às questões da prova.

Em Novembro foram aplicados novos testes aos mesmos alunos e professores, com alguns itens comuns à prova anterior, de modo a medir o conhecimento no final do ano letivo. As provas de cada disciplina foram compostas de 10 itens aplicados em abril e 30 novos itens pré-testados.

Além dos testes de Português e de Matemática, foram aplicados questionários aos alunos, aos seus professores e aos diretores dessas escolas. O objetivo foi correlacionar algumas características dos envolvidos com a aprendizagem dos alunos e o desempenho destes nas provas. Estes questionários possuíam questões, em geral, de cunho sócio-econômico e para os professores e diretores foram incluídas questões pedagógicas e outras referentes à gestão escolar. Para a escola, as questões estavam basicamente relacionadas a infra-estrutura física e os serviços disponíveis.

Foi criado um banco de dados com todas as informações coletadas nos questionários e nos testes. Além disso, foi elaborado um instrumento - as Fichas Escolares - contendo as Médias Finais em Português e em Matemática, Resultado Final (Aprovado, Reprovado, Evadido, Transferido) e faltas no ano letivo.

2.3 Amostra

O banco de dados da pesquisa possui informações das áreas urbanas das micro-regiões de seis Estados brasileiros, Sergipe, Pernambuco, Pará, Rondônia, Goiás e Mato Grosso do Sul - representantes das regiões Nordeste, Norte e Centro-Oeste, com informações coletadas para os anos de 1999, 2000, 2001, 2002 e 2003, respectivamente para os alunos de 4a. a 8a. série do Ensino fundamental.

Para a primeira amostra, em abril de 1999, foi analisado, primeiramente, o perfil censitário das 400 escolas que participavam do PDE. Foram selecionadas 120 dessas escolas para constituir o Grupo PDE, e em seguida, selecionadas outras 120, sem o PDE, para compor o Grupo Controle. As escolas selecionadas abrangiam, aproximadamente, 18.000 alunos matriculados nas 4as. séries dos turnos manhã, tarde e intermediário. Procedeu, então, a uma redução no número das escolas até chegar ao total de 158 escolas, 400 professores e 12.641 alunos.

Em Novembro de 1999 uma nova prova foi aplicada para 10.686 alunos, com interesse de avaliar o desempenho dos alunos, ao final do ano letivo. Esta prova era composta de 40 itens de matemática, sendo 10 itens comuns com a prova aplicada em abril e 30 itens novos.

As provas aplicadas em Novembro de 2000, 2001, 2002 e 2003, tiveram como objetivo principal avaliar o rendimento escolar dos alunos referente a disciplina matemática ao longo das séries, a qual foram respondidas por 15.494, 15.280, 12.191 e 10.101 alunos, respectivamente, da 5a., 6a., 7a. e 8a. séries.

Durante todo o processo, muitos alunos foram excluídos e outros incluídos no estudo. O número de alunos que participou dos seis instantes de observação foi 1978, que comporão a base de dados a ser analisada neste trabalho para efeito de estimação de curvas de crescimento.

No próximo capítulo trataremos do processo de estimação para obtenção das equações de estimação dos parâmetros populacionais do modelo.

Estimação

3.1 Modelo dicotômico para várias populações

Vamos iniciar considerando a situação em que um mesmo grupo de N alunos responde a T testes em cada condição de avaliação. O conjunto de itens aplicados a cada teste não é o mesmo, por se tratar de diferentes graus de escolaridade, porém temos itens comuns entre esses testes.

Para fins de aplicação utilizamos o Modelo Logístico de 3 parâmetros (ML3), que tem sido o mais utilizado pelos pesquisadores da área.

Vamos considerar a situação em que indivíduos são submetidos a um teste t com I_t itens, de múltipla escolha, selecionados de um conjunto total de I itens, dicotômicos ou dicotomizados. A resposta do indivíduo j no teste t ao item i , será representado pela variável aleatória de Bernoulli:

$$U_{jit} = \begin{cases} 1, & \text{se o item } i \text{ é respondido corretamente pelo indivíduo } j \text{ no teste } t; \\ 0, & \text{se o item } i \text{ não é respondido corretamente pelo indivíduo } j \text{ no teste } t. \end{cases}$$

com $i = 1, \dots, I_t$, $j = 1, \dots, N$ e $t = 1, \dots, T$.

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:

$$P(U_{jit} = 1 | \theta_{jt}, \zeta_i) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_{jt} - b_i)}}, \quad (3.1)$$

onde, θ_{jt} é a habilidade do j -ésimo respondente ao teste t , ζ_i é o conjunto de características associadas ao item i , com $\zeta_i = (a_i, b_i, c_i)'$, b_i é o parâmetro de dificuldade (ou de posição), a_i é o parâmetro de discriminação, c_i é o parâmetro do item que representa a probabilidade (aproximada) de indivíduos com baixa habilidade responderem corretamente o item, e D é um fator de escala, constante e igual a 1. Utiliza-se o valor 1,7 quando deseja-se que a função logística forneça resultados semelhantes ao da função ogiva normal.

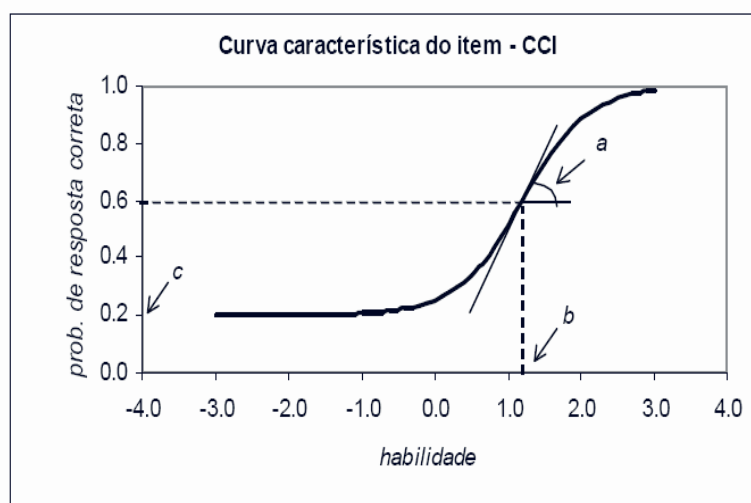


Figura 3.1 Curva Característica do Item – CCI

A Figura 3.1 nos mostra a relação entre a probabilidade de resposta correta a determinado item e os parâmetros do modelo. Podemos dizer que quanto maior o parâmetro b_i mais difícil será o item, maior a habilidade exigida e menor probabilidade terá um indivíduo de acertar o item, fixada sua habilidade. O parâmetro c_i nos dá a probabilidade de acerto ao acaso. Para itens com m alternativas, espera-se que c_i esteja próximo de $1/m$. O parâmetro a_i distingue indivíduos com habilidades diferentes no nível da escala em torno do nível de dificuldade do item.

3.1.1 Suposições do Modelo

Os modelos de Resposta ao Item incluem um conjunto de hipóteses sobre os dados em que o modelo é aplicado.

A primeira hipótese é chamada de *unidimensionalidade* de teste, isto é, o conjunto de itens que mede um único traço latente. É evidente que dentro das instâncias do ser humano existe uma variedade de habilidades responsáveis por um processo de execução de uma tarefa. Porém, para satisfazer tal postulado, é suficiente admitir que haja uma habilidade dominante, responsável pelo conjunto de itens. Esse fator é que se supõe estar sendo medido pelo teste (Hambleton, Swaminathan & Rogers, 1991).

Uma outra suposição é a chamada *independência local* (ou *independência condicional*), na qual os itens que compõem cada teste são respondidos de forma independente por cada

indivíduo com habilidade θ . Em outras palavras, considera-se que para o indivíduo responder aos itens, ele necessita apenas da sua habilidade, não importando se ele respondeu ou não a outros itens, e sua habilidade não se altera durante o teste.

Temos ainda a *independência temporal*, que declara que os itens de um teste serão respondidos de maneira independentes dos itens de outros testes, fixadas as habilidades nos T testes.

Além disso, na presença de várias populações temos a necessidade de mais uma suposição. Suporemos que as respostas oriundas de indivíduos de testes diferentes serão independentes (ver Tavares, 2001).

3.2 Máxima Verossimilhança Marginal

O método utilizado para estimação é o de Máxima Verossimilhança Marginal, no qual adotamos uma determinada distribuição para a habilidade dos indivíduos da população em estudo, permitindo criar estruturas para os parâmetros das respectivas funções densidade de probabilidade (fdp).

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, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_T)'$, 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 $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_T)'$ é o vetor de médias e $\boldsymbol{\Sigma} = (\sigma_{tj})_{t,j}$, é a matriz de covariâncias.

3.2.1 Construção da Verossimilhança

Embora tenhamos T testes, devemos notar que alguns itens estão em dois ou mais testes. Por conta disso vamos fazer uma ordenação nos I itens que compõem o conjunto dos T testes, representando-os por $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_I)$ e denotando por $\mathbf{I}_t$ o conjunto dos índices dos itens presentes no teste t . Considerando I_t o número de itens no teste t , teremos que $I < \sum_{t=1}^T I_t$. Sejam N_t o número de indivíduos que responderão o teste t e N o número total de indivíduos na amostra. Sejam ainda, $\mathbf{U}_{j,t} = (U_{j1t}, U_{j2t}, \dots, U_{jI_t t})'$ o vetor aleatório de respostas do indivíduo j no teste t , $\mathbf{U}_{j..} = (\mathbf{U}'_{j,1}, \mathbf{U}'_{j,2}, \dots, \mathbf{U}'_{j,T})'$ o vetor

aleatório de respostas do indivíduo j em todos os testes, e $\mathbf{U}_{..} = (\mathbf{U}'_{..1}, \mathbf{U}'_{..2}, \dots, \mathbf{U}'_{..N})'$ o vetor total de respostas.

Assumindo a independência local das respostas para os itens no teste t , dada a habilidade θ_t , temos que

$$P(\mathbf{U}_{j.t}|\theta_t, \boldsymbol{\zeta}) = \prod_{i \in \mathbf{I}_t} P(U_{jit}|\theta_t, \zeta_i). \quad (3.2)$$

Com a independência temporal podemos escrever

$$P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta}) = \prod_{t=1}^T \prod_{i \in \mathbf{I}_t} P(U_{jit}|\theta_t, \zeta_i). \quad (3.3)$$

A probabilidade marginal do vetor de respostas $\mathbf{U}_{j..}$ é dada por

$$P(\mathbf{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \int_{\mathbb{R}} P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta}) g(\boldsymbol{\theta}|\boldsymbol{\eta}) d\boldsymbol{\theta}. \quad (3.4)$$

E a verossimilhança marginal será representada por

$$L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \prod_{j=1}^N P(\mathbf{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta}). \quad (3.5)$$

Consequentemente, a log-verossimilhança será

$$\log L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \sum_{j=1}^N \log P(\mathbf{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta}). \quad (3.6)$$

Vale ressaltar que, embora neste trabalho os parâmetros dos itens $\boldsymbol{\zeta}$ sejam conhecidos de estudos prévios, o símbolo $\boldsymbol{\zeta}$ será mantido na notação para fins de uniformização com a literatura.

3.3 Estimação dos Parâmetros Populacionais

Na presença de várias populações, temos mais um conjunto de parâmetro a estimar: $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_T)$, que são referidos como *Parâmetros Populacionais*. Como são conhecidos os parâmetros dos itens, o interesse está em estimar apenas estes parâmetros. As equações de estimação para os parâmetros populacionais são obtidas através de:

$$\frac{\partial \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \mathbf{0}. \quad (3.7)$$

Para chegarmos a equação de estimação em (3.7), precisaremos de alguns resultados intermediários. Notemos que,

$$\frac{\partial \log L(\zeta, \eta)}{\partial \eta} = \sum_{j=1}^N \frac{1}{P(\mathbf{U}_{j..}|\zeta, \eta)} \frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta} \quad (3.8)$$

e

$$\begin{aligned} \frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta} &= \int_{\mathbb{R}} P(\mathbf{U}_{j..}|\theta, \zeta) \left(\frac{\partial g(\theta|\eta)}{\partial \eta} \right) d\theta \\ &= \int_{\mathbb{R}} P(\mathbf{U}_{j..}|\theta, \zeta) g(\theta|\eta) \left(\frac{\partial \log g(\theta|\eta)}{\partial \eta} \right) d\theta. \end{aligned} \quad (3.9)$$

Então, segue-se que a equação de estimação para η é

$$\begin{aligned} \frac{\partial \log L(\zeta, \eta)}{\partial \eta} &= \sum_{j=1}^N \frac{1}{P(\mathbf{U}_{j..}|\zeta, \eta)} \int_{\mathbb{R}} P(\mathbf{U}_{j..}|\theta, \zeta) g(\theta|\eta) \left(\frac{\partial \log g(\theta|\eta)}{\partial \eta} \right) d\theta \\ &= \sum_{j=1}^N \int_{\mathbb{R}} \left(\frac{\partial \log g(\theta|\eta)}{\partial \eta} \right) g_j^*(\theta) d\theta = 0 \end{aligned} \quad (3.10)$$

onde,

$$g_j^*(\theta) = \frac{P(\mathbf{U}_{j..}|\theta, \zeta) g(\theta|\eta)}{P(\mathbf{U}_{j..}|\zeta, \eta)}. \quad (3.11)$$

Se utilizarmos a distribuição Normal Multivariada $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, onde $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_T)'$ é o vetor de médias nos T testes e $\boldsymbol{\Sigma} = (\sigma_{tj})_{t,j}$, é a matriz de covariâncias associada ao teste t . Teremos que

$$\log g(\theta|\eta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} (\theta - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\theta - \boldsymbol{\mu}).$$

Pensando agora que o vetor de médias, assim como a matriz de covariância, pode ser função de um subfator, é conveniente desenvolver este subcaso. Fazendo $\eta = \eta(\alpha)$, onde α é um vetor qualquer, pela Regra da Cadeia temos

$$\frac{\partial \log g(\theta|\eta)}{\partial \alpha} = \left(\frac{\partial \log g(\theta|\eta)}{\partial \boldsymbol{\mu}} \right) \left(\frac{\partial \boldsymbol{\mu}}{\partial \alpha} \right). \quad (3.12)$$

Dessa forma, precisamos encontrar

$$\frac{\partial \log g(\theta|\eta)}{\partial \boldsymbol{\mu}}$$

Temos que

$$\frac{\partial \log g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\mu}} = -\frac{1}{2} \frac{\partial}{\partial \boldsymbol{\mu}} [(\boldsymbol{\theta} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu})].$$

Fazendo $\mathbf{V} = \boldsymbol{\Sigma}^{-1} = (v_{tj})$, teremos

$$\begin{aligned} \frac{\partial \log g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\mu}_t} &= -\frac{1}{2} \frac{\partial}{\partial \boldsymbol{\mu}} [(\boldsymbol{\theta} - \boldsymbol{\mu})' \mathbf{V} (\boldsymbol{\theta} - \boldsymbol{\mu})] \\ &= \mathbf{v}_t' (\boldsymbol{\theta} - \boldsymbol{\mu}), \quad t = 1, \dots, T, \end{aligned} \quad (3.13)$$

onde $\mathbf{v}_t$ é a t -ésima coluna de $\mathbf{V}$. Podemos escrever a equação (3.13) em forma matricial

$$\frac{\partial \log g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\mu}} = \begin{pmatrix} \frac{\partial \log g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \mu_1} \\ \frac{\partial \log g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \mu_2} \\ \vdots \\ \frac{\partial \log g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \mu_T} \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1' \\ \mathbf{v}_2' \\ \vdots \\ \mathbf{v}_T' \end{pmatrix} (\boldsymbol{\theta} - \boldsymbol{\mu}) = \mathbf{V} (\boldsymbol{\theta} - \boldsymbol{\mu}). \quad (3.14)$$

Assim, a equação de estimação para $\boldsymbol{\alpha}$ é

$$\boldsymbol{\alpha} : \sum_{j=1}^N \int_{\mathbb{R}} \mathbf{V}(\boldsymbol{\theta} - \boldsymbol{\mu}) \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\alpha}} \right) g_j^*(\boldsymbol{\theta}) d\boldsymbol{\theta} = \mathbf{0}. \quad (3.15)$$

Observe que as integrais descritas em (3.15) não possuem solução explícita, sendo necessário o uso de algum método numérico para a aproximação de integrais. O método *Hermite-Gauss*, também denominado de *método de quadratura gaussiana*, tem sido o mais usado na TRI.

Se $g(\boldsymbol{\theta}|\boldsymbol{\eta})$ é uma função contínua com integral finita, ela pode ser aproximada, para qualquer grau de precisão, por uma função discreta com um número finito de pontos (isto é, um histograma). Dessa forma, podemos substituir o cálculo de integral pela soma das áreas de um número finito, digamos q , de retângulos. Os pontos médios de cada retângulo, $\bar{\boldsymbol{\theta}}_l, l = 1, \dots, q$, são denominados de nós, ou pontos de quadratura. Cada nó tem um peso $A_l \equiv A(\bar{\boldsymbol{\theta}}_l)$ associado que leva em conta a altura $g(\bar{\boldsymbol{\theta}}_l|\boldsymbol{\eta})$ e a largura (Δ_l) do respectivo intervalo, tal como $A_l = g(\bar{\boldsymbol{\theta}}_l|\boldsymbol{\eta}) \times \Delta_l$. Para obtenção desses valores, ver Hildebrand (1956), páginas 327-330. Conhecidos os nós $\bar{\boldsymbol{\theta}}_l$ e os pesos, $A_l \equiv A(\bar{\boldsymbol{\theta}}_l)$, podemos substituir as aproximações por pontos de quadratura na equação (3.15), concluindo que

$$\sum_{j=1}^N \sum_{l=1}^q \mathbf{V}(\boldsymbol{\theta}_l - \boldsymbol{\mu}) \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\alpha}} \right) g_j^*(\bar{\boldsymbol{\theta}}_l) = \mathbf{0}. \quad (3.16)$$

E trocando a ordem dos somatórios,

$$\sum_{l=1}^q \mathbf{V}(\boldsymbol{\theta}_l - \boldsymbol{\mu}) \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\alpha}} \right) \bar{\mathbf{g}}^*(\bar{\boldsymbol{\theta}}_l) = \mathbf{0} \quad (3.17)$$

onde

$$\bar{\mathbf{g}}^*(\bar{\boldsymbol{\theta}}_l) = \frac{1}{N} \sum_{j=1}^N g_j^*(\bar{\boldsymbol{\theta}}_l) \quad (3.18)$$

com

$$g_j^*(\bar{\boldsymbol{\theta}}_l) = \frac{P(\mathbf{U}_{j..} | \bar{\boldsymbol{\theta}}_l, \boldsymbol{\zeta}) A_l}{\sum_{l=1}^q P(\mathbf{U}_{j..} | \bar{\boldsymbol{\theta}}_l, \boldsymbol{\zeta}) A_l}. \quad (3.19)$$

Vale ressaltar que na expressão (3.17) dividimos a expressão (3.16) pela constante N , o que não altera a solução da equação de estimação. No entanto, o processo de estimação pode se tornar mais rápido, uma vez que a parcela somável fica mais simples. Ainda, a expressão (3.18) caracteriza a densidade $\bar{\mathbf{g}}^*(\bar{\boldsymbol{\theta}}_l)$ como a média das densidades posteriores dos indivíduos. Se os parâmetros dos itens forem conhecidos, como é o caso tratado neste trabalho, bem como os pontos de quadratura, então $\bar{\mathbf{g}}^*(\bar{\boldsymbol{\theta}}_l), l = 1, \dots, q$, são fixos durante todo o processo de estimação.

3.4 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.

3.4.1 Matriz de Covariância Diagonal

Se considerarmos que a habilidade dos alunos em um tempo é não-correlacionada das habilidades nos outros $T - 1$ tempos, podemos dizer que a matriz de covariâncias das

habilidades nas T condições de avaliação é dada por

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_T^2 \end{pmatrix} \equiv \mathbf{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_T^2\}. \quad (3.20)$$

Logo,

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

3.4.2 Matriz de Covariância Uniforme

Com essa estrutura temos que as variâncias são sempre as mesmas, iguais a σ^2 , ao longo do tempo, e que a correlação entre as habilidades dos alunos é igual a ρ . É dada por:

$$\mathbf{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{pmatrix}, \quad (3.21)$$

com $\rho \in (-1, 1)$.

Esta matriz pode ser escrita como

$$\mathbf{\Sigma} = \sigma^2(1 - \rho)\mathbf{I}_T + \sigma^2\rho\mathbf{J}_T$$

onde, $\mathbf{I}_T$ é a matriz identidade e $\mathbf{J}_T$ é uma matriz ($T \times T$) de elementos unitários. Com isso, teremos que

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

3.4.3 Matriz de Covariância Bandas

Essa estrutura é um pouco semelhante a anterior, porém a diferença está no fato de que apenas as habilidades nos tempos anterior e posterior, a um determinado tempo, são correlacionadas entre si.

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & 0 & \dots & 0 \\ \rho & 1 & \rho & \dots & 0 \\ 0 & \rho & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}, \quad (3.22)$$

onde $\rho \in (-k, k)$, com k dependendo do valor de T . Para $T = 3$, $k \approx 1/\sqrt{2}$, e para T grande, $k = 1/2$.

Para esta estrutura temos que,

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

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

3.4.4 Matriz de Covariância AR(1)

Essa estrutura assume que as correlações entre as habilidades diminuem ao longo que a distância entre os instantes das observações aumenta. As variâncias são sempre as mesmas em todos os instantes. A forma dessa matriz é

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \rho & 1 & \rho & \dots & \rho^{T-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \dots & 1 \end{pmatrix}. \quad (3.23)$$

com $\rho \in (-1, 1)$.

Temos que,

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

Com,

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

Apresentadas as principais estruturas de covariância constantes na literatura (Tavares, 2001), resta-nos tratar efetivamente das curvas de crescimento. No próximo capítulo serão apresentados alguns modelos para Curvas de Crescimento, tais como Linear, Quadrático, Logístico e Gompertz.

Curvas de Crescimento com Estrutura de Dependência

Neste capítulo vamos explorar algumas estruturas para os parâmetros populacionais de localização. O interesse agora se concentra no vetor de médias, $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_T)'$, que representa as habilidades médias para cada teste.

Os grupos envolvidos na análise corresponderão as séries $\mathbf{t} = (s_1, s_2, \dots, s_T)'$. Se essas séries são sequenciais, podemos adotar $s_1 = 1, s_2 = 2, \dots, s_T = T$. Quando consideramos modelos de curvas de crescimento para as habilidades médias em anos escolares diferentes, temos

$$\mu_t = f(s_t|\boldsymbol{\alpha}),$$

onde f é uma função duplamente diferenciável com p parâmetros e $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_p)'$ o vetor de parâmetros da função. O vetor de médias nos T testes pode ser escrito como

$$\boldsymbol{\mu} = (f(s_1|\boldsymbol{\alpha}), f(s_2|\boldsymbol{\alpha}), \dots, f(s_T|\boldsymbol{\alpha}))'.$$

Nosso principal objetivo é, com base em cada curva de crescimento, determinar $\partial\boldsymbol{\mu}/\partial\boldsymbol{\alpha}$ presente na expressão (3.16).

Abaixo vamos apresentar as principais funções propostas na teoria de curvas de crescimento, e suas respectivas equações de estimação.

4.1 Curva de Crescimento Linear

O modelo para curva de crescimento Linear é dado por

$$\mu_t = \alpha_0 + \alpha_1 s_t. \tag{4.1}$$

Para obter as equações de estimação dos respectivos parâmetros, temos que encontrar

as derivadas em relação a cada parâmetro. Assim,

$$\begin{aligned}\frac{\partial \mu_t}{\partial \alpha_0} &= 1 \\ \frac{\partial \mu_t}{\partial \alpha_1} &= s_t, \quad t = 1, \dots, T.\end{aligned}$$

Portanto,

$$\frac{\partial \boldsymbol{\mu}}{\partial \alpha_0} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{(T \times 1)} = \mathbf{l}_0 \quad \text{e} \quad \frac{\partial \boldsymbol{\mu}}{\partial \alpha_1} = \begin{pmatrix} 1 \\ 2 \\ \vdots \\ T \end{pmatrix}_{(T \times 1)} = \mathbf{l}_1. \quad (4.2)$$

4.2 Curva de Crescimento Quadrática

O modelo para curva de crescimento Quadrática é dado por

$$\mu_t = \alpha_0 + \alpha_1 s_t + \alpha_2 s_t^2. \quad (4.3)$$

Para obter as equações de estimação, precisamos das seguintes expressões:

$$\begin{aligned}\frac{\partial \mu_t}{\partial \alpha_0} &= 1 \\ \frac{\partial \mu_t}{\partial \alpha_1} &= s_t \\ \frac{\partial \mu_t}{\partial \alpha_2} &= s_t^2.\end{aligned}$$

$t = 1, \dots, T$. Portanto,

$$\frac{\partial \boldsymbol{\mu}}{\partial \alpha_0} = \mathbf{l}_0, \quad \frac{\partial \boldsymbol{\mu}}{\partial \alpha_1} = \mathbf{l}_1, \quad \frac{\partial \boldsymbol{\mu}}{\partial \alpha_2} = \begin{pmatrix} 1^2 \\ 2^2 \\ \vdots \\ T^2 \end{pmatrix}_{(T \times 1)} = \mathbf{l}_2, \quad (4.4)$$

onde $\mathbf{l}_0$ e $\mathbf{l}_1$ são dados por (4.2).

4.3 Curva de Crescimento Logística

O modelo para curva de crescimento Logística é dado por

$$\mu_t = \frac{\alpha_0}{1 + e^{-\alpha_1(s_t - \alpha_2)}}. \quad (4.5)$$

Para obter as equações de estimação, precisamos das seguintes expressões:

$$\begin{aligned} \frac{\partial \mu_t}{\partial \alpha_0} &= \frac{1}{1 + e^{-\alpha_1(s_t - \alpha_2)}} \\ \frac{\partial \mu_t}{\partial \alpha_1} &= \frac{\alpha_0(s_t - \alpha_2)e^{-\alpha_1(s_t - \alpha_2)}}{(1 + e^{-\alpha_1(s_t - \alpha_2)})^2} \\ \frac{\partial \mu_t}{\partial \alpha_2} &= -\frac{\alpha_0\alpha_1 e^{-\alpha_1(s_t - \alpha_2)}}{(1 + e^{-\alpha_1(s_t - \alpha_2)})^2}, \end{aligned}$$

para $t = 1, \dots, T$.

4.4 Curva de Crescimento Gompertz

O modelo para curva de crescimento Gompertz é dado por

$$\mu_t = \alpha_0 + \alpha_1 e^{-e^{-\alpha_2(s_t - \alpha_3)}}. \quad (4.6)$$

Para obter as equações de estimação, precisamos das seguintes expressões:

$$\begin{aligned} \frac{\partial \mu_t}{\partial \alpha_0} &= 1 \\ \frac{\partial \mu_t}{\partial \alpha_1} &= e^{-e^{-\alpha_2(s_t - \alpha_3)}} \\ \frac{\partial \mu_t}{\partial \alpha_2} &= \alpha_1(s_t - \alpha_3)e^{-\alpha_2(s_t - \alpha_3) - e^{-\alpha_2(s_t - \alpha_3)}} \\ \frac{\partial \mu_t}{\partial \alpha_3} &= -\alpha_1\alpha_2 e^{-\alpha_2(s_t - \alpha_3) - e^{-\alpha_2(s_t - \alpha_3)}}, \end{aligned}$$

para $t = 1, \dots, T$.

4.5 Chutes iniciais

Uma etapa importante usada no processo de estimação é estabelecer valores iniciais para os parâmetros das curvas de crescimento. Vamos considerar que conhecemos estimativas iniciais razoáveis para μ_t , $t = 1, \dots, T$. Para as curvas Linear e Quadrática, as

estimativas iniciais serão obtidas através do Método de Mínimos Quadrados. Os elementos do vetor $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots, \alpha_p)'$, com $p = 0, 1, 2, 3$, serão encontrados obtendo-se a seguinte estimativa para $\boldsymbol{\alpha}$:

$$\hat{\boldsymbol{\alpha}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}, \quad (4.7)$$

onde,

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & T \end{pmatrix} \quad \text{e} \quad \mathbf{Y} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_T \end{pmatrix}.$$

Para a curva de crescimento Logística, o valor inicial para o parâmetro α_0 (assíntota à direita) será próximo de μ_T , ou seja, $\hat{\alpha}_0 \simeq \hat{\mu}_T$. As estimativas para os parâmetros α_1 e α_2 serão encontradas através da equação (4.17), onde:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & T \end{pmatrix} \quad \text{e} \quad \mathbf{Y} = \begin{pmatrix} \ln\left(\frac{\alpha_0}{\mu_1} - 1\right) \\ \vdots \\ \ln\left(\frac{\alpha_0}{\mu_T} - 1\right) \end{pmatrix}.$$

Vale ressaltar, neste caso, que quando $t = T$ o último elemento fica $\ln(0)$, de forma que é conveniente multiplicar cada estimativa inicial de μ_t por algum fator superior a 1, tal como 1,05. O mesmo artifício vale para outras curvas.

Finalmente, para a curva de crescimento Gompertz, teremos que o valor inicial para α_0 será o valor mínimo de $\boldsymbol{\mu}$, $\hat{\alpha}_0 = \min(\hat{\mu}_1, \dots, \hat{\mu}_T)$. Para o parâmetro α_1 , o valor inicial será o valor máximo de $\boldsymbol{\mu}$, $\hat{\alpha}_1 = \max(\hat{\mu}_1, \dots, \hat{\mu}_T)$. Já as estimativas de α_2 e α_3 serão encontradas através da equação (4.17), onde:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & T \end{pmatrix} \quad \text{e} \quad \mathbf{Y} = \begin{pmatrix} \ln\left[-\ln\left(\frac{\mu_1 - \alpha_0}{\alpha_1}\right)\right] \\ \vdots \\ \ln\left[-\ln\left(\frac{\mu_T - \alpha_0}{\alpha_1}\right)\right] \end{pmatrix}.$$

Apresentados até agora a abordagem teórica do trabalho, no próximo capítulo serão feitas as aplicações a dados simulados e dados reais.

Aplicações

5.1 Aplicação a dados simulados

Nesta seção apresentaremos os resultados de simulações da metodologia proposta, considerando cada curva de crescimento apresentada e cada uma das estruturas de covariância adotadas.

Considerou-se que $N=1000$ indivíduos responderam a testes em $T=5$ condições de avaliação. Os resultados foram obtidos a partir de $M=100$ réplicas. Os indivíduos responderam cada teste composto de 24 itens com seis itens comuns entre dois testes consecutivos, ou seja, 6 itens comuns entre os testes 1 e 2, 6 entre os testes 2 e 3, 6 entre os testes 3 e 4 e 6 entre os testes 4 e 5. O teste 1 contou com a presença dos itens 1 a 24, o teste 2 com os itens 19 a 42, o teste 3 com os itens 37 a 60, o teste 4 pelos itens 55 a 78 e o teste 5 com os itens 73 a 96.

Os valores adotados para os parâmetros dos itens estão apresentados no Apêndice A, na Tabela A.1, utilizou-se 1,7 para o fator de escada D . O parâmetro de discriminação, a_i , variou de 0.6 (baixa discriminação) a 1.4 (alta discriminação) e os valores do parâmetro de dificuldade, b_i , variam de -0.7 a 4.7. Para o parâmetro c_i consideramos apenas o valor 0.20.

Assumimos que a distribuição das habilidades nas T condições de avaliação é Normal Multivariada, $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. O vetor de médias $\boldsymbol{\mu}$, considerando cada curva de crescimento, assim como os valores dos parâmetros dessas curvas, serão mostrados na Tabela 5.2. Para a matriz de covariância Diagonal consideramos para os parâmetros os seguintes valores: $\sigma_1^2 = 1$, $\sigma_2^2 = 0.8$, $\sigma_3^2 = 0.9$, $\sigma_4^2 = 1.1$ e $\sigma_5^2 = 1.2$.

Para as estruturas de covariância Uniforme, Bandas e AR(1), será adotado o valor 1 para σ^2 . Já ρ será de 0.5 para as estruturas Uniforme e AR(1), e $\rho = 0.25$ para Bandas, isso deve ao fato de que a estrutura de covariância de Bandas só é positiva definida quando $|\rho| \leq 0.5$.

Tabela 5.1 *Valores dos parâmetros populacionais, Vetor de médias, e dos parâmetros das Curvas de Crescimento usados na simulação.*

Curvas de Crescimento	Parâmetros				Vetor de Médias				
	α_0	α_1	α_2	α_3	μ_1	μ_2	μ_3	μ_4	μ_5
Linear	-1	1			0	1	2	3	4
Quadrática	-1.8	2	-0.2		0	1.4	2.4	3	3.2
Logística	4	2	3.5		0.03	0.19	1.08	2.92	3.81
Gompertz	0	5	1	3	0.01	0.33	1.84	3.46	4.37

5.1.1 Resultados das simulações

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

Nas Tabelas 5.2 a 5.5 estão as estimativas médias dos parâmetros das curvas de crescimento, das médias populacionais e os desvios-padrões, entre parênteses. Em cada tabela foi adotada uma estrutura de covariância. No Apêndice A, Tabelas A.2 e A.3, estão as estimativas médias dos parâmetros das matrizes estudadas no trabalho.

Observamos nas Tabelas 5.2 a 5.5, A.2 e A.3, que as estimativas médias dos parâmetros das curvas de crescimento, das médias e dos parâmetros das matrizes apresentaram valores muito próximos dos valores verdadeiros, indicando uma eficiência do processo de estimação. Nota-se, ainda, que não há nenhuma predominância em termos de qualidade das estimativas para um particular tipo de estrutura de covariância.

Observamos ainda, que os desvios-padrões das estimativas são muito pequenos, indicando uma grande estabilidade do processo de estimação.

Tabela 5.2 *Resultados da simulação, usando a matriz Diagonal*

	μ_1	μ_2	μ_3	μ_4	μ_5	α_0	α_1	α_2	α_3
Linear	-0.0032 (0.0106)	0.9981 (0.0073)	1.9994 (0.0063)	3.0008 (0.0084)	4.0021 (0.0121)	-1.0045 (0.0147)	1.0013 (0.0047)		
Quadrática	-0.0018 (0.0125)	1.3981 (0.0091)	2.3990 (0.0102)	3.0009 (0.0104)	3.2039 (0.0237)	-1.8006 (0.0317)	1.9983 (0.0269)	-0.1995 (0.0049)	
Logística	0.0275 (0.0028)	0.1919 (0.0115)	1.0766 (0.0228)	2.9201 (0.0162)	3.8132 (0.0171)	4.0078 (0.0285)	1.9900 (0.0409)	3.5035 (0.0136)	
Gompertz	0.0068 (0.0129)	0.3294 (0.0151)	1.8361 (0.0150)	3.4621 (0.0139)	4.3718 (0.0147)	0.0037 (0.0132)	5.0041 (0.0526)	1.0009 (0.0216)	3.0047 (0.0141)

Tabela 5.3 *Resultados da simulação, usando a matriz Uniforme*

	μ_1	μ_2	μ_3	μ_4	μ_5	α_0	α_1	α_2	α_3
Linear	-0.0027 (0.0402)	0.9994 (0.0345)	2.0016 (0.0295)	3.0037 (0.0256)	4.0058 (0.0236)	-1.0049 (0.0463)	1.0021 (0.0074)		
Quadrática	-0.0196 (0.1884)	1.4015 (0.0699)	2.3724 (0.2614)	2.9819 (0.2540)	3.1986 (0.2340)	-1.7927 (0.0680)	1.9880 (0.0332)	-0.1970 (0.0052)	
Logística	0.0305 (0.0034)	0.2049 (0.0134)	1.1033 (0.0239)	2.9259 (0.0189)	3.8214 (0.0195)	4.0256 (0.0341)	1.9536 (0.0469)	3.4989 (0.0135)	
Gompertz	0.0042 (0.0448)	0.3443 (0.0343)	1.8437 (0.0256)	3.4552 (0.0256)	4.3700 (0.0242)	0.0001 (0.0453)	5.0294 (0.0711)	0.9839 (0.0267)	3.0037 (0.0148)

Tabela 5.4 *Resultados da simulação, usando a matriz Bandas*

	μ_1	μ_2	μ_3	μ_4	μ_5	α_0	α_1	α_2	α_3
Linear	0.0050 (0.0332)	1.0042 (0.0237)	2.0033 (0.0185)	3.0025 (0.0209)	4.0016 (0.0292)	-0.9942 (0.0440)	0.9992 (0.0126)		
Quadrática	0.0006 (0.0395)	1.3894 (0.0253)	2.3910 (0.0248)	3.0055 (0.0224)	3.2329 (0.0315)	-1.7752 (0.0774)	1.9694 (0.0530)	-0.1936 (0.0083)	
Logística	0.0310 (0.0044)	0.2073 (0.0167)	1.1098 (0.0308)	2.9257 (0.0318)	3.8133 (0.0297)	4.0166 (0.0470)	1.9513 (0.0606)	3.4939 (0.0240)	
Gompertz	0.0097 (0.0388)	0.3488 (0.0338)	1.8428 (0.0308)	3.4502 (0.0290)	4.3633 (0.0331)	0.0054 (0.0398)	5.0194 (0.1149)	0.9840 (0.0433)	3.0052 (0.0303)

Tabela 5.5 *Resultados da simulação, usando a matriz AR(1)*

	μ_1	μ_2	μ_3	μ_4	μ_5	α_0	α_1	α_2	α_3
Linear	-0.0032 (0.0418)	0.9987 (0.0306)	2.0006 (0.0213)	3.0025 (0.0175)	4.0044 (0.0222)	-1.0051 (0.0536)	1.0019 (0.0129)		
Quadrática	-0.0032 (0.0482)	1.3929 (0.0357)	2.3957 (0.0290)	3.0053 (0.0226)	3.2218 (0.0256)	-1.7925 (0.0727)	1.9859 (0.0392)	-0.1966 (0.0060)	
Logística	0.0304 (0.0034)	0.2048 (0.0136)	1.1045 (0.0257)	2.9267 (0.0193)	3.8177 (0.0209)	4.0198 (0.0303)	1.9562 (0.0442)	3.4963 (0.0128)	
Gompertz	0.0055 (0.0427)	0.3443 (0.0353)	1.8449 (0.0216)	3.4551 (0.0175)	4.3663 (0.0196)	0.0015 (0.0430)	5.0181 (0.0742)	0.9865 (0.0246)	3.0016 (0.0164)

5.2 Aplicação em Dados Reais

Para a obtenção dos resultados 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.

No Apêndice A, Tabela A.4, serão mostrados os valores adotados para os parâmetros dos itens, obtidos em estudos prévios. Utilizou-se 1,7 para o fator de escada D, pois os parâmetros dos itens estavam na escala logística (Andrade, Tavares & Valle, 2000).

Para tornar os resultados comparáveis foram usados itens comuns, como mostra a Tabela 5.6.

Tabela 5.6 *Tabela de Itens Comuns*

	Teste 1	Teste 2	Teste 3	Teste 4	Teste 5	Teste 6
Teste 1	34	10	5	1	0	0
Teste 2	10	38	10	4	0	0
Teste 3	5	10	36	7	3	1
Teste 4	1	4	7	34	10	2
Teste 5	0	0	3	10	40	10
Teste 6	0	0	1	2	10	34

O Teste 2 utilizou 10 itens do Teste 1, o Teste 3 utilizou 5 e 5 itens dos Testes 1 e 2, respectivamente, o Teste 4 utilizou 1, 3 e 5 itens dos Testes 1, 2 e 3, o Teste 5 utilizou 3 e 7 itens dos testes 3 e 4, e o Teste 6 utilizou 1, 1 e 8 itens dos Testes 3, 4 e 5, respectivamente.

Nas Tabelas 5.8 a 5.11 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.

Relativamente à estrutura de covariância, na Tabela 5.7 apresentamos as covariâncias entre os escores nos seis instantes de avaliação. Nota-se claramente uma correlação entre os escores, certamente presente também nas proficiências individuais. Percebe-se um suave decaimento das correlações ao longo do tempo, menos acentuado que a estrutura AR(1), mas suficientemente claro a ponto de descartar as estruturas Diagonal e Bandas. No entanto, é possível que as diferentes estruturas tenham um impacto modesto nas estimativas dos parâmetros das curvas de crescimento e das proficiências médias.

Tabela 5.7 *Matriz de correlações observada*

	$Teste_1$	$Teste_2$	$Teste_3$	$Teste_4$	$Teste_5$	$Teste_6$
$Teste_1$	1.000	0.717	0.628	0.593	0.580	0.485
$Teste_2$	0.717	1.000	0.686	0.650	0.641	0.544
$Teste_3$	0.628	0.686	1.000	0.654	0.647	0.581
$Teste_4$	0.593	0.650	0.654	1.000	0.674	0.601
$Teste_5$	0.580	0.641	0.647	0.674	1.000	0.640
$Teste_6$	0.485	0.544	0.581	0.601	0.640	1.000

Tabela 5.8 *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 5.9 *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 5.10 *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 5.11 *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

Nota-se que, de forma geral, os parâmetros estão próximos, mas com alguma diferença identificável. A estrutura de Bandas foi a que mais se diferenciou, apresentando seus parâmetros das curvas de crescimento claramente diferente das demais. Isso pode ter ocorrido pela estabilidade do processo de estimação, visto que para valores altos de correlação essa matriz não é positiva definida. As estruturas Uniforme e Auto-Regressiva de ordem 1 são as que estão mais próximas quanto às estimativas dos parâmetros das curvas de crescimento, como seria de se esperar com base na Tabela 5.7.

Para identificarmos, dentre estas duas estruturas de covariância, qual curva de crescimento melhor se ajusta aos dados, utilizamos o critério AIC (*Akaike's Information Criterion*), obtido por

$$AIC = -2 \log L(\boldsymbol{\zeta}, \hat{\boldsymbol{\eta}}) + 2n_p,$$

onde a $L(\boldsymbol{\zeta}, \hat{\boldsymbol{\eta}})$ 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.

A Tabela 5.12 apresenta os principais valores de AIC para as duas estruturas, Uniforme e AR(1). Nota-se, de forma geral, que realmente 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

Tabela 5.12 *Log-verossimilhança associada a cada estrutura de covariância e curva*

Modelo	Curva	$-2 \times \text{Log-verossimilhança}$	Num. Parâmetros	AIC
Uniforme	Linear	499.598,2295	4	499.606,2295
	Quadrática	499.589,3954	5	499.599,3954
	Logística	499.533,4151	5	499.543,4151
	Gompertz	499.532,2438	6	499.544,2438
AR(1)	Linear	499.464,4986	4	499.472,4986
	Quadrática	499.460,5648	5	499.470,5648
	Logística	499.389,8326	5	499.399,8326
	Gompertz	499.389,5142	6	499.401,5142

das proficiências médias do grupo ao longo dos seis instantes de avaliação, com estimativas dos parâmetros da curva de crescimento dadas por $\hat{\alpha}_0 = 1.8115$, $\hat{\alpha}_1 = 0.0697$ e $\hat{\alpha}_2 = 21.2647$, e da estrutura de covariância $\hat{\rho} = 0.9000$ e $\hat{\sigma}^2 = 1.1690$. As estimativas médias geradas com esta curva de crescimento para os tempos adotados são dadas por $\hat{\mu}_4 = 0.4180$, $\hat{\mu}_{11} = 0.5947$, $\hat{\mu}_{23} = 0.9605$, $\hat{\mu}_{35} = 1.3092$, $\hat{\mu}_{47} = 1.5534$ e $\hat{\mu}_{59} = 1.6899$.

Conclusões e Sugestões

6.1 Conclusões

Neste trabalho foram propostos modelos para 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. Foram apresentadas as principais curvas constantes na literatura estatística, tais como: Linear, Quadrática, Logística e Gompertz.

Também foi abordado o fato de termos uma possível dependência entre os resultados nos vários instantes de avaliação, e propomos estruturas de covariância para modelar tal dependência, particularmente a Diagonal, Uniforme, Auto-Regressiva de ordem 1, e estrutura de Bandas.

Consideramos conhecidos os parâmetros dos itens envolvidos no estudo, pois o foco do trabalho foi a abordagem de curvas de crescimento com dependência. Foram feitas aplicações a dados simulados e os resultados foram bastante promissores, com um processo de estimação realmente satisfatório, tanto para os parâmetros das estruturas de dependência como das curvas de crescimento.

Também foi feita uma aplicação a dados reais, 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. Utilizou-se apenas o conjunto de indivíduos que participou de todas as condições de avaliação, totalizando 1.987 indivíduos. Os itens foram previamente calibrados, de forma que seus parâmetros já eram conhecidos e foram obtidos do Banco Nacional de Itens (BNI) do Inep/MEC.

Observou-se descritivamente que, em termos da dependência entre os escores dos estudantes, as estruturas teóricas que mais se aproximam da matriz de covariância dos escores foram a Uniforme e a AR(1). No entanto, a correlação na estrutura observada decaiu com o tempo, fugindo um pouco da Uniforme, mas decaiu lentamente comparada à AR(1).

Relativamente aos resultados da análise propriamente, verifica-se um bom ajuste da curva de crescimento logística, em geral esperada em dados educacionais, associada à estrutura auto-regressiva de ordem 1.

6.2 Sugestões para trabalhos futuros

Como comentado na seção anterior, a estrutura de dependência dos dados pode ser um pouco diferente daquelas usadas no trabalho, apresentando um certo decaimento, mas lento comparativamente à $AR(1)$. Sugere-se que novos estudos sejam realizados de forma a sugerir outras estruturas que apresentem uma taxa de decaimento diferente da $AR(1)$. Em termos de curvas de crescimento, sugere-se a aplicação a dados de outros estudos de forma a confirmar uma determinada forma funcional para dados educacionais, particularmente a forma logística, alterando apenas os parâmetros da curva, e que estes sirvam para identificar situações, áreas de estudos etc.

Sugere-se ainda a extensão da abordagem para situações em que os parâmetros dos itens também devam ser estimados, de forma identificar o impacto nas estimativas dos parâmetros das curvas de crescimento neste caso, bem como quando os itens não são dicotômicos.

Apêndice A

A.1 Parâmetros de Itens - Simulações

Tabela A.1 *Valores para os parâmetros dos itens usados na simulação*

Item	Teste	a_i	b_i	c_i	Item	Teste	a_i	b_i	c_i	Item	Teste	a_i	b_i	c_i
1	1	0.6	-0.7	0.2	33	2	1.4	1.1	0.2	65	4	1.0	3.0	0.2
2	1	1.0	-0.7	0.2	34	2	0.6	1.3	0.2	66	4	1.4	3.0	0.2
3	1	1.4	-0.7	0.2	35	2	0.6	1.5	0.2	67	4	0.6	3.1	0.2
4	1	0.6	-0.5	0.2	36	2	0.6	1.7	0.2	68	4	1.0	3.1	0.2
5	1	1.0	-0.5	0.2	37	2,3	1.0	1.3	0.2	69	4	1.4	3.1	0.2
6	1	1.4	-0.5	0.2	38	2,3	1.0	1.5	0.2	70	4	0.6	3.3	0.2
7	1	0.6	-0.3	0.2	39	2,3	1.0	1.7	0.2	71	4	0.6	3.5	0.2
8	1	1.0	-0.3	0.2	40	2,3	1.4	1.3	0.2	72	4	0.6	3.7	0.2
9	1	1.4	-0.3	0.2	41	2,3	1.4	1.5	0.2	73	4,5	1.0	3.3	0.2
10	1	0.6	-0.1	0.2	42	2,3	1.4	1.7	0.2	74	4,5	1.0	3.5	0.2
11	1	1.0	-0.1	0.2	43	3	0.6	1.9	0.2	75	4,5	1.0	3.7	0.2
12	1	1.4	-0.1	0.2	44	3	1.0	1.9	0.2	76	4,5	1.4	3.3	0.2
13	1	0.6	0.1	0.2	45	3	1.4	1.9	0.2	77	4,5	1.4	3.5	0.2
14	1	1.0	0.1	0.2	46	3	0.6	2.0	0.2	78	4,5	1.4	3.7	0.2
15	1	1.4	0.1	0.2	47	3	1.0	2.0	0.2	79	5	0.6	3.9	0.2
16	1	0.6	0.3	0.2	48	3	1.4	2.0	0.2	80	5	1.0	3.9	0.2
17	1	0.6	0.5	0.2	49	3	0.6	2.1	0.2	81	5	1.4	3.9	0.2
18	1	0.6	0.7	0.2	50	3	1.0	2.1	0.2	82	5	0.6	4.0	0.2
19	1,2	1.0	0.3	0.2	51	3	1.4	2.1	0.2	83	5	1.0	4.0	0.2
20	1,2	1.0	0.5	0.2	52	3	0.6	2.3	0.2	84	5	1.4	4.0	0.2
21	1,2	1.0	0.7	0.2	53	3	0.6	2.5	0.2	85	5	0.6	4.1	0.2
22	1,2	1.4	0.3	0.2	54	3	0.6	2.7	0.2	86	5	1.0	4.1	0.2
23	1,2	1.4	0.5	0.2	55	3,4	1.0	2.3	0.2	87	5	1.4	4.1	0.2
24	1,2	1.4	0.7	0.2	56	3,4	1.0	2.5	0.2	88	5	0.6	4.3	0.2
25	2	0.6	0.9	0.2	57	3,4	1.0	2.7	0.2	89	5	0.6	4.5	0.2
26	2	1.0	0.9	0.2	58	3,4	1.4	2.3	0.2	90	5	0.6	4.7	0.2
27	2	1.4	0.9	0.2	59	3,4	1.4	2.5	0.2	91	5	1.0	4.3	0.2
28	2	0.6	1.0	0.2	60	3,4	1.4	2.7	0.2	92	5	1.0	4.5	0.2
29	2	1.0	1.0	0.2	61	4	0.6	2.9	0.2	93	5	1.0	4.7	0.2
30	2	1.4	1.0	0.2	62	4	1.0	2.9	0.2	94	5	1.4	4.3	0.2
31	2	0.6	1.1	0.2	63	4	1.4	2.9	0.2	95	5	1.4	4.5	0.2
32	2	1.0	1.1	0.2	64	4	0.6	3.0	0.2	96	5	1.4	4.7	0.2

Tabela A.2 *Estimativas dos parâmetros da matriz Diagonal obtidos na simulação*

Curvas	Parametros Populacionais (Σ)				
	σ_1^2	σ_2^2	σ_3^2	σ_4^2	σ_5^2
Linear	1.0013 (0.0438)	0.8026 (0.0298)	0.9009 (0.0395)	1.1052 (0.0446)	1.2028 (0.0521)
Quadrática	1.0011 (0.0438)	0.8026 (0.0324)	0.8984 (0.0420)	1.1052 (0.0447)	1.2020 (0.0671)
Logística	1.0016 (0.0442)	0.8016 (0.0401)	0.9055 (0.0565)	1.1023 (0.0453)	1.1977 (0.0574)
Gompertz	1.0018 (0.0444)	0.8037 (0.0411)	0.9014 (0.0439)	1.1070 (0.0534)	1.2019 (0.0563)

Tabela A.3 *Estimativas dos parâmetros das matrizes Uniforme, Bandas e AR(1), obtidos na simulação*

Curvas	Uniforme		Bandas		AR(1)	
	ρ	σ^2	ρ	σ^2	ρ	σ^2
Linear	0.4840 (0.0111)	0.9095 (0.0220)	0.2864 (0.0151)	0.8764 (0.0190)	0.5024 (0.0112)	0.9121 (0.0205)
Quadrática	0.4823 (0.0111)	0.9082 (0.0231)	0.2825 (0.0154)	0.8794 (0.0193)	0.4985 (0.0124)	0.9128 (0.0197)
Logística	0.4797 (0.0111)	0.9120 (0.0237)	0.2837 (0.0181)	0.8898 (0.0218)	0.4988 (0.0115)	0.9187 (0.0204)
Gompertz	0.4833 (0.0105)	0.9168 (0.0220)	0.2849 (0.0152)	0.8853 (0.0186)	0.5026 (0.0097)	0.9220 (0.0203)

A.2

Tabela A.4 *Valores adotados para os parâmetros dos itens*

Item	Teste	a	b	c	Item	Teste	a	b	c	Item	Teste	a	b	c
101	1	0.78	0.84	0.35	229	2	1.23	-0.55	0.21	440	4	1.44	2.48	0.26
102	1	0.68	0.22	0.22	230	2 e 3	1.46	1.48	0.27	501	5 e 6	0.58	0.53	0.38
103	1	1.26	0.37	0.35	233	2	1.05	0.34	0.24	502	5	0.93	1.83	0.17
104	1 e 2	1.12	0.54	0.34	234	2	1.29	0.06	0.10	503	5	0.57	1.53	0.23
105	1	0.49	-2.57	0.15	236		1.16	-0.16	0.13	504	5	0.78	2.27	0.50
106	1	0.57	-0.70	0.08	240		0.88	-0.10	0.17	506	5	1.56	2.14	0.17
108	1	0.49	-0.83	0.16	302	3	0.49	1.83	0.15	507	5	1.18	2.60	0.32
109	1, 2 e 3	1.17	0.01	0.31	303	3, 4, 5 e 6	0.57	-0.29	0.02	508	5	0.98	2.77	0.23
110	1 e 2	1.07	0.25	0.15	308	3	0.96	1.67	0.19	511	5 e 6	0.84	2.93	0.36
111	1	0.70	-0.80	0.09	309	3, 4 e 5	0.63	0.12	0.11	513	5	0.78	1.33	0.29
112	1	0.43	-0.84	0.19	311	3	0.50	1.70	0.15	514	5	0.68	2.03	0.25
114	1, 2 e 3	0.75	0.39	0.17	312	3, 4 e 5	1.01	1.01	0.30	515	5 e 6	1.24	3.17	0.37
115	1	0.45	-0.40	0.20	315	3	1.06	1.46	0.24	517	5	1.00	1.57	0.22
116	1	0.76	-0.62	0.12	316	3	1.10	-0.07	0.23	518	5	1.50	3.27	0.18
117	1	0.66	0.52	0.07	317	3	1.05	1.04	0.21	519	5	0.54	2.11	0.19
118	1	0.64	0.02	0.14	318	3	1.07	1.44	0.25	521	5	0.93	1.40	0.38
119	1, 2, 3 e 4	0.80	0.57	0.29	321	3	0.52	0.93	0.12	522	5	1.08	2.89	0.28
120	1	0.55	0.03	0.07	323	3	0.44	0.42	0.08	523	5	1.61	2.12	0.15
121	1	0.77	-1.03	0.06	324	3 e 4	0.91	0.37	0.17	526	5 e 6	1.06	2.50	0.24
122	1 e 2	1.08	0.66	0.11	326	3	0.61	0.45	0.08	527	5 e 6	0.45	3.86	0.25
123	1	0.56	0.17	0.22	327	3	1.42	1.91	0.22	528	5	2.20	2.73	0.19
124	1, 2 e 3	1.10	0.62	0.16	329	3	1.17	2.19	0.28	530	5	1.23	2.69	0.18
125	1	0.97	-0.33	0.15	330	3	1.34	1.32	0.26	531	5	1.28	2.47	0.19
126	1 e 2	0.95	0.12	0.11	331	3	0.71	0.03	0.08	532	5 e 6	0.98	1.63	0.45
127	1	0.79	0.60	0.10	332	3	1.10	1.05	0.13	534	5 e 6	0.94	3.69	0.39
128	1, 2 e 3	1.36	0.67	0.21	334	3 e 4	0.72	0.27	0.12	535	5	1.10	1.99	0.19
129	1	1.04	0.32	0.08	335	3	0.91	0.86	0.13	536	5	0.52	2.86	0.19
130	1	0.43	0.46	0.11	336	3	0.70	1.29	0.21	537	5 e 6	1.44	2.77	0.21
131	1	0.42	1.16	0.02	337	3	1.21	1.43	0.13	538	5	1.04	1.38	0.12
132	1	1.94	1.00	0.35	338	3	0.84	0.75	0.20	539	5	0.91	2.33	0.12
133	1	0.75	1.97	0.13	339	3	1.26	1.29	0.19	540	5	2.21	2.78	0.20
134	1 e 2	0.77	-0.99	0.07	340	3	0.95	1.63	0.20	602	6	0.55	2.78	0.34
135	1	1.07	-0.09	0.05	401	4	0.74	-0.21	0.17	603	6	0.52	4.22	0.13
136	1	0.78	0.80	0.25	403	4 e 5	1.06	1.70	0.23	607	6	0.41	1.19	0.14
201	2	1.37	2.35	0.20	406	4 e 5	1.02	1.90	0.17	608	6	0.82	2.93	0.19
203	2	0.74	0.74	0.17	408	4	0.95	2.38	0.22	613	6	0.43	2.66	0.12
204	2 e 3	0.53	-0.63	0.11	409	4	1.78	2.41	0.16	616	6	1.00	2.07	0.26
207	2	1.24	0.72	0.18	418	4	0.72	1.97	0.29	618	6	0.69	3.49	0.21
208	2	1.55	1.44	0.10	419	4	1.38	2.80	0.14	619	6	0.64	-0.61	0.10
209	2	0.87	-0.27	0.14	420	4 e 5	1.42	1.63	0.15	621	6	0.54	-0.07	0.08
210	2	0.78	0.37	0.17	422	4 e 5	1.01	1.82	0.29	622	6	0.93	3.14	0.31
212	2	1.00	1.85	0.26	423	4 e 5	0.70	2.05	0.18	623	6	1.08	2.50	0.31
213	2	0.81	0.53	0.32	424	4	1.08	2.27	0.22	626	6	1.13	2.07	0.20
214	2	1.49	1.56	0.14	425	4	0.74	1.82	0.20	627	6	0.72	3.61	0.24
215	2	1.51	1.47	0.22	427	4	1.75	2.06	0.21	628	6	0.72	4.59	0.28
216	2	0.87	-1.05	0.05	428	4	0.87	0.67	0.05	629	6	1.04	2.81	0.29
218	2 e 3	0.97	-0.16	0.17	430	4	2.17	2.74	0.17	630	6	0.81	4.12	0.30
219	2, 3 e 4	0.93	0.88	0.27	431	4	1.44	2.37	0.17	632	6	0.60	2.04	0.13
220	2	1.41	1.50	0.18	432	4, 5 e 6	1.29	1.96	0.27	633	6	0.45	4.74	0.28
221	2	0.71	0.56	0.17	433	4	0.83	1.23	0.18	634	6	0.40	3.46	0.22
222	2	0.89	-0.82	0.05	434	4 e 5	1.06	2.14	0.14	635	6	1.16	4.26	0.20
223	2 e 3	1.50	0.87	0.19	435	4	1.29	2.97	0.29	636	6	0.85	2.02	0.33
225	2	0.97	-0.96	0.12	436	4	1.20	1.84	0.31	638	6	1.01	3.77	0.26
226	2	0.63	0.25	0.16	437	4	1.98	2.81	0.12	639	6	0.87	2.39	0.24
227	2	1.15	1.30	0.19	438	4	1.89	2.35	0.26	640	6	0.69	0.10	0.11
228	2	2.00	1.10	0.25	439	4	0.67	3.59	0.21					

Referências Bibliográficas

- [1] Andrade, D. F. (1999). *Comparando o Desempenho de Grupos (Populações) de Respondentes Através da Teoria da Resposta ao Item*. Tese apresentada ao Departamento de Estatística e Matemática Aplicada da UFC para o concurso de professor titular.
- [2] Andrade, D.F, Tavares, H.R., Valle, R.C. (2000). *Teoria da Resposta ao Item: Conceitos e Aplicações. Associação Brasileira de Estatística: São Paulo*.
- [3] Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In F.M. Lord & M.R. Novick. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.
- [4] Bock, R. D, Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of a EM algorithm. *Psychometrika*, 46, 433-459.
- [5] Bock, R. D, Lieberman, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika*, 35, 179-197.
- [6] Bock, R. D, Zimowski, M. F. (1997). Multiple Group IRT. In *Handbook of Modern Item Response Theory*. W.J. van der Linder e R.K. Hambleton Eds. New York: Springer-Verlag.
- [7] Dempster, A. P. , Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B*, 39, 1-38.
- [8] 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.
- [9] Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park: Sage Publications.

-
- [10] Hedges, L. V, Vevea, J. L. (1997). A study of equating in NAEP. *Paper presented at The NAEP Validity Studies Panel*. Palo Alto : American Institutes for Research.
- [11] Hildebrand, F. B. (1956). *Introduction to Numerical Analysis*. New-York: McGraw-Hill.
- [12] Issac, E, Keller, H. B. (1966). *Analysis of Numerical Methods*. New York: Wiley & Sons.
- [13] Jones, R. H. (1993). *Longitudinal Data with Serial Correlation: A State-Space Approach*. Chapman and Hall: London.
- [14] Linden, W. J. van der, Hambleton, R. K. (1997). *Handbook of Modern Item Response Theory*. New York : Springer-Verlag.
- [15] Lord, F. M. (1952). A theory of test scores (No. 7). *Psychometric Monograph*.
- [16] Mislevy, R. J. (1986). Bayes modal estimation in item response models. *Psychometrika*, 51, 177-195.
- [17] Mislevy, R. J. and Bock, R. D. (1990). *BILOG 3 : Item Analysis and Test Scoring with Binary Logistic Models*. Chicago : Scientific Software, Inc.
- [18] Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. New York: Wiley & Sons.
- [19] Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen : Danish Institute for Educational Research.
- [20] Tavares, H.R. (2001). *Teoria da Resposta ao Item para Dados Longitudinais*. Tese de Doutorado (IME/USP).
- [21] Wright, B. D. (1968). *Sample-free test calibration and person measurement*. Proceedings of the 1967 Invitational Conference on Testing Problems. Princeton, N. J. : Educational Testing Service.