UNIVERSIDADE FEDERAL DO PARÁ

INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA

ESCALA DE PROFICIÊNCIA PARA O ENEM UTILIZANDO TEORIA DA RESPOSTA AO ITEM

Francisco Fialho Guedes Ferreira

Orientador: Prof. Dr. Héliton Ribeiro Tavares

Belém
2009

Francisco Fialho Guedes Ferreira

ESCALA DE PROFICIÊNCIA PARA O ENEM UTILIZANDO TEORIA DA 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. Héliton Ribeiro Tavares

Belém
2009

Francisco Fialho Guedes Ferreira

ESCALA DE PROFICIÊNCIA PARA O ENEM UTILIZANDO TEORIA DA 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, 06 de Fevereiro de 2009

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

Banca Examinadora

Prof. Dr. Héliton Ribeiro Tavares
Universidade Federal do Pará
Orientador

Prof. Dr. Joaquim José Soares Neto
Universidade de Brasília
Examinador

Prof. Dr. Joaquim Carlos Barbosa Queiroz
Universidade Federal do Pará
Examinador

A minha mãe Rita Guedes Ferreira,
por tudo que me ensinou.

Aos meus irmãos Francisco de Assis Guedes Ferreira (Gotão), Francisco Idelfonso Guedes Ferreira(Ruben), Francisco Eudes Guedes Ferreira e Semária Idelfonso Guedes Ferreira, pelo apoio.

E em especial a meu pai Raimundo Idelfonso Ferreira e aos meus irmãos Francisco Guedes Ferreira (Nico) e Francimar Guedes Ferreira (in memoriam),
que na dimensão onde se encontram, nos acompanha
e protege nessa caminhada.

AGRADECIMENTOS

⋆ À Deus Todo Poderoso, que me concedeu a vida e a força necessária para chegar ao final deste trabalho;

⋆ À minha esposa Silvia Kátia, pelo insentivo, dedicação, compreensão e amor;

⋆ Aos meus filhos Ramon, Jullieth e Phellipe, por compreender a minha ausência em alguns momentos de suas vidas;

⋆ Ao meu orientador, Prof. Dr. Heliton Ribeiro Tavares, pela atenção e paciência;

⋆ À Profa Dr. Regina Madruga Tavares pelo ensino oferecido;

⋆ À Universidade Federal do Pará;

⋆ À Coordenação do Curso de Pós-Graduação em Matemàtica e Estatística, representado pelo Prof. Dr. Mauro Lima;

⋆ Ao INEP pela colaboração;

⋆ Aos amigos do Cespe-UNB pela ajuda;

⋆ Aos amigos da Secretaria de Educação do Estado da Bahia;

⋆ Aos meus colegas de curso;

⋆ À equipe de professores do C. E. F. M. Isaura Amorim a qual faço parte;

⋆ À prefeitura municipal de Cidelândia;

⋆ À Secretaria de Educação do Estado do Maranhão;

⋆ Finalmente, a todos, que de alguma forma contribuíram para a realização deste trabalho.

"Quem passou pela vida em branca nuvem
E em plácido repouso adormeceu
Quem não sentiu o frio da desgraça
Quem passou pela vida e não sofreu;
Foi espectro de homem - não foi homem
Só passou pela vida não viveu".

Francisco Octaviano

RESUMO

FERREIRA, Francisco Fialho Guedes. Escala de Proficiência para o ENEM utilizando Teoria da Resposta ao Item. Dissertação de Mestrado - (Mestrado em Matemática e Estatística - UFPA, Belém - PA, Brasil).

O presente trabalho tem como objetivo criar uma escala de proficiências para o ENEM e colocá-lo numa mesma métrica ao longo de suas edições de 1998 a 2008, utilizando a Teoria de Resposta ao Item. Essa teoria originou-se na metade da década de 30, mas só recentemente está sendo utilizada em situações práticas na área de educação no Brasil. Vários são os modelos de resposta ao item, sendo que os mais utilizados são os que medem apenas um traço latente representados por uma única dimensão. Já o ENEM, em princípio, foi construído com a finalidade de medir vários traços latentes posto que a estrutura do teste é composta por uma matriz de 5 competências. Primeiramente, este trabalho apresenta uma visão geral do ENEM e em seguida um brve comentário sobre a Teoria Clássica de Testes apontando suas limitações. Para a verificação da dimensionalidade do teste, foi utilizada a Análise Fatorial de Informação Plena. Como critério para a determinação da dimensionalidade foram utilizadas a variação explicada por cada fator. O resultado desse estudo mostrou que o ENEM pode ser bem aproximado por um teste unidimensional que mede um constructo* bastante geral, que denominaremos apenas por constructo ENEM. Finalmente, são apresentados e comentados os resultados desse constructo.

Palavras Chave: Enem, escala e proficiências.

* Construção puramente mental, criada a partir de elementos mais simples para ser parte de uma teoria.

ABSTRACT

FERREIRA, Francisco Fialho Guedes. Proficiency Scale for ENEM using Item Response Theory. Dissertation Masters - (Masters in Mathematics and Statistics - UFPA, Belém - PA, Brazil).

This paper aims create a scale of proficiency and to place it ENEM in the same metric over its editions from 1998 to 2008, using the Theory of Item Response. This theory was originated in the mid-1930, but only recently has been used in practical situations in the field of education in Brazil. There are several models of response to the item, and the most used are those which measure only a latent trait represented by a single dimension. The ENEM, in principle, was built for the purpose of measuring several latent traits, since the structure of the test consists of a matrix of 5 (five) skills. First, this work presents an overview of ENEM, and then a brief comment on the Classical Test Theory, pointing out its limitations. In order to verify the dimensionality of the search was used as a Factor Analysis of Full Information. As a criterion for determining the dimensionality was used the variation explained by each factor. This research result showed that the ENEM can be well approximated by a one-dimensional test that measures a construct (Construction purely mental, created from simple elements to be part of a theory) fairly general, which call for only construct ENEM. Finally, are presented and commented on the results of this construct.

Words Keys: Enem, scale and proficiency.

SUMÁRIO

LISTA DE TABELAS

LISTA DE FIGURAS

Capítulo 1

Introdução

1.1 Aspectos Gerais

O Exame Nacional do Ensino Médio (ENEM), criado em 1998 pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), do Ministério da Educação, é um exame individual e de caráter voluntário, oferecido anualmente aos concluintes e egressos do ensino médio, com o objetivo principal de possibilitar uma referência para auto-avaliação, a partir das competências e habilidades que o estruturam. Além disso, ele serve como modalidade alternativa ou complementar aos processos de seleção para o acesso ao ensino superior e ao mercado de trabalho. Realizado anualmente, ele se constitui um valioso instrumento de avaliação, fornecendo uma imagem realista e sempre atualizada da educação no Brasil. O modelo de avaliação do Enem foi desenvolvido com ênfase na aferição das estruturas mentais com as quais construímos continuamente o conhecimento e não apenas na memória, que, importantíssima na constituição dessas estruturas, sozinha não consegue fazer-nos capazes de compreender o mundo em que vivemos (Eixos cognitivos do ENEM 2004).

O ENEM é estruturado por uma matriz de competências que define claramente os pressupostos do exame e delineia suas características operacionais. O modelo da matriz contempla a indicação das competências gerais próprias do aluno, na fase de desenvolvimento cognitivo correspondente ao término da escolaridade básica, associadas aos conteúdos do ensino fundamental e médio.

As cinco competências tratadas pelo ENEM são:

1. Dominar linguagens (DL);
2. Compreender fenômenos (CF);
3. Enfrentar situações-problema (ES);

4. Construir argumentação (CA);

5. Elaborar propostas (EP).

Cada uma dessas competências pode contemplar até 21 habilidades que estão enumeradas a seguir:

1. Dada a descrição discursiva ou por ilustração de um experimento ou fenômeno, de natureza científica, tecnológica ou social, "identificar"variáveis relevantes e "selecionar"os instrumentos necessários para realização ou interpretação do mesmo;

2. Em um gráfico cartesiano de variável socioeconômica ou técnico científico, "identificar"e "analisar"valores das variáveis, intervalos de crescimento ou decréscimo e taxas de variação;

3. Dada uma distribuição estatística de variável social, econômica, física, química ou biológica, traduzir e interpretar as informações disponíveis,ou reorganizá-las, objetivando interpolações ou extrapolações;

4. Dada uma situação-problema, apresentada em uma linguagem de determinada área de conhecimento, "relacioná-la"com sua formulação em outras linguagens ou vice-versa;

5. A partir da leitura de textos literários consagrados e de informações sobre concepções artísticas, estabelecer relações entre eles e seu contexto histórico, social, político ou cultural, inferindo as escolhas dos temas, gêneros discursivos e recursos expressivos dos autores;

6. Com base em um texto, analisar as funções da linguagem, identificar marcas de variantes lingüísticas de natureza sociocultural, regional, de registro ou de estilo, e explorar as relações entre as linguagens coloquial e formal;

7. "Identificar"e "caracterizar"a conservação e as transformações de energia em diferentes processos de sua geração e uso social, e "comparar"diferentes recursos e opções energéticas;

8. Analisar criticamente, de forma qualitativa ou quantitativa, as implicações ambientais, sociais e econômicas, dos processos de utilização dos recursos naturais, materiais ou energéticos;

9. "Compreende" o significado e a importância da água e de seu ciclo para a manutenção da vida, em sua "relação" com condições socioambientais, sabendo "quantificar" variações de temperatura e mudanças de fase em processos naturais e de "intervenção" humana;

10. "Utilizar" e "interpretar" diferentes escalas de tempo para "situar" e "descrever" transformações na atmosfera, biosfera, hidrosfera e litosfera, origem e evolução da vida, variações populacionais e modificações no espaço geográfico;

11. Diante da diversidade da vida, analisar, do ponto de vista biológico, físico ou químico, padrões comuns nas estruturas e nos processos que garantem continuidade e a evolução dos seres vivos;

12. "Analisar" fatores socioeconômicos e ambientais associados ao desenvolvimento, às condições de vida e saúde de populações humanas, por meio da "interpretação" de diferentes indicadores;

13. Compreender o caráter sistêmico do planeta e reconhecer a importância da biodiversidade para preservação da vida, relacionando condições do meio e intervenção humana;

14. Diante da diversidade de formas geométricas planas e espaciais, presentes na natureza ou imaginadas, "caracterizá-las" por meio de propriedades, "relacionar" seus elementos, "calcular" comprimentos, áreas ou volumes, e "utilizar" o conhecimento geométrico para "leitura", "compreensão" e "ação" sobre a realidade;

15. "Reconhecer" o caráter aleatório de fenômenos naturais ou não e "utilizar" em situações-problema processos de contagem, representação de freqüências relativas, construção de espaços amostrais, distribuição e cálculo de probabilidades;

16. "Analisar", de forma qualitativa ou quantitativa, situações-problema referentes a perturbações ambientais, identificando fonte, transporte e destino dos poluentes, "reconhecendo" suas transformações; "prever" efeitos nos ecossistemas e no sistema produtivo e "propor" formas de intervenção para reduzir e controlar os efeitos da poluição ambiental;

17. Na obtenção e produção de materiais e de insumos energéticos, "identificar" etapas,

"calcular" rendimentos, taxas e índices, e "analisar" implicações sociais, econômicas e ambientais;

18. Valorizar a diversidade dos patrimônios etnoculturais e artísticos, identificando-a nas suas manifestações e representações em diferentes sociedades, épocas e lugares;

19. "Confrontar" interpretações diversas de situações ou fatos de natureza histórico-geográfica, técnico-científica, artístico-cultural ou do cotidiano, "comparando" diferentes pontos de vista, "identificando" os pressupostos de cada interpretação e "analisando" a validade dos argumentos utilizados;

20. Comparar processos de formação socioeconômica, relacionando-os com seu contexto histórico e geográfico;

21. Dado um conjunto de informações sobre uma realidade histórico-geográfica, "contextualizar" e "ordenar" os eventos registrados, "compreendendo" a importância dos fatores sociais, econômicos, políticos ou culturais.

Estas 21 habilidades se combinam compondo as cinco competências de acordo com a estrutura apresentada abaixo.

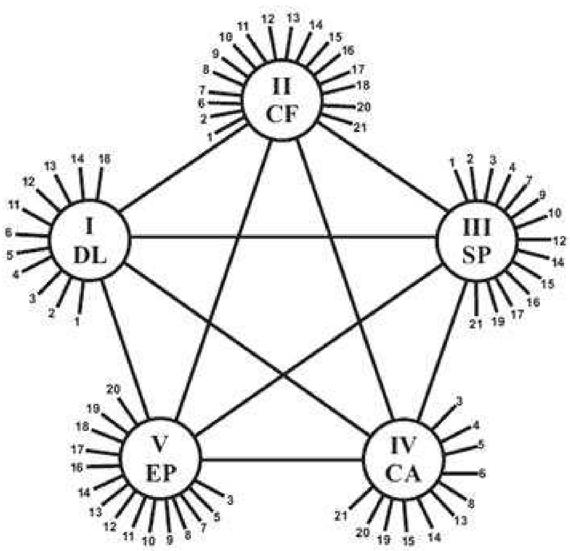

Figura 1.1 *Diagrama das cinco competências e 21 habilidades.*

É importante observar que cada item (questão), componente de qualquer edição do ENEM é elaborado de forma a contemplar até cinco competências, o que significa que existe a possibilidade da existência de uma variação quanto ao número de parâmetros a serem estimados.

Dado que o ENEM é uma avaliação que vem ocorrendo desde 98, na qual os testes são compostos por itens totalmente distintos e não calibrados em cada ano, torna-se impossibilitada a comparação do desempenho dos alunos no decorrer dos anos.

Desde o início de sua realização, os resultados do ENEM veem sendo construídos com a utilização da Teoria Clássica dos Testes, que apresenta algumas deficiências que serão mencionadas mais adiante. Portanto, com o intuito de resolver o problema de comparabilidade, uma prova de ligação (link) foi plicada para um grupo de alunos do Ensino Médio com itens de todos os anos do ENEM, de 1998 a 2008, a partir do qual cria-se uma estrutura de ligação necessária para que possamos considerar uma avaliação geral única, de forma que todas as proficiências individuais estimadas estejam na mesma escala, portanto comparáveis, o que também acontecerá com a proficiência média de cada edição do ENEM.

1.2 Justificativa e Importância do Trabalho

A justificativa para à realização deste trabalho deve-se ao fato de que este instrumento avaliativo (ENEM) não compara desempenho no decorrer dos anos, portanto, com a criação de uma escala e a colocação de todas as edições do ENEM numa mesma métrica pode-se ter resultados relevantes no que diz respeito à comparabilidade.

Além disso, é importante mencionar que a realização deste trabalho se justifica também pelas contribuições na área da Teoria da Resposta ao Item, observando sua relevância em muitos aspectos. Sob este ponto de vista, Andrade, Tavares & Vale (2000) afirmam que o uso da teoria da Resposta ao Item é uma alternativa vantajosa em relação aos métodos clássicos, pois é livre de certos paradoxos ou violações de princípios, que estão associados com a Teoria Clássica.

1.3 A Hipótese Básica do Trabalho

Parte-se da hipótese de que as estatísticas que se tem atualmente com relação ao ENEM não são satisfatórias para fazer comparação do desempenho desses grupos no decorrer dos anos.

1.4 Objetivos

1.4.1 Objetivo Geral

O objetivo deste trabalho é criar uma escala de proficiências para o ENEM e, considerando todas as suas edições, colocá-lo numa mesma métrica (escala), que possibilite a comparação destas edições.

1.4.2 Objetivos Específicos

Como objetivos específicos podem-se enumerar:

- Adotar uma prova de ligação de modo que itens em comum conectem todas as edições do exame, com o intuito de possibilitar a utilização da TRI;
- Utilizar a metodologia da TRI para encontrar estimativas de interesse; parâmetros dos itens e das proficiências;
- Utilizar a Analise Fatorial de Informação Plena (AFIP) para verificar a dimensionalidade dos itens e consequentemente dos testes;
- Através dos resultados obtidos com a Análise Fatorial de Informação Plena, verificar a adequação aos modelos Unidimensional ou Multidimensional;
- Aplicar o modelo Unidimensional para a obtenção das estimativas ótimas;
- Utilizar as proficiências dos indivíduos já numa mesma métrica para verificar a evolução dos grupos.

1.5 Estrutura do Trabalho

Este trabalho encontra-se dividido em oito capítulos, a saber:

- Capítulo 1: refere-se à introdução do trabalho, onde estão englobados os aspectos gerais, a justificativa e importância do trabalho, os objetivos geral e específico e a hipótese do trabalho.
- Capítulo 2: Um rápido comentário sobre Teoria Clássica de Medidas e um breve histórico da TRI;
- Capítulo 3: Este foi dedicado à revisão de alguns conceitos da TRI unidimensional para uma e/ou duas ou mais populações;
- Capítulo 4: Trata da estimação dos parâmetros dos itens e das proficiências dos respondentes para uma, duas ou mais populações;
- Capítulo 5: Uma pequena abordagem sobre Equalização e Escala;
- Capítulo 6: Uma breve abordagem sobre a Análise Fatorial para dados dicotômicos ou dicotomizados;
- Capítulo 7: Aplicabilidade e resultados;
- Capítulo 8: Considerações finais e sugestões.

Capítulo 2

Teoria Clássica

Qualquer modelo matemático inclui um conjunto de suposições sobre os dados para os quais o modelo é aplicado e especifica o relacionamento entre construtos observados e não observados descritos no modelo. No modelo clássico, dois construtos são introduzidos: o escore verdadeiro e o erro de medida. O escore verdadeiro para um indivíduo pode ser definido como um valor esperado dos seus escores obtidos em vários testes. O erro de medida pode ser definido como a diferença entre o escore verdadeiro e o observado. O modelo clássico supõe que: (1) os erros de medida são aleatórios com média zero e não correlacionados entre si e com os escores verdadeiros e (2) os escores verdadeiros, os observados e os erros de medida são linearmente relacionados. Matematicamente tem-se o modelo

$$x = t + \epsilon, \tag{2.1}$$

onde x, t e ϵ são, respectivamente, o escore observado, o escore verdadeiro e o erro de medida.

As suposições seriam:

1. $\mathrm{E}(\epsilon) = 0$;

2. $\rho(t, \epsilon) = 0$;

3. $\rho(\epsilon_1, \epsilon_2) = 0$,

onde ϵ_1 e ϵ_2 são os erros de medida em duas aplicações de um teste.

Os principais índices calculados na Teoria Clássica para cada item são: índice de dificuldade, discriminação, correlação bisserial e correlação ponto bisserial.

2.1 Índice de dificuldade

O Índice de dificuldade (I) representa a proporção de acerto em um item. Quanto maior a proporção de acerto, mais fácil é esse item. Esse índice varia de 0 (niguém acertou o item) a 1 (todos acertaram o item). Em geral testes que atingem um índice médio de dificuldade em torno de 0,5 produzem distribuições de escores no teste com maior variação Pasquali (1997). Condé (2001) sugere a seguinte interpretação:

- Item fácil: $I > 0,70$;
- Item de dificuldade média: $0,30 < I \leq 0,70$;
- Item difícil $I \leq 0,30$.

2.2 Índice de discriminação

Diferença entre a porcentagem de acerto do grupo superior e a porcentagem de acerto do grupo inferior varia de -1 a 1.

- Grupo superior: 27% dos respondentes com os escores mais altos;
- Grupo inferior: 27% dos respondentes com os escores mais baixos;
- Grupo intermediário 46% dos respondentes restantes.

Esperamos que, para cada item, os indivíduos do Grupo Superior apresentem uma proporção de acerto maior que o do Grupo Intermediário, e que este apresente uma proporção de acertos maior que do Grupo Inferior. Quanto maior a diferença Psup - Pinf, maior será o potencial de discriminação do item. Vamos, então, adotar Disc = Psup - Pinf.

Escala:

- ≥ 40 Bom;
- 30 a 39 Bom, mas sujeito a aprimoramento;
- 20 a 29 Item marginal, sujeito a reelaboração;
- ≤ 19 Item deficiente, que deve ser rejeitado.

2.3 Correlação ponto bisserial

Quando o cálculo do Coeficiente Bisserial é efetuado para cada uma das alternativas, tem-se a correlçao da opção de respostas do indivíduo ao item com o seu desempenho no teste como um todo. Assim, espera-se que alunos que se desempenham bem no teste, tenham feito a opção pela alternativa correta de um determinado item. Caso esses alunos tenham sido atraídos a responder qualquer uma das outras alternativas que não a certa, o item não é discriminativo e não consegue diferenciar os alunos que construíram proficiências, daqueles que as não construíram. Esse índice é uma medida estatística capaz de identificar itens com formulação inadequada ou com erro de gabarito. Sua fórmula é:

$$\overline{r}_{pbis} = \frac{\overline{S}_p - \overline{S}}{\overline{\sigma}_s}\sqrt{\frac{\overline{p}}{\overline{q}}} \tag{2.2}$$

onde,

- $\overline{S}_p$ = escore médio no teste para os que acertaram o item;
- $\overline{S}$ = escore médio no teste para todos;
- $\overline{\sigma}_s$ = desvio padrão dos escores obtidos no teste pelos respondentes;
- $\overline{p}$ = proporção de indivíduos que acertam o item no teste;
- $\overline{q}$ = complementar de $\overline{p}$.
- $\overline{r}_{pbis}$ = estimativa para a correlação de Pearson e o que frequentemente se denomina na literatura de correlação ponto bisserial.

2.4 Correlação bisserial

A Correlação Bisserial é um índice de discriminação que indica a correlação entre o desempenho no item e o desempenho no teste como um todo. Espera-se de uma resposta a um item discriminativo que alunos que vão bem na prova como um todo, acertem-no, e por sua vez, aqueles que não vão, errem-no.

Quanto maior o valor desse índice, maior a capacidade do item de discriminar grupos de estudantes que construíram determinada proficiência, daqueles que não a construíram.

Os itens com coeficiente baixo não diferenciam o indivíduo que construiu, daquele que não construiu determinada proficiência. A correlação bisserial é menos influenciada pela dificuldade do item e tende a apresentar menos variação de uma situação de testagem para outra Wilson et al (1991). Sua fórmula é:

$$\overline{r}_{bis} = \overline{r}_{pbis}\frac{\sqrt{(\overline{p})(\overline{q})}}{h(z_{\overline{p}})}$$

onde,

- $\overline{r}_{pbis}$ = correlação ponto bisserial ou correlação de Pearson;
- $h(z_{\overline{p}})$ = valor da função de densidade normal padrão em $z_{\overline{p}}$.

2.5 Limitações

Pasquali (2004) relata que o enorme impacto que a TRI vem tendo em Psicometria se deve ao fato dela superar certas limitações teóricas graves que a Psicometria Clássica contém. Hambleton, Swaminathan e Rogers (1991) salientam quatro dessas limitações:

1. Os parâmetros clássicos dos itens (dificuldade e discriminação) dependem diretamente da amostra de sujeitos utilizada para estabelecê-los. Daí, se a amostra não for rigorosamente representativa da população, aqueles parâmetros dos itens não podem ser considerados válidos para esta população. Como conseguir amostras representativas é um problema prático grave para os construtores de testes, a dependência dos parâmetros dos itens na amostra obtida se torna um empecilho de grandes proporções para a elaboração de instrumentos psicométricos não viesados.

2. A avaliação das aptidões dos testandos também depende do teste utilizado. Assim, testes diferentes que medem a mesma aptidão irão produzir escores diferentes da mesma aptidão para sujeitos idênticos. Testes com índices de dificuldades diferentes evidentemente produzirão escores diferentes. No caso das formas paralelas de testes, é preciso observar que, em primeiro lugar, conseguir formas estritamente paralelas é uma tarefa quase impossível e, em segundo lugar, mesmo conseguindo formas paralelas, é difícil pressupor que elas produzem o mesmo montante de erro, o que vem afetar a estimação de escore verdadeiro dos sujeitos.

3. A definição do conceito de fidedignidade ou precisão na teoria clássica dos testes constitui também uma fonte de dificuldades. Ela é concebida como a correlação entre os escores obtidos de formas paralelas de um teste ou, mais genericamente, como o oposto do erro de medida. Ambos os conceitos apresentam dificuldades. Em primeiro lugar, é praticamente impossível satisfazer as condições de definição de formas paralelas e, no caso do erro de medida, é postulado que seja idêntico em todos os examinandos, postulado improvável Lord (1984), uma vez que fica difícil presumir que sujeitos de baixa aptidão, por exemplo, cometam erros iguais aos de proficiências superiores.

4. Outro problema da teoria clássica dos testes consiste em que ela é orientada para o teste total e não para o item individual. Toda a informação do item deriva de considerações do teste geral, não se podendo assim determinar como examinando se comportaria diante de cada item individual. Ademais, a análise de cada item é feita em função do escore total, do qual cada item faz parte. Então, fica um tanto incongruente avaliar a qualidade do item quando ele próprio contribui para a mesma e, ainda, a admissão de escore total já supõe que os itens sejam adequados; em sendo o caso, para que fazer a análise individual de cada item em função de todos os outros, que, aliás, ainda não foram analisados em sua adequação?

Após esse breve relato, fica claro que o desejável seria ter (a) estatísticas de itens não dependentes do grupo, (b) escores que não dependessem da dificuldade do teste para descrever a proficiência dos indivíduos e (c) modelos que não requeiram testes estritamente paralelos para avaliar a confiança ou fidedignidade. Há agora evidências substanciais para sugerir que essas propriedades desejáveis, além de outras, possam ser obtidas em uma outra estrutura de teoria de medidas, conhecida como Teoria da Resposta ao Item Andrade et al (2000).

2.6 TRI - Histórico

Segundo Pasquali (2004) a Teoria de Resposta ao Item já tem uma longa história. Ela se baseia no modelo de traço latente, que possui uma história mais longa ainda, dentro do qual pode-se encontar autores dos anos 30, tais como Thurstone, Lumsden (1980), Likert, Andersen (1977), Andrich (1978), Lawley (1943), Guttman(1941, 1944) Lazarfeld (1950).

Contudo, a TRI começou a ser formalizada mais tecnicamente com o trabalho de Lord (1952, 1953) nos Estados Unidos, onde foram primeiramente desenvolvidos os modelos na forma de uma função ogiva normal e, depois, foram descritos para uma forma matemática mais conveniente, e que vem sendo usada até então: a logística e, Rach (1960) na Dinamarca, que a utilizaram para testes de desempenho e de aptidão. Entretanto, apenas ultimamente, a partir de meados dos anos 80, a TRI vem se tornando a técnica predominante no campo dos testes. A razão da demora desta teoria em ser amplamente utilizada se deve ao fato da enorme complexidade de manipulação de seus modelos matemáticos, inviáveis sem os complexos programas de computador e estes só começaram efetivamente a entrar no mercado nos anos 80, com os procedimenos de estimação dos parâmetros do modelo desenvolvidos por Wood, Wingersky & Lord (1978) e por Wingersky, Barton & Lord (1982).

Atualmente, a TRI parece que veio para ficar e substituir grande parte da Teoria Clássica de Medidas. Isto é um fato em países do Primeiro Mundo (Estados Unidos, Canadá, Europa, Japão, Israel, Austrália); no restante do mundo ela é ainda raramente utilizada e no Brasil (América Latina em geral) ela chegou a ser conhecida apenas a partir dos anos 90. Recentemente, Bock & Zimowski (1997) introduziram os modelos logísticos de 1, 2 e 3 parâmetros para duas ou mais populações de respondentes. A introdução desses modelos 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.

Nojosa (2003) ressalta que um ponto crítico na TRI é a estimação dos parâmetros envolvidos nos modelos, em particular quando necessita-se estimar tanto os parâmetros dos itens quanto as proficiências. Inicialmente, a estimação era feita através do método da máxima verossimilhança conjunta que envolve um número muito grande de parâmetros a serem estimados simultaneamente e, conseqüentemente, grandes problemas computacionais.

Em 1970, Bock & Lieberman introduziram o método da máxima verossimilhança marginal para a estimação dos parâmetros em duas etapas.

Em 1981, Bock & Aitkin propuseram uma modificação no método acima, utilizando o algoritmo EM de Dempster, Laird & Rubin (1977), de modo a permitir que os itens pudessem ter seus parâmetros estimados em separado, facilitando em muito o aspecto computacional do processo de estimação.

Mais recentemente, métodos bayesianos foram propostos para, entre outras coisas, resolver o problema de estimação dos parâmetros dos itens respondidos corretamente ou incorretamente por todos os respondentes, e também o problema da estimação das proficiências dos respondentes que acertaram ou erraram todos os itens da prova Nojosa (2003).

Nas últimas décadas, a TRI vem tornando-se a técnica predominante no campo de testes em vários países. Aqui no Brasil, a TRI foi usada pela primeira vez em 1995 na análise dos dados do Sistema Nacional de Ensino Básico - SAEB. A introdução da TRI permitiu que os desempenhos de alunos de 4ª e 8ª séries do Ensino Fundamental e de 3ª série do Ensino Médio pudessem ser comparados e colocados em uma escala única de conhecimento. Esta técnica também tem sido adotada nas avaliações educacionais de vários outros estados brasileiros, tais como São Paulo, Rio de Janeiro, Minas Gerais, Rio Grande do Sul, Ceará, dentre outros. O INEP/MEC também usa em outras avaliações a saber, Enceja, Provinha Brasil e Prova Brasil.

Capítulo 3

Modelos Unidimensionais

3.1 Introdução

A TRI é um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da proficiência (ou proficiências) do respondente. Essa relação é sempre expressa de tal forma que quanto maior a proficiência, maior a probabilidade de acerto do item. Os vários modelos propostos na literatura dependem fundamentalmente de três fatores:

1. da natureza do item - dicotômicos ou não dicotômicos;
2. do número de populações envolvidas - apenas uma ou mais de uma;
3. e da quantidade de traços latentes que está sendo medida - apenas um ou mais de um.

Nesse capítulo estaremos somente considerando modelos que avaliam apenas um traço latente ou proficiência, os chamados modelos unidimensionais. Modelos que consideram que mais de uma proficiência está sendo medida, os chamados modelos multidimensionais não serão contemplados neste trabalho.

Mais adiante mostraremos os modelos unidimensionais mais utilizados para um único grupo. Os modelos para dois ou mais grupos serão abordados na Seção 3.3.

3.2 Modelos envolvendo um único grupo

Em Tavares (2000) pode-se ver a definição dos conceitos de grupo e população, que serão largamente utilizados neste e nos demais capítulos, da seguinte forma: Quando o termo *grupo*, for usado, se estará dando referência a uma amostra de indivíduos de uma *população*. O conceito de grupo está diretamente ligado ao processo de amostragem —

e estaremos sempre considerando o processo de amostragem aleatória simples. Portanto, quando falarmos em um único grupo de respondentes, nos referimos a uma amostra de indivíduos retirada de uma mesma população. Consequentemente, dois grupos – ou mais – de respondentes são dois conjuntos distintos de indivíduos, que foram amostrados de duas – ou mais – populações.

Na área de Avaliação Educacional é comum que uma população seja definida por determinadas características que podem variar, dependendo dos objetivos do estudo, e portanto, podem ou não ser relevantes para a diferenciação de populações. Por exemplo, pode-se considerar que a 3ª série do Ensino Médio do Brasil é a população alvo. Daí, toma-se uma única amostra dos alunos dessa população, composta de alunos do período diurno e do noturno. Nesse caso, temos então um único grupo de respondentes. Já em outro estudo, poderíamos considerar a 3ª série *diurna* e a 3ª série *noturna* do Ensino Médio do Brasil como duas populações de interesse. Então, seriam tomadas duas amostras: uma dos alunos do período diurno e outra dos alunos do noturno. Nessa situação, teríamos dois grupos de alunos. Portanto, é pelo próprio processo de amostragem do estudo que identifica-se quantas (e quais) populações estão envolvidas.

A seguir, serão mostrados os modelos mais utilizados quando um teste é aplicado a um único grupo de respondentes.

3.2.1 Modelos para itens dicotômicos ou dicotomizados

Atualmente para analisar testes de multipla escolha, dicotomizados, (corrrigidos como certo ou errado) os modelos logísticos para itens dicotômicos são os modelos de resposta ao item mais utilizados, sendo que há basicamente três tipos, que se diferenciam pelo número de parâmetros que utilizam para descrever o item. Eles são conhecidos como os modelos logísticos de 1, 2 e 3 parâmetros, que consideram, respectivamente somente a dificuldade do item, a dificuldade e a discriminação, a dificuldade, a discriminação e a probabilidade de resposta correta dada por indivíduos de baixa proficiência. Porém, os primeiros modelos desenvolvidos foram na forma de uma função distribuição da normal.

Nesta seção, será apresentado o modelo ogiva normal e o logístico de 3 parâmetros, na qual será dado ênfase ao segundo.

O modelo normal

$$\Phi(\theta) = \int_{-\infty}^{a_i(\theta-b_i)} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt \tag{3.1}$$

Esses modelos foram depois descritos por uma forma mais conveniente computacionalmente que é a função logística.

O modelo logístico de 3 parâmetros (ML3)

Definição

Dos modelos propostos pela TRI, o *modelo logístico unidimensional de 3 parâmetros (ML3)* é atualmente o mais utilizado e é dado por:

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i)\frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \tag{3.2}$$

com $i = 1, 2, \cdots, I$, e $j = 1, 2, \cdots, n$, onde:

U_{ij} é uma variável dicotômica que assume os valores 1, quando o indivíduo j responde corretamente o item i, ou 0 quando o indivíduo j não responde corretamente ao item i.

θ_j representa a proficiência (traço latente) do j-ésimo indivíduo.

$P(U_{ij} = 1|\theta_j)$ é a probabilidade de um indivíduo j com proficiência θ_j responder corretamente o item i e é chamada de Função de Resposta do Item – FRI.

b_i é o parâmetro de dificuldade (ou de posição) do item i, medido na mesma escala da proficiência.

a_i é o parâmetro de discriminação (ou de inclinação) do item i, com valor proporcional à inclinação da Curva Característica do Item — CCI no ponto b_i.

c_i é o parâmetro do item que representa a probabilidade de indivíduos com baixa proficiência responderem corretamente o item i (muitas vezes referido como a probabilidade de acerto casual).

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.

Interpretação e representação gráfica

Note que $P(U_{ij} = 1|\theta_j)$ pode ser vista como a proporção de respostas corretas ao item i dentre todos os indivíduos da população com proficiência θ_j. A relação existente entre $P(U_{ij} = 1|\theta_j)$ e os parâmetros do modelo é mostrada na Figura 3.1, que é chamada

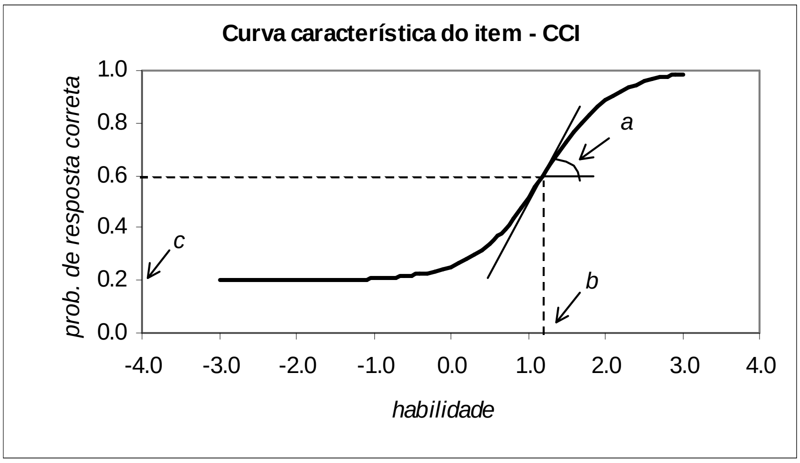

Figura 3.1 *Exemplo de uma Curva Característica do Item - CCI*

O modelo proposto baseia-se no fato de que indivíduos com maior proficiência possuem maior probabilidade de acertar o item e que esta relação não é linear. De fato, pode-se perceber a partir do gráfico acima que a CCI tem forma de "S"com inclinação e deslocamento na escala de proficiência definidos pelos parâmetros do item.

A escala da proficiência é uma escala *arbitrária* onde o importante são as relações de ordem existentes entre seus pontos e não necessariamente sua magnitude. O parâmetro b é medido na mesma unidade da proficiência e o parâmetro c não depende da escala, pois trata-se de uma probabilidade, e como tal, assume sempre valores entre 0 e 1.

Na realidade, o parâmetro b representa a proficiência necessária para uma probabilidade de acerto igual a $(1 + c)/2$. Assim, quanto maior o valor de b, mais difícil é o item, e vice-versa.

O parâmetro c representa a probabilidade de um aluno com baixa proficiência responder corretamente o item e é muitas vezes referido como a probabilidade de acerto ao acaso. Então, quando não é permitido "chutar", c é igual a 0 e b representa o ponto na escala da proficiência onde a probabilidade de acertar o item é 0,5.

O parâmetro a é proporcional à inclinação da curva no ponto de inflexão. Assim, itens com a negativo não são esperados sob esse modelo, uma vez que indicariam que a probabilidade de responder corretamente o item diminui com o aumento da proficiência. Baixos valores de a indicam que o item tem pouco poder de discriminação (alunos com proficiências bastante diferentes têm aproximadamente a mesma probabilidade de responder corretamente ao item) e valores muito altos indicam itens com curvas características muito "íngremes", que discriminam os alunos basicamente em dois grupos: os que possuem proficiências abaixo do valor do parâmetro b e os que possuem proficiências acima do valor do parâmetro b Andrade, Tavares & Vale (2000).

Uma medida bastante utilizada em conjunto com a CCI é a *função de informação do item*. Ela permite analisar quanto um item (ou teste) contém de informação para a medida de proficiência. A função de informação de um item é dada por:

$$I_i(\theta) = \frac{\left[\frac{d}{d\theta} P_i(\theta)\right]^2}{P_i(\theta) Q_i(\theta)},$$

onde,

$I_i(\theta)$ é a "informação" fornecida pelo item i no nível de proficiência θ;

$P_i(\theta) = P(X_{ij} = 1|\theta)$ e $Q_i(\theta) = 1 - P_i(\theta)$.

No caso do modelo logístico de 3 parâmetros, a equação pode ser escrita como:

$$I_i(\theta) = D^2 a_i^2 \frac{Q_i(\theta)}{P_i(\theta)} \left[\frac{P_i(\theta) - c_i}{1 - c_i}\right]^2.$$

Esta equação mostra a importância que têm os três parâmetros sobre o montante de informação do item. Isto é, a informação é maior quando b_i se aproxima de θ quanto maior for o a_i e quanto mais c_i se aproximar de 0.

A informação fornecida pelo teste é simplesmente a soma das informações fornecidas por cada item que compõe o mesmo:

$$I(\theta) = \sum_{i=1}^{I} I_i(\theta).$$

Outra maneira de representar esta função de informação do teste é através do erro-padrão de medida, chamado na TRI de erro-padrão de estimação, que é dado por

$$EP(\theta) = \frac{1}{\sqrt{I(\theta)}}.$$

É importante notar que essas medidas de informação dependem do valor de θ. Assim, a amplitude do intervalo de confiança para θ dependerá também do seu valor.

3.3 Modelos envolvendo dois ou mais grupos de respondentes

Alguns modelos já foram desenvolvidos para serem aplicados quando um teste envolve mais de uma população, sendo basicamente, extensões dos modelos até aqui discutidos. No entanto, um dos poucos modelos que já se encontram implementados computacionalmente e que portanto, já estão sendo utilizados na prática, quando um teste é aplicado a dois ou mais grupos de respondentes, é uma generalização dos modelos logísticos unidimensionais de 1, 2 e 3 parâmetros, que foi recentemente proposta por Bock & Zimowski (1997). O modelo é dado por:

$$P(U_{ijk} = 1|\theta_{jk}) = c_i + (1 - c_i)\frac{1}{1 + e^{-Da_i(\theta_{jk} - b_i)}}, \tag{3.3}$$

com $i = 1, 2, \cdots, I,\ j = 1, 2, \cdots, n_k,$ e $k = 1, 2, \cdots, K$, onde:

U_{ijk} é uma variável dicotômica que assume os valores 1, quando o indivíduo j da população k responde corretamente ao item i, ou 0 quando o indivíduo não responde corretamente ao item.

θ_{jk} representa a proficiência do j-ésimo indivíduo da população k.

$P(U_{ijk} = 1|\theta_{jk})$ é a probabilidade de um indivíduo j da população k, com proficiência θ_{jk}, responder corretamente ao item i.

Os demais parâmetros já foram descritos anteriormente.

Para que seja possível a comparação entre grupos de respondentes, é sugerido que haja pelo menos vinte porcento de itens comuns entre os grupos. Assim, I representa o número total de itens *distintos* apresentados.

Andrade, Tavares & Vale (2000) revelam que a recente implementação computacional desse modelo para mais de um grupo de respondentes foi um dos maiores avanços da TRI nos últimos anos. Através dele a comparação de indivíduos de grupos distintos, submetidos a provas diferentes mas com itens comuns, passou a ser feita de uma maneira bem mais eficiente do que era feito até então, uma vez que diminui possíveis erros de modelagem que a metodologia anterior poderia vir a ter. A seguir faremos uma abordagem sobre alguns métodos de estimação dos parâmetros dos modelos para um ou mais grupos. Maiores detalhes sobre estimação de parâmetros dos itens e das proficiências, podem ser encontrados em Andrade, Tavares & Vale (2000) e Azevedo (2003).

Capítulo 4

Métodos de Estimação

4.1 Introdução

Uma vez determinado o modelo da TRI a ser utilizado, é necessário determinar os valores dos parâmetros dos itens e das proficiências dos indivíduos. Nos modelos unidimensionais, cada indivíduo é caracterizado apenas por um parâmetro, θ, e para a ca-racterização dos itens utilizam-se de 1 a 3 parâmetros, dependendo do modelo utilizado. De modo geral, obtêm-se estimativas para os parâmetros do modelo com erros-padrão pequenos quando o número de itens é de pelo menos 30 e o número de respondentes para cada item é de pelo menos 300 Nojosa (2003). No processo de estimação, nos defrontamos com 3 casos. O primeiro é quando os parâmetros dos itens são conhecidos e deseja-se apenas estimar as proficiências dos indivíduos. No segundo caso são conhecidas as proficiências e deseja-se apenas estimar os parâmetros dos itens. No terceiro, nem os parâmetros dos itens e nem as proficiências dos indivíduos são conhecidos; deseja-se estimar ambos. O primeiro caso começa a ser visto frequente na prática, e a solução é dada empregando o método da máxima verossimilhança ou métodos bayesianos, ambos através da aplicação de procedimentos interativos, como, por exemplo, o método de Newton-Raphson ou scoring de Fisher. O segundo caso tem apenas caráter teórico e é solucionado usando o método da máxima verossimilhança. O terceiro caso, provavelmente o mais encontrado na prática, é abordado de duas formas: a estimação conjunta dos parâmetros de itens e das proficiências dos indivíduos; ou em duas etapas, primeiro a estimação dos parâmetros dos itens e, em seguida, a das proficiências Baker (1992).

A seguir serão abordados apenas os métodos de Máxima Verossimilhança e Máxima Virossimilhança Maginal, com o intuito de melhor ser entendido quanto a questão de que método de estimação foi utilizado neste trabalho. Maiores detalhes sobre esses e outros métodos de estimação podem ser vistos em Andrade, Tavares & Vale (2000) e Azevedo (2003), por exemplo. Nos métodos de estimação descritos a seguir, algumas notações e suposições serão necessárias para o desenvolvimento do modelo. Em particular, sejam θ_j a proficiência e U_{ji} a variável aleatória que representa a resposta (binária) do indivíduo j ao item i, com

$$U_{ji} = \begin{cases} 1, & \text{resposta correta,} \\ 0, & \text{resposta incorreta.} \end{cases}$$

Sejam $\mathbf{U}_{j.} = (U_{j1}, U_{j2}, \cdots, U_{jI})$ o vetor aleatório de respostas do indivíduo j e $\boldsymbol{U}_{..} = (\boldsymbol{U}_{1.}, \cdots, \boldsymbol{U}_{n.})$ o conjunto integral de respostas. De forma similar, representaremos as observações por $u_{ji}, u_{j.}$ e $u_{..}$. Ainda, $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_n)$ representará o vetor de proficiências dos n indivíduos e $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \cdots, \boldsymbol{\zeta}_I)$ o conjunto de parâmetros dos itens.

As duas principais suposições que serão usadas em todo o restante deste trabalho, são as seguintes:

(S1) as respostas oriundas de indivíduos diferentes são independentes,

(S2) os itens são respondidos de forma independente por cada indivíduo (Independência Local), fixada sua proficiência.

4.2 Estimação dos parâmetros dos itens

Nesta seção trataremos da estimação dos parâmetros dos itens pelo método da máxima verossimilhança, quando as proficiências são conhecidas. Por estar muito bem descrito na literatura disponível, não iremos aquí dar ênfase no desenvolvimento destas equações, pois as mesmas podem ser encontradas em Baker (1992), Andrade, Tavares & Vale (2000) e Azevedo (2003) por exemplo. Para fins de aplicação, o ML3 tem sido amplamente empregado e, por isso, o usaremos para efeito de exemplificação. Os modelos ML1 e ML2 são casos particulares do ML3; a escolha desse último leva a resultados que servem para os dois primeiros.

Pela independência entre as respostas de diferentes indivíduos (S1) e a independência local (S2), podemos escrever a verossimilhança, $L(\boldsymbol{\zeta}) = P(\boldsymbol{U}_{..} = \boldsymbol{u}_{..}|\boldsymbol{\theta}, \boldsymbol{\zeta})$, como

$$\begin{aligned} L(\boldsymbol{\zeta}) &= \prod_{j=1}^{n} P(U_{j.} = u_{j.}|\theta_j, \boldsymbol{\zeta}) \\ &= \prod_{j=1}^{n}\prod_{i=1}^{I} P(U_{ji} = u_{ji}|\theta_j, \boldsymbol{\zeta}_i), \end{aligned}$$

onde na última igualdade usamos que a distribuição de U_{ji} só depende de $\boldsymbol{\zeta}$ através de $\boldsymbol{\zeta}_i$. Usando a notação $P_{ji} = P(U_{ji} = 1|\theta_j, \boldsymbol{\zeta}_i)$ e $Q_{ji} = 1 - P_{ji}$, temos que

$$\begin{aligned} P(U_{ji} = u_{ji}|\theta_j, \boldsymbol{\zeta}_i) &= P(U_{ji} = 1|\theta_j, \boldsymbol{\zeta}_i)^{u_{ji}} P(U_{ji} = 0|\theta_j, \boldsymbol{\zeta}_i)^{1-u_{ji}} \\ &= P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}. \end{aligned}$$

Portanto,

$$L(\boldsymbol{\zeta}) = \prod_{j=1}^{n} \prod_{i=1}^{I} P_{ji}^{u_{ji}} Q_{ji}^{1-u_{ji}}.$$

Segue que a log-verossimilhança pode ser escrita como

$$\log L(\boldsymbol{\zeta}) = \sum_{j=1}^{n} \sum_{i=1}^{I} \{u_{ji} \log P_{ji} + (1 - u_{ji}) \log Q_{ji}\}. \tag{4.1}$$

Os Estimadores de Máxima verossimilhança (EMV) de $\boldsymbol{\zeta}_i$, $i = 1, \cdots, I$, são os valores que maximizam a verossimilhança, ou equivalente, são as soluções da equação

$$\frac{\partial \log L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i} = 0, \quad i = 1, \cdots, I.$$

Notemos que

$$\begin{aligned} \frac{\partial \log L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i} &= \sum_{j=1}^{n} \left\{ u_{ji} \frac{\partial (\log P_{ji})}{\partial \boldsymbol{\zeta}_i} + (1 - u_{ji}) \frac{\partial (\log Q_{ji})}{\partial \boldsymbol{\zeta}_i} \right\} \\ &= \sum_{j=1}^{n} \left\{ u_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) - (1 - u_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \right\} \\ &= \sum_{j=1}^{n} \left\{ u_{ji} \frac{1}{P_{ji}} - (1 - u_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) \\ &= \sum_{j=1}^{n} \left\{ \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right). \end{aligned}$$

Por conveniência, consideremos a seguinte ponderação:

$$W_{ji} = \frac{P_{ji}^* Q_{ji}^*}{P_{ji} Q_{ji}}, \tag{4.2}$$

onde

$$P^*_{ji} = \{1 + e^{-Da_i(\theta_j - b_i)}\}^{-1} \quad \text{e} \quad Q^*_{ji} = 1 - P^*_{ji}. \tag{4.3}$$

Em resumo, as equações de estimação para os parâmetros a_i, b_i e c_i são, respectivamente,

$$a_i \quad : \quad D(1 - c_i)\sum_{j=1}^{n}(u_{ji} - P_{ji})(\theta_j - b_i)W_{ji} = 0, \tag{4.4}$$

$$b_i \quad : \quad -Da_i(1 - c_i)\sum_{j=1}^{n}(u_{ji} - P_{ji})W_{ji} = 0, \tag{4.5}$$

$$c_i \quad : \quad \sum_{j=1}^{n}(u_{ji} - P_{ji})\frac{W_{ji}}{P^*_{ji}} = 0. \tag{4.6}$$

Estas equações não possuem solução explícita, por isso será usado algum método iterativo. Os métodos iterativos mais usados em problemas de estimação paramétrica na TRI são os métodos de Newton-Raphson e o método “Scoring”de Fisher.

Como mencionado anteriormente outros procedimentos de estimação alternativos podem ser encontrados em Andrade, Tavares & Vale (2000) e Azevedo (2003) com bastante propriedade.

4.2.1 Newton-Raphson

Seja $l(\boldsymbol{\zeta}) = \log L(\boldsymbol{\zeta})$ a log-verossimilhança, onde $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \cdots, \boldsymbol{\zeta}_I)$, com $\boldsymbol{\zeta}_i = (a_i, b_i, c_i)'$. Se valores iniciais $\widehat{\boldsymbol{\zeta}}_i^{(0)} = (a_i^{(0)}, b_i^{(0)}, c_i^{(0)})'$ podem ser encontrados para $\boldsymbol{\zeta}_i$, então uma estimativa atualizada será $\widehat{\boldsymbol{\zeta}}_i^{(1)} = \widehat{\boldsymbol{\zeta}}_i^{(0)} + \Delta\widehat{\boldsymbol{\zeta}}_i^{(0)}$, ou seja:

Considerando $\widehat{\boldsymbol{\zeta}}_i^{(t)}$ a estimativa de $\boldsymbol{\zeta}_i$ na iteração t, então na iteração $t+1$ do algoritmo Newton-Raphson teremos que

$$\widehat{\boldsymbol{\zeta}}_i^{(t+1)} = \widehat{\boldsymbol{\zeta}}_i^{(t)} - [H(\widehat{\boldsymbol{\zeta}}_i^{(t)})]^{-1}h(\widehat{\boldsymbol{\zeta}}_i^{(t)}). \tag{4.7}$$

$$\tag{4.8}$$

onde,

$$
\begin{aligned}
h(\boldsymbol{\zeta}_i) &\equiv \frac{\partial \log L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i} && (4.9)\\
&= \sum_{j=1}^{n} \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P^*_{ji} Q^*_{ji}} \right\} (P^*_{ji} Q^*_{ji}) h_{ji} && (4.10)\\
&= \sum_{j=1}^{n} (u_{ji} - P_{ji}) W_{ji} \boldsymbol{h}_{ji}. && (4.11)
\end{aligned}
$$

e

$$
\begin{aligned}
\boldsymbol{H}(\boldsymbol{\zeta}_i) &\equiv \frac{\partial \log L(\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} && (4.12)\\
&= \sum_{j=1}^{n} \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) (P^*_{ji} Q^*_{ji}) \boldsymbol{H}_{ji} - \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 (P^*_{ji} Q^*_{ji})^2 \boldsymbol{h}_{ji} \boldsymbol{h}_{ji}' \right\} && (4.13)\\
&= \sum_{j=1}^{n} (u_{ji} - P_{ji}) W_{ji} \left\{ \boldsymbol{H}_{ji} - (u_{ji} - P_{ji}) W_{ji} \boldsymbol{h}_{ji} \boldsymbol{h}_{ji}' \right\}. && (4.14)\\
& && (4.15)
\end{aligned}
$$

com

$$
\boldsymbol{h}_{ji} = (P^*_{ji} Q^*_{ji})^{-1} \left(\frac{\partial P_{ji}}{\partial \boldsymbol{\zeta}_i} \right) = \begin{pmatrix} D(1 - c_i)(\theta_j - b_i) \\ -D a_i (1 - c_i) \\ \frac{1}{P^*_{ji}} \end{pmatrix},
$$

e

$$
\begin{aligned}
\boldsymbol{H}_{ji} &= (P^*_{ji} Q^*_{ji})^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \right)\\
&= \begin{pmatrix} D^2 (1 - c_i)(\theta_j - b_i)^2 (1 - 2P^*_{ji}) & . & . \\ -D(1 - c_i)\{1 + D a_i (\theta_j - b_i)(1 - 2P^*_{ji})\} & D^2 a_i^2 (1 - c_i)(1 - 2P^*_{ji}) & . \\ -D(\theta_j - b_i) & D a_i & 0 \end{pmatrix}.
\end{aligned}
$$

4.2.2 Aplicação do método "Scoring"de Fisher

Para aplicação do método "Scoring"de Fisher, devemos substituir os componentes da matriz de derivadas segundas usadas no processo iterativo de Newton-Raphson pelos seus valores esperados. Notando que a variável U_{ji} só pode assumir dois valores: 1, com probabilidade P_{ji} e 0 com probabilidade Q_{ji}, então U_{ji} tem distribuição $Bernoulli(P_{ji})$. Segue que $E(U_{ji}) = P_{ji}$ e $E(U_{ji} - P_{ji})^2 = Var(U_{ji}) = P_{ji} Q_{ji}$. Logo, de (4.14), temos que

$$\begin{aligned}
\boldsymbol{\Delta}(\boldsymbol{\zeta}_i) &\equiv E(\boldsymbol{H}(\boldsymbol{\zeta}_i)) && (4.16)\\
&= \sum_{j=1}^{N}\left\{E(U_{ji}-P_{ji})W_{ji}\boldsymbol{H}_{ji} - E(U_{ji}-P_{ji})^2W_{ji}^2\boldsymbol{h}_{ji}\boldsymbol{h}'_{ji}\right\} && (4.17)\\
&= \sum_{j=1}^{N}\left\{-P_{ji}Q_{ji}W_{ji}^2\boldsymbol{h}_{ji}\boldsymbol{h}'_{ji}\right\} && (4.18)\\
&= -\sum_{j=1}^{N}\left\{P^*_{ji}Q^*_{ji}W_{ji}\boldsymbol{h}_{ji}\boldsymbol{h}'_{ji}\right\}. && (4.19)
\end{aligned}$$

Para as proficiências agrupadas, a expressão acima fica

$$\boldsymbol{\Delta}(\boldsymbol{\zeta}_i) = -\sum_{k=1}^{q}\left\{P^*_{ki}Q^*_{ki}W_{ki}\boldsymbol{h}_{ki}\boldsymbol{h}'_{ki}\right\}. \tag{4.20}$$

A expressão para estimativa de $\boldsymbol{\zeta}_i$ na iteração $t+1$ será

$$\widehat{\boldsymbol{\zeta}}_i^{(t+1)} = \widehat{\boldsymbol{\zeta}}_i^{(t)} - [\boldsymbol{\Delta}(\widehat{\boldsymbol{\zeta}}_i^{(t)})]^{-1}\boldsymbol{h}(\widehat{\boldsymbol{\zeta}}_i^{(t)}).$$

4.2.3 Estimativas iniciais

Um ponto importante no processo de estimação é a obtenção de valores iniciais para os parâmetros. Se o item i tem m_i alternativas possíveis, um chute razoável para o parâmetro de acerto ao acaso é $c_i = 1/m_i$. Richardson (1936) e Tucker (1946) mostraram que se adotarmos a FRI Normal, então

$$\rho_{T,U_i} = \frac{a_i}{\sqrt{1+a_i^2}}, \quad -1 < \rho_{T,U_i} < 1, \tag{4.21}$$

onde ρ_{T,U_i} é o coeficiente de correlação bisserial, utilizado na Teoria Clássica de Medidas. Este coeficiente é estimado pelo coeficiente de correlação de Pearson entre os escores, T_j, e as respostas ao item i. Com isso, obtemos $\widehat{a}_i^{(0)}$.

Em complemento, Tucker (1946) expressou o parâmetro de dificuldade associado ao item i da teoria clássica de itens π_i (proporção verdadeira de respostas corretas) como

$$\pi_i = \Phi(-\nu_i), \quad \nu_i = b_i\rho_{T,U_i}, \tag{4.22}$$

onde Φ é a função de distribuição associada à N(0,1). Vale notar que no caso de usar a função Logística para a FRI, o fator $D = 1,702$ torna os modelos Normal e Logístico muito próximos (ver Halley (1952)) de forma que as expressões (4.21) e (4.22) produzem bons resultados para o modelo logístico.

4.3 Estimação das proficiências

Conhecido os parâmetros dos itens, trataremos agora da estimação das proficiências. Na prática, essa situação ocorre quando os itens já foram calibrados (estimados) em outros testes. Como a calibração dos itens deve ser feita com um número grande de indivíduos, a estimação das proficiências de um grupo pequeno de indivíduos é mais confiável se forem utilizados itens já calibrados.

Pela independência entre as respostas de diferentes indivíduos (S1) e a independência local (S2), podemos escrever a log-verossimilhança como em (4.1), agora como função de θ e não de ζ, ou seja,

$$\log L(\boldsymbol{\theta}) = \sum_{j=1}^{n}\sum_{i=1}^{I}\{u_{ji}\log P_{ji} + (1-u_{ji})\log Q_{ji}\}. \tag{4.23}$$

O EMV de θ_j é o valor que maximiza a verossimilhança, ou equivalentemente, é a solução da equação

$$\frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_j} = 0, \quad j = 1, \cdots, n. \tag{4.24}$$

Notemos, de (4.23), que

$$\begin{aligned}
\frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_j} &= \sum_{i=1}^{I} \left\{ u_{ji} \frac{\partial (\log P_{ji})}{\partial \theta_j} + (1 - u_{ji}) \frac{\partial (\log Q_{ji})}{\partial \theta_j} \right\} && (4.25) \\
&= \sum_{i=1}^{I} \left\{ u_{ji} \frac{1}{P_{ji}} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) - (1 - u_{ji}) \frac{1}{Q_{ji}} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) \right\} && (4.26) \\
&= \sum_{i=1}^{I} \left\{ u_{ji} \frac{1}{P_{ji}} - (1 - u_{ji}) \frac{1}{Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) && (4.27) \\
&= \sum_{i=1}^{I} \left\{ \frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) && (4.28) \\
&= \sum_{i=1}^{I} \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P^*_{ji} Q^*_{ji}} \right\} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right), && (4.29) \\
& && (4.30)
\end{aligned}$$

onde a última igualdade segue de (4.2). Como

$$\frac{\partial P_{ji}}{\partial \theta_j} = D a_i (1 - c_i) P^*_{ji} Q^*_{ji}, \qquad (4.31)$$

obtêm-se

$$\begin{aligned}
\frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_j} &= \sum_{i=1}^{I} \left\{ (u_{ji} - P_{ji}) D a_i (1 - c_i) P^*_{ji} Q^*_{ji} \frac{W_{ji}}{P^*_{ji} Q^*_{ji}} \right\} \\
&= D \sum_{i=1}^{I} a_i (1 - c_i)(u_{ji} - P_{ji}) W_{ji}. && (4.32) \\
& && (4.33)
\end{aligned}$$

Segue então que a equação de estimação (4.24) para $\theta_j, \ j = 1, \cdots, n$, é

$$\theta_j \quad : \quad D \sum_{i=1}^{I} a_i (1 - c_i)(u_{ji} - P_{ji}) W_{ji} = 0. \qquad (4.34)$$

Novamente, esta equação não apresenta solução explícita para θ_j e, por isso, precisamos de algum método iterativo para obter as estimativas desejadas. A seguir, obteremos as expressões necessárias para aplicações dos processos iterativos Newton-Raphson e "Scoring"de Fisher.

4.3.1 Newton-Raphson

Considerando $\widehat{\theta}_j^{(t)}$ a estimativa de θ_j na iteração t, então na iteração $t+1$ do algoritmo Newton-Raphson teremos que

$$\widehat{\theta}_j^{(t+1)} = \widehat{\theta}_j^{(t)} - [H(\widehat{\theta}_j^{(t)})]^{-1} h(\widehat{\theta}_j^{(t)}) \tag{4.35}$$

na qual, [ver Andrade, Tavares e Vale (2000)] para maiores detalhes das demonstrações dos resultados,

$$\begin{aligned} h(\theta_j) &\equiv \frac{\partial \log L(\boldsymbol{\theta})}{\partial \theta_j} && (4.36)\\ &= \sum_{i=1}^{I} \left\{ (u_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^* Q_{ji}^*} \right\} (P_{ji}^* Q_{ji}^*) h_{ji} && (4.37)\\ &= \sum_{i=1}^{I} (u_{ji} - P_{ji}) W_{ji} h_{ji} && (4.38) \end{aligned}$$

e

$$\begin{aligned} H(\theta_j) &\equiv \frac{\partial^2 \log L(\boldsymbol{\theta})}{\partial \theta_j^2} && (4.39)\\ &= \sum_{i=1}^{I} \left\{ \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right) (P_{ji}^* Q_{ji}^*) H_{ji} - \left(\frac{u_{ji} - P_{ji}}{P_{ji} Q_{ji}} \right)^2 (P_{ji}^* Q_{ji}^*)^2 h_{ji}^2 \right\} && (4.40)\\ &= \sum_{i=1}^{I} (u_{ji} - P_{ji}) W_{ji} \left\{ H_{ji} - (u_{ji} - P_{ji}) W_{ji} h_{ji}^2 \right\} && (4.41) \end{aligned}$$

com

$$h_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial P_{ji}}{\partial \theta_j} \right) = D a_i (1 - c_i) \tag{4.42}$$

e

$$H_{ji} = (P_{ji}^* Q_{ji}^*)^{-1} \left(\frac{\partial^2 P_{ji}}{\partial \theta_j^2} \right) = D^2 a_i^2 (1 - c_i)(1 - 2P_{ji}^*). \tag{4.43}$$

4.3.2 Aplicação do método "Scoring"de Fisher

Para aplicação do método "Scoring"de Fisher, devemos substituir os componentes da matriz de derivadas segundas usadas no processo iterativo de Newton-Raphson pelos seus valores esperados. Por (4.41) temos que

$$\begin{aligned}
\Delta(\theta_j) &\equiv E(H(\theta_j)) && (4.44)\\
&= \sum_{i=1}^{I} \left\{ E(U_{ji} - P_{ji}) W_{ji} H_{ji} - E(U_{ji} - P_{ji})^2 W_{ji}^2 h_{ji}^2 \right\} && (4.45)\\
&= -\sum_{i=1}^{I} P_{ji}^{*} Q_{ji}^{*} W_{ji} h_{ji}^2. && (4.46)
\end{aligned}$$

Neste caso, a expressão para estimativa de $\theta_j, j = 1, \cdots, n$, na iteração $t+1$ será

$$\widehat{\theta}_j^{(t+1)} = \widehat{\theta}_j^{(t)} - [\Delta(\widehat{\theta}_j^{(t)})]^{-1} h(\widehat{\theta}_j^{(t)}).$$

4.3.3 Estimativas iniciais

A obtenção de estimativas (valores) iniciais para o início do processo de estimação pode ser feita com os escores padronizados. Se T_j é o escore do indivíduo j, m o escore médio e s o desvio-padrão dos escores dos n indivíduos, então $\widehat{\theta}_j^{(0)} = (T_j - m)/s$.

4.4 Máxima verossimilhança marginal

Basicamente o método da Máxima Verossimilhança Marginal, proposto por Bock & Lieberman (1970) apresenta algumas vantagens em relação ao método da Máxima Verossimilhança Conjunta [(detalhes sobre este e outros método de estimação podem ser visto em Andrade, Tavares & Vale (2000)]. A proposta desse método é fazer a estimação em duas etapas: primeiro os parâmetros dos itens e, posteriormente, as proficiências. Como as proficiências não são conhecidas, precisaremos usar algum artifício de forma que a verossimilhança não seja mais função das proficiências. Anderson (1980) argumenta que se considerarmos uma população Π composta por n indivíduos com proficiências $\theta_j,\ j = 1, \cdots, n$, e construirmos a distribuição de frequência acumulada $G(\theta)$=(número de $j : \theta_j \leq \theta)/n$, então, se n for suficientemente grande os θ_j estarão bastante próximos de forma que $G(\theta)$ pode ser aproximada por uma distribuição contínua. A densidade $g(\theta)$, relativa à $G(\theta)$, pode realmente ser considerada a função densidade para θ no experimento de retirar um indivíduo ao acaso da população Π e observar seu parâmetro θ. Neste contexto, é importante ressaltar que, quando atribuímos uma distribuição de probabilidade para θ *não estamos aplicando nenhum argumento bayesiano.* A distribuição de θ realmente existe, no sentido explicado acima, como a densidade relativa à distribuição $G(\theta)$ Andrade, Tavares & Valle (2000).

De acordo com isso, um artifício para eliminar as proficiências na verossimilhança consiste em marginalizar a verossimilhança integrando-a com relação à distribuição da proficiência. De forma geral, consideremos que as proficiências, $\theta_j,\ j = 1, \cdots, n$, são realizações de uma variável aleatória θ com distribuição contínua e função densidade de probabilidade (fdp) $g(\theta|\boldsymbol{\eta})$, duplamente diferenciável, com as componentes de $\boldsymbol{\eta}$ conhecidas e finitas. Para o caso em que θ tem distribuição Normal, temos $\boldsymbol{\eta} = (\mu, \sigma^2)$, onde μ é a média e σ^2 a variância das proficiências dos indivíduos de Π. Portanto, se desejarmos que os itens sejam estimados na métrica (0,1), deveremos adotar $\mu = 0$ e $\sigma = 1$.

4.4.1 Abordagem de Bock & Aitkin

Uma reformulação da abordagem de Bock & Lieberman, que foi considerada satisfatória do ponto de vista computacional, foi proposta por Bock & Aitkin (1981). Esta reformulação teve como base a suposição de que os itens são independentes, de forma que

$$\frac{\partial^2 \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_l} = 0, \quad \text{para } i \neq l. \tag{4.47}$$

Essa suposição modifica a matriz $\boldsymbol{H}_{PI}(\boldsymbol{\zeta})$ tornando-a bloco diagonal, uma situação similar ao exposto em Andrade, Tavares & Vale (2000), onde eram estimados os parâmetros dos itens e as proficiências conjuntamente. Naquele caso, a independência local foi suficiente para garantir (4.47) e, assim, possibilitar que os itens fossem estimados individualmente, fixadas as proficiências. A proposta de Bock & Aitkin foi adotar a independência entre os itens de forma a possibilitar que os itens sejam estimados individualmente. Vale notar que as suposições de independência local e a suposição de independência dos itens são completamente diferentes. A primeira está relacionada às respostas dos indivíduos, enquanto a segunda se refere apenas aos itens.

Com esta construção, a estimação pode ser feita adotando as mesmas expressões desenvolvidas por Andrade, Tavares & Vale (2000), fazendo a adaptação devida a (4.47).

Entretanto, Bock & Aitkin sugerem que a obtenção das estimativas de máxima verossimilhança seja feita através do algorítimo EM, introduzido por Dempster, Laird & Rubin (1977), e por isso algumas alterações nas expressões citadas anteriormente serão necessárias. Bock & Lieberman assumiram que, de acordo com as notações acima, a probabilidade marginal de $U_{j.}$ é dada por

$$\begin{aligned} P(u_{j.}|\boldsymbol{\zeta}, \boldsymbol{\eta}) &= \int_{\Re} P(u_{j.}|\theta, \boldsymbol{\zeta}, \boldsymbol{\eta}) g(\theta|\boldsymbol{\eta}) d\theta && (4.48) \\ &= \int_{\Re} P(u_{j.}|\theta, \boldsymbol{\zeta}) g(\theta|\boldsymbol{\eta}) d\theta && (4.49) \\ & && (4.50) \end{aligned}$$

onde na última igualdade usamos que a distribuição de $\boldsymbol{U}_{j.}$ não é função de $\boldsymbol{\eta}$. Pela independência entre as respostas de diferentes indivíduos, podemos escrever a verossimilhança como:

$$\begin{aligned} L(\boldsymbol{\zeta}, \boldsymbol{\eta}) &= P(u_{..}|\boldsymbol{\zeta}, \boldsymbol{\eta}) && (4.51) \\ &= \Pi_{j=1}^{n} P(u_{j.}|\boldsymbol{\zeta}, \boldsymbol{\eta}) && (4.52) \\ & && (4.53) \end{aligned}$$

Em Andrade, Tavares & Valle (2000) pode-se ver que usando a abordagem de Padrões de Resposta, irá facilitar os cálculos computacionais. Como temos I itens no total, com 2 possíveis respostas para cada item (0 ou 1), há $S = 2^I$ possíveis respostas (padrões de resposta). Quando o número de indivíduos é grande com relação ao número de itens, pode haver vantagens computacionais em trabalhar com o número de ocorrências dos diferentes padrões de resposta. Neste sentido, daqui em diante vamos trabalhar considerando este raciocínio. O índice já não mais representará um indivíduo, mas sim um padrão de resposta. Seja r_j o número de ocorrências distintas do padrão de resposta j, do tipo $(u_{j1}, ..., u_{jn})$, e ainda $s \leq$min(n, S) o número de padrões de respostas com $r_j > 0$. Segue disso que

$$\sum_{j=1}^{s} r_j = n \tag{4.54}$$

Pela independência entre as respostas dos diferentes indivíduos, temos que os dados seguem uma distribuição Multinomial, isto é,

$$\text{Ł}(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \frac{n!}{\Pi_{j=1}^{s} r_j} \Pi_{j=1}^{s} [P(u_{j.}|\boldsymbol{\zeta}, \boldsymbol{\eta})]^{r_j}. \tag{4.55}$$

Igualando-se a derivada da log-verossimilhança, com relação a $\boldsymbol{\zeta}_i$, ao vetor nulo, obtém-se as equações de estimação para os parâmetros dos itens: a_i, b_i e c_i, respectivamente,

$$\frac{\partial \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial a_i} = D(1-c_i)\Sigma_{j=1}^{s} r_j \int_{\Re} [(u_{ji} - P_i)(\theta - b_i)] W_i g_j^*(\theta) d\theta = 0, \quad (4.56)$$

$$\frac{\partial \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial b_i} = -Da_i(1-c_i)\Sigma_{j=1}^{s} r_j \int_{\Re} [(u_{ji} - P_i)] W_i g_j^*(\theta) d\theta = 0, \quad (4.57)$$

$$\frac{\partial \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial c_i} = \Sigma_{j=1}^{s} r_j \int_{\Re} [(u_{ji} - P_i)] \frac{W_i}{P_i} g_j^*(\theta) d\theta = 0. \quad (4.58)$$

as quais não possuem solução explícita, sendo portanto necessário métodos iterativos para resolvê-las. As equações envolvem integrais que não apresentam soluções analíticas e devem ser resolvidas por algum método de quadratura. Na TRI costuma-se usar o método Hermit-Gauss [Ver Andrade, Tavares & Valle (2000) e Azevedo (2003)]. Considerando conhecidos os nós $\overline{\theta}_k$ e os pesos, A_k, $k = 1, \cdots, q$, temos que as equações de estimação em forma de quadratura para os parâmetros a_i, b_i e c_i são, respectivamente,

$$a_i \; : \; D(1-c_i) \sum_{k=1}^{q} (\overline{\theta}_k - b_i) [r_{ki} - P_{ki} f_{ki}] W_{ki} = 0, \quad (4.59)$$

$$b_i \; : \; -Da_i(1-c_i) \sum_{k=1}^{q} [r_{ki} - P_{ki} f_{ki}] W_{ki} = 0, \quad (4.60)$$

$$c_i \; : \; \sum_{k=1}^{q} [r_{ki} - P_{ki} f_{ki}] \frac{W_{ki}}{P_{ki}^*} = 0. \quad (4.61)$$

onde

$$r_{ki} = \sum_{j=1}^{s} r_j u_{ji} g_{jk}^*, \quad f_{ki} = \sum_{j=1}^{s} r_j g_{jk}^* \quad \text{e} \quad g_{jk}^* = g_j^*(\overline{\theta}_k). \quad (4.62)$$

Essas equações são resolvidas aplicando-se o algoritmo EM Dempster, Laird & Rubin (1977) que é um processo iterativo realizado em dois passos: Esperança (E) e Maximização (M). Mais especificamente, os passos E e M são

Passo E: Usar os pontos de quadratura $\overline{\theta}_k$, os pesos A_k, $k = 1, ..., q$ e estimativas iniciais dos parâmetros dos itens, $\widehat{\boldsymbol{\zeta}}_i, i = 1, ..., I$, para gerar $g_j^*(\theta_k)$ e, posteriormente, r_{ki} e $f_k, i = 1, ..., n$ e $k = 1, ...q$.

Passo M Com $r = (r_{ki})$ e $f = (f_k)$ obtidos no Passo E, resolver as equações de estimação para $\boldsymbol{\zeta}_i, i = 1, ..., n$, usando o algoritmo Newton-Raphson ou "Scoring" de Fisher.

Estes passos compõem cada iteração do algoritmo EM, as quais serão repetidas até que algum critério de parada seja alcançado. Após a finalização do processo, os erros-padrão são obtidos com o uso da matriz de informação de Fisher. Para mais informações sobre o algoritmo EM [ver Andrade, Tavares & Valle (2000) e Azevedo (2003)]. Concluída a fase de estimação dos parâmetros dos itens, inicia-se a etapa de estimação das proficiências individuais. Nesta etapa considera-se que os parâmetros dos itens já são conhecidos, e são adotadas as estimativas usadas na primeira etapa. Na próxima seção aborda-se os principais métodos dessa estimação.

4.4.2 Estimação das proficiêncas

A estimação das proficiências é feita em uma segunda etapa, considerando os parâmetros dos itens fixos. Através da suposição de independência entre as proficiências de diferentes indivíduos, pode-se fazer as estimações em separado para cada indivíduo. Desta forma, a estimação da proficiência do indivíduo j é baseada na função,

$$g_j^*(\theta_j) \equiv g(\theta_j|u_{j.}, \boldsymbol{\zeta}, \boldsymbol{\eta}) \propto P(u_{j.}|\theta_j, \boldsymbol{\zeta})g(\theta_j|\boldsymbol{\eta}). \tag{4.63}$$

Novamente, podemos adotar alguma característica de $g_j^*(\theta_j)$ como estimador de θ_j, sendo que as mais adotadas são a média e a moda. A seguir, trataremos unicamente da obtenção das proficiências pela média da posteriori, por ser o método mais aplicado na área de avaliação educacional.

Estimação pela média da posteriori - EAP

A estimação de θ_j pela média da posteriori (ou EAP: expected a posteriori) consiste em obter a esperança da posteriori, que pode ser escrita como

$$g(\theta|u_{j.}, \boldsymbol{\zeta}, \boldsymbol{\eta}) = \frac{P(u_{j.}|\boldsymbol{\theta}, \boldsymbol{\zeta})g(\theta|\boldsymbol{\eta})}{P(u_{j.}|\zeta, \boldsymbol{\eta})}. \tag{4.64}$$

Segue que a esperança da posteriori é

$$\widehat{\theta}_j \equiv E[\theta|u_{j.}, \boldsymbol{\zeta}, \boldsymbol{\eta}] = \frac{\int_{\Re} \theta g(\theta|\boldsymbol{\eta})P(u_{j.}|\theta, \boldsymbol{\zeta})d\theta}{\int_{\Re} g(\theta|\boldsymbol{\eta})P(u_{j.}|\theta, \boldsymbol{\zeta})d\theta}. \tag{4.65}$$

Esta forma de estimação tem a vantagem de ser calculada diretamente, não necessitando da aplicação de métodos iterativos. Além disso, as quantidades necessárias para o seu cálculo são um produto final da etapa de estimação. Por conta disso alguns autores [por exemplo, Mislevy & Stocking (1989)] recomendam esta escolha para a estimação das proficiências. Algumas vantagens e desvantagens dos métodos citados a cima são descritos em Andrade, Tavares & Vale (2000).

4.5 Estimação: duas ou mais populações

Como descrito nas seções anteriores, é freqüente a situação em que temos duas ou mais populações envolvidas na análise. Estas populações podem ser caracterizadas por diferentes anos, graus de escolaridade, região, sexo, tipo de escola, etc. O primeiro passo para que os resultados relativos às várias populações possam ser comparáveis é a exigência de itens comuns nos testes aplicados a estas populações, criando uma estrutura de ligação entre as mesmas. Nessa situação, o procedimento usual é fazer a estimação para cada população e utilizar uma das técnicas de equalização que serão descritas no próximo capítulo.

Uma abordagem alternativa é o Modelo para Várias Populações proposto por Bock & Zimowski (1997), exposto na Seção 2.3, que representou um grande avanço na TRI. Nesse modelo considera-se que há K populações independentes em estudo e é feita uma análise conjunta das respostas amostrais dessas populações. Considera-se que a distribuição da proficiência dos indivíduos da população k segue uma determinada distribuição com vetor de parâmetros $\boldsymbol{\eta}_k$. Frequentemente adota-se a distribuição Normal com $\boldsymbol{\eta}_k = (\mu_k, \sigma_k^2)'$, sendo que estes parâmetros representam, respectivamente, a média e a variância das proficiências da população k, $k = 1, \cdots, K$.

A grande vantagem da abordagem de Bock & Zimowski está no fato que a equalização é feita automaticamente no próprio processo de estimação. Desta forma, não estamos mais sujeitos a diferenças nas estimativas dos parâmetros devidas ao método de equalização escolhido. Além disso, na presença de várias populações (digamos, $K \geq 5$), com as equalizações sendo feitas entre os testes k e $k+1$, temos erros (relativos à regressão, por exemplo) associados a cada equalização entre duas populações, que serão acumulados para a estimação de (μ_2,σ_2^2), (μ_3,σ_3^2), $\cdots$, e principalmente de (μ_K,σ_K^2), podendo levar a uma má estimação destes parâmetros. Além disso, essa abordagem requer um número menor de itens comuns, em comparação com outros métodos, para produzir resultados similares, conforme será discutido no próximo capítulo.

Sejam u_{kji} a resposta (binária) ao item i oriunda do j-ésimo indivíduo do grupo k, e θ_{kj} a proficiência do j-ésimo indivíduo do grupo k. (Por *grupo* k entenderemos a amostra relativa à população k.) Embora no desenvolvimento que segue a função de resposta possa assumir qualquer uma das formas descritas no Capítulo 2, para fins de aplicação utilizaremos a função ML3, que tem sido a função mais utilizada pelos pesquisadores da área, dada abaixo

$$P(u_{kji} = 1|\theta_{kj}) = c_i + (1 - c_i)\frac{1}{1 + e^{-Da_i(\theta_{kj} - b_i)}}. \tag{4.66}$$

Algumas suposições serão necessárias para a construção do modelo. Além da independência local, assumiremos que as respostas oriundas de indivíduos diferentes serão independentes. Vamos considerar a mesma função de resposta para todos os itens.

4.6 Notações e definições

Embora tenhamos K testes, devemos notar que alguns itens estarão em dois ou mais testes. Por conta disso vamos fazer uma ordenação nos I itens que compõem o conjunto dos K testes, representando-os por $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1, \cdots, \boldsymbol{\zeta}_I)$ e denotando por I_k o conjunto dos índices dos itens aplicados ao grupo k. Considerando I_k o número de itens no teste k, teremos que $I \leq \sum_{k=1}^{K} I_k$. Sejam n_k o número de indivíduos do grupo k e n o número total de indivíduos na amostra. Sejam ainda, $\boldsymbol{U}_{kj.} = (U_{kj1}, U_{kj2}, \cdots, U_{kjI_k})$ o vetor aleatório de respostas do indivíduo j do grupo k; $\boldsymbol{U}_{k..} = (\boldsymbol{U}_{1..}, \boldsymbol{U}_{2..}, \cdots, \boldsymbol{U}_{n_k..})$ o vetor aleatório de respostas do grupo k e $\boldsymbol{U}_{...} = (\boldsymbol{U}_{1..}, \boldsymbol{U}_{2..}, \cdots, \boldsymbol{U}_{n..})$ o vetor total de respostas. De forma similar, representaremos as respostas observadas por $\boldsymbol{u}_{kji}, \boldsymbol{u}_{kj.}, \boldsymbol{u}_{k..}$ e $\boldsymbol{u}_{...}$. Com esta notação e a independência local, podemos escrever a probabilidade associada ao vetor de respostas $\boldsymbol{U}_{kj.}$ como

$$P(\boldsymbol{u}_{kj.}|\theta_{kj}, \boldsymbol{\zeta}) = \prod_{i \in I_k} P(u_{kji}|\theta_{kj}, \boldsymbol{\zeta}_i). \tag{4.67}$$

Pelo exposto anteriormente, o método da Máxima Verossimilhança Marginal, bem como o Bayesiano, têm sido preferidos ao método da Máxima Verossimilhança Conjunta para a estimação dos parâmetros de interesse. Além disso, o fato de podermos associar distribuições para a proficiência da população em estudo nos permite criar estruturas para os parâmetros das respectivas funções densidade de probabilidade, que serão fundamentais nesse modelo. De forma geral, consideremos que as proficiências dos indivíduos da população k, $\theta_{jk},\ j = 1, \cdots, n_k$, são realizações de uma variável aleatória, θ_k, com distribuição contínua e função densidade de probabilidade $g(\theta|\boldsymbol{\eta}_k)$, duplamente diferenciável, com as componentes de $\boldsymbol{\eta}_k$ conhecidas e finitas. Para o caso em que θ_k tem distribuição Normal, temos $\boldsymbol{\eta}_k = (\mu_k, \sigma_k^2)$, onde μ_k é a média e σ_k^2 a variância das proficiências dos indivíduos da população k, $k = 1, \cdots, K$.

Na situação em que temos uma única população em estudo, não há necessidade de estimação dos parâmetros populacionais. Isso ocorre porque a métrica é estabelecida fixando-se os parâmetros populacionais, geralmente em $\mu = 0$ e $\sigma = 1$, onde μ é a média e σ é o desvio-padrão das proficiências da população considerada. Na presença de várias populações, temos mais um conjunto de parâmetro a estimar: $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, \cdots, \boldsymbol{\eta}_K)$, que serão referidos como *Parâmetros Populacionais*. Entretanto, ainda há a necessidade do estabelecimento da métrica e isso pode ser resolvido fixando-se os parâmetros relativos a qualquer uma das populações. Neste trabalho adotaremos a seguinte referência:

$$\mu_1 = 0, \qquad \sigma_1 = 1. \tag{4.68}$$

Logo, resta apenas a estimação de $\boldsymbol{\eta}_2, \cdots, \boldsymbol{\eta}_K$. Novamente, a estimação neste modelo é feita por máxima verossimilhança marginal, com o diferencial que a primeira etapa envolve a estimação dos parâmetros dos itens e dos parâmetros populacionais; as proficiências individuais são estimadas na segunda etapa. Cabe notar aqui uma grande contribuição do modelo de Bock & Zimowski, a de que as médias populacionais podem ser estimadas diretamente, ao passo que o procedimento anterior era fazer a estimação das proficiências para cada grupo, adotar um método de equalização para colocá-las na mesma escala e, finalmente, obter a média amostral das proficiências de cada grupo.

4.7 Estimação dos parâmetros dos itens

Com as notações definidas acima, temos que a probabilidade marginal de $\boldsymbol{U}_{kj.}$ é dada por

$$P(\boldsymbol{u}_{kj.}|\boldsymbol{\zeta}, \boldsymbol{\eta}_k) = \int_{\Re} P(\boldsymbol{u}_{kj.}|\theta, \boldsymbol{\zeta}, \boldsymbol{\eta}_k) g(\theta|\boldsymbol{\eta}_k) d\theta = \int_{\Re} P(\boldsymbol{u}_{kj.}|\theta, \boldsymbol{\zeta}) g(\theta|\boldsymbol{\eta}_k) d\theta, \tag{4.69}$$

onde na última igualdade usamos que a distribuição de $\boldsymbol{U}_{kj.}$ não é função de $\boldsymbol{\eta}_k$.

Usando a independência entre as respostas de diferentes indivíduos, podemos escrever a probabilidade associada ao vetor de respostas $\boldsymbol{U}_{...}$ como

$$P(\boldsymbol{u}_{...}|\boldsymbol{\zeta}, \boldsymbol{\eta}) = \prod_{k=1}^{K} \prod_{j=1}^{n_k} P(\boldsymbol{u}_{kj.}|\boldsymbol{\zeta}, \boldsymbol{\eta}_k). \tag{4.70}$$

Embora a verossimilhança possa ser escrita como (4.70), tem sido freqüente utilizar a abordagem de *Padrões de Resposta*. Como temos I_k itens no teste k, com duas possíveis respostas para cada item (0 ou 1), há $S_k = 2^{I_k}$ possíveis respostas (padrões de resposta) associados a esse teste. Seja r_{kj} o número de ocorrências distintas do padrão de resposta j no grupo k, e ainda $s_k \leq \min(n_k, S_k)$ o número de padrão de respostas com $r_{kj} > 0$. Segue que

$$\sum_{j=1}^{s_k} r_{kj} = n_k. \tag{4.71}$$

Pela independência entre as respostas dos diferentes indivíduos, temos que os dados seguem uma distribuição $Produto - Multinomial$, isto é,

$$L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \prod_{k=1}^{K} \left\{ \frac{n_k!}{\prod_{j=1}^{s_k} r_{jk}!} \prod_{j=1}^{s_k} [P(\boldsymbol{u}_{jk.}|\boldsymbol{\zeta}, \boldsymbol{\eta}_k)]^{r_{jk}} \right\}. \tag{4.72}$$

E, portanto, a log-verossimilhança é

$$\log L(\boldsymbol{\zeta}, \boldsymbol{\eta}) = \sum_{k=1}^{K} \log \left\{ \frac{n_k!}{\prod_{j=1}^{s_k} r_{jk}!} \right\} + \sum_{k=1}^{K} \sum_{j=1}^{s_k} r_{jk} \log P(\boldsymbol{u}_{jk.}|\boldsymbol{\zeta}, \boldsymbol{\eta}_k). \tag{4.73}$$

Com os desenvolvimentos que podem ser vistos em Andrade, Tavares & Vale (2000), as equações finais de estimação para os parâmetros dos itens a_i, b_i e c_i são, respectivamente,

$$a_i \ : \ D(1-c_i) \sum_{k=1}^{K} \sum_{j=1}^{s_k} r_{kj} \int_{\Re} [(u_{kji} - P_i)(\theta - b_i) W_i] \, g^*_{kj}(\theta) d\theta = 0, \tag{4.74}$$

$$b_i \ : \ -Da_i(1-c_i) \sum_{k=1}^{K} \sum_{j=1}^{s_k} r_{kj} \int_{\Re} [(u_{kji} - P_i) W_i] \, g^*_{kj}(\theta) d\theta = 0, \tag{4.75}$$

$$c_i \ : \ \sum_{k=1}^{K} \sum_{j=1}^{s_k} r_{kj} \int_{\Re} \left[(u_{kji} - P_i) \frac{W_i}{P^*_i} \right] g^*_{kj}(\theta) d\theta = 0, \tag{4.76}$$

Mais uma vez essas equações não possuem solução explícita.

4.8 Estimação dos parâmetros populacionais

Considerando a log-verossimilhança obtida em (4.73), as equações de estimação para as proficiências médias e variâncias das populações são obtidas por

$$\frac{\partial \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \mu_k} = 0 \qquad \text{e} \qquad \frac{\partial \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \sigma_k^2} = 0, \quad k = 2, \cdots, K.$$

Mas,

$$\begin{aligned}
\frac{\partial \log L(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}_k} &= \sum_{j=1}^{s_k} r_{jk} \frac{1}{P(\boldsymbol{u}_{kj.}|\boldsymbol{\zeta}, \boldsymbol{\eta}_k)} \int_{\Re} P(\boldsymbol{u}_{kj.}|\theta, \boldsymbol{\zeta}) \left(\frac{\partial g(\theta|\boldsymbol{\eta}_k)}{\partial \boldsymbol{\eta}_k} \right) d\theta \\
&= \sum_{j=1}^{s_k} r_{jk} \frac{1}{P(\boldsymbol{u}_{kj.}|\boldsymbol{\zeta}, \boldsymbol{\eta}_k)} \int_{\Re} P(\boldsymbol{u}_{kj.}|\theta, \boldsymbol{\zeta}) \left(\frac{\partial \log g(\theta|\boldsymbol{\eta}_k)}{\partial \boldsymbol{\eta}_k} \right) g(\theta|\boldsymbol{\eta}_k) d\theta \\
&= \sum_{j=1}^{s_k} r_{jk} \int_{\Re} \left(\frac{\partial \log g(\theta|\boldsymbol{\eta}_k)}{\partial \boldsymbol{\eta}_k} \right) g_{kj}^*(\theta) d\theta.
\end{aligned}$$

Se utilizarmos a distribuição $N(\mu_k, \sigma_k^2)$ para θ_k, teremos que

$$\frac{\partial \log g(\theta|\boldsymbol{\eta}_k)}{\partial \mu_k} = \frac{\theta - \mu_k}{\sigma_k^2}$$

e

$$\frac{\partial \log g(\theta|\boldsymbol{\eta}_k)}{\partial \sigma_k^2} = -\frac{\sigma_k^2 - (\theta - \mu_k)^2}{2\sigma_k^4}.$$

Assim, as formas finais das equações de estimação são

$$\begin{aligned}
\mu_k &: \quad (\sigma_k^2)^{-1} \sum_{j=1}^{s_k} r_{kj} \int_{\Re} (\theta - \mu_k) g_{kj}^*(\theta) d\theta = 0, \\
\sigma_{kj}^2 &: \quad -(2\sigma_k^4)^{-1} \sum_{j=1}^{s_k} r_{kj} \int_{\Re} \left[\sigma_k^2 - (\theta - \mu_k)^2 \right] g_{kj}^*(\theta) d\theta = 0. \qquad (4.77)
\end{aligned}$$

Note que, se fizermos

$$\mu_{kj} = \int_{\Re} \theta g_{kj}^*(\theta) d\theta, \sigma_{kj}^2 = \int_{\Re} (\theta - \mu_{kj})^2 g_{kj}^*(\theta) d\theta, \qquad (4.78)$$

que representam a média e a variância da distribuição condicional da proficiência da população k, dado $\{U_{kj.} = u_{kj.}\}$, respectivamente, então, por (4.71), (4.77) e (4.78), segue que

$$\begin{aligned} 0 &= \sum_{j=1}^{s_k} r_{kj} \int_{\Re} \theta g_{kj}^*(\theta) d\theta - \widehat{\mu}_k \sum_{j=1}^{s_k} r_{kj} \int_{\Re} g_{kj}^*(\theta) d\theta \\ &= \sum_{j=1}^{s_k} r_{kj} \widehat{\mu}_{kj} - \widehat{\mu}_k \sum_{j=1}^{s_k} r_{kj} \\ &= \sum_{j=1}^{s_k} r_{kj} \widehat{\mu}_{kj} - n_k \widehat{\mu}_k, \end{aligned} \tag{4.79}$$

de onde conclui-se que

$$\widehat{\mu}_k = \frac{1}{n_k} \sum_{j=1}^{s_k} r_{kj} \widehat{\mu}_{kj}. \tag{4.80}$$

Também, por (4.77) e usando que $\theta - \mu_k = (\theta - \mu_{kj}) + (\mu_{kj} - \mu_k)$, temos

$$\begin{aligned} 0 &= \sum_{j=1}^{s_k} r_{kj} \widehat{\sigma}_k^2 \int_{\Re} g_{kj}^*(\theta) d\theta - \sum_{j=1}^{s_k} r_{kj} \int_{\Re} (\theta - \widehat{\mu}_k)^2 g_{kj}^*(\theta) d\theta \\ &= \widehat{\sigma}_k^2 \sum_{j=1}^{s_k} r_{kj} - \sum_{j=1}^{s_k} r_{kj} \left[\int_{\Re} (\theta - \widehat{\mu}_{kj})^2 g_{kj}^* d\theta + \int_{\Re} (\widehat{\mu}_{kj} - \widehat{\mu}_k)^2 g_{kj}^*(\theta) d\theta \right] \\ &= n_k \widehat{\sigma}_k^2 - \sum_{j=1}^{s_k} r_{kj} \left[\widehat{\sigma}_{kj}^2 + (\widehat{\mu}_{kj} - \widehat{\mu}_k)^2 \right], \end{aligned}$$

de onde concluímos que

$$\widehat{\sigma}_k^2 = \frac{1}{n_k} \sum_{j=1}^{s_k} r_{kj} \left[\widehat{\sigma}_{kj}^2 + (\widehat{\mu}_{kj} - \widehat{\mu}_k)^2 \right]. \tag{4.81}$$

Note que $g_{kj}^*(\theta)$ depende dos parâmetros dos itens e também dos parâmetros populacionais e, consequentemente, seu valor nas expressões acima deve ser calculado a partir de estimativas desses parâmetros.

Representando por $\overline{\mu}_{k.}$ a média das esperanças condicionais μ_{kj}, por $\overline{\sigma}_{k.}^2$ a média das variâncias condicionais σ_{kj}^2 e por $\delta_{k.}^2$ uma medida adequada de variabilidade entre as médias condicionais, todas associadas ao grupo k, ou seja,

$$\overline{\mu}_{k.} = \frac{1}{n_k} \sum_{j=1}^{s_k} r_{kj} \mu_{kj}, \quad \overline{\sigma}_{k.}^2 = \frac{1}{n_k} \sum_{j=1}^{s_k} r_{kj} \sigma_{kj}^2 \quad \text{e} \quad \delta_{k.}^2 = \frac{1}{n_k} \sum_{j=1}^{s_k} r_{kj} (\mu_{kj} - \mu_k)^2,$$

podemos escrever as equações (4.80) e (4.81) como

$$\widehat{\mu}_k = \widehat{\overline{\mu}}_{k.} \quad \text{e} \quad \widehat{\sigma}_k^2 = \widehat{\overline{\sigma}}_{k.}^2 + \widehat{\delta_{k.}}^2, \quad k = 2, \cdots, K. \tag{4.82}$$

Estas expressões nos permitem interpretações bastante intuitivas. Primeiro, notemos que os somatórios nas definições acima podem ser adaptados de forma a considerar as respostas individuais ao invés dos padrões de respostas. Com isso, o estimador para a proficiência média da população k é a média obtida com os estimadores das médias da distribuição condicional da proficiência, dados os vetores de respostas individuais $u_{kj.}$. Por outro lado, o estimador para a variância das proficiências da população k não é simplesmente a média entre estimadores das variâncias da distribuição condicional da proficiência, dados os vetores de respostas individuais $u_{kj.}$. Existe também uma outra contribuição relativa à variabilidade entre os estimadores das médias da distribuição condicional da proficiência com relação ao estimador da média populacional associada.

4.8.1 Estimação conjunta: aplicação do algoritmo EM

Podemos escrever a equação de estimação para $\boldsymbol{\zeta}_i$ como

$$\boldsymbol{\zeta}_i \quad : \quad \sum_{k=1}^{K} \int_{\Re} \left[(r_{ki}(\theta) - P_i f_{ki}(\theta)) \left(\frac{\partial P_i}{\partial \boldsymbol{\zeta}_i} \right) \frac{W_i}{P_i^* Q_i^*} \right] d\theta = 0, \tag{4.83}$$

onde

$$f_{ki}(\theta) = \sum_{j=1}^{s_k} r_{kj} g_j^*(\theta) e \quad r_{ki}(\theta) = \sum_{j=1}^{s_k} r_{kj} u_{kji} g_{kj}^*(\theta)$$

representam, respectivamente, o números de indivíduos do grupo k com proficiência θ respondendo ao item i e o número destes indivíduos que respondem corretamente ao item i. Novamente, as integrais são aproximadas através de quadratura Gaussiana. Fixados os q nós $\overline{\theta}_{kl}$ e os pesos A_{kl}, $l = 1, \cdots, q$, $k = 1, \cdots, K$, e com estimativas iniciais dos parâmetros dos itens, $\widehat{\boldsymbol{\zeta}}_i$, $i = 1, \cdots, I$, as equações (4.82) podem ser resolvidas diretamente para obtenção das estimativas desejadas. A estimação é feita em separado para cada item, e por isso poderemos utilizar o desenvolvimento da Seção 4.2. Reformulando-se os passos do algoritmo EM descritos na Seção 3.5.3, para a situação de duas ou mais populações, teremos

Passo E

1. Usar os pontos de quadratura $\overline{\theta}_{kl}$, os pesos A_{kl}, $l = 1, \cdots, q$ e estimativas iniciais dos parâmetros dos itens , $\boldsymbol{\zeta}_i$, $i = 1, \cdots, I$, e dos parâmetros populacionais, $\boldsymbol{\mu}_k$ e σ_k^2, $k = 1, \cdots, K$, para gerar $g_{kj}^*(\overline{\theta}_{kl})$ e, posteriormente, $\overline{r}_{kli}$ e $\overline{f}_{kli}$, $i = 1, \cdots, I$ e $k = 1, \cdots, q$.

2. Usar os pontos de quadratura e $g_{kj}^*(\overline{\theta}_{kl})$ para obter $\widehat{\mu}_{kj}$ e $\widehat{\sigma}_{kj}^2$ por (4.78) e (4.77), e poteriormente, $\widehat{\mu}_k$ e $\widehat{\sigma}_k^2$ por (4.82).

Passo M Com r, f e η obtidos no Passo E, resolver as equações de estimação para ζ_i, $i = 1, \cdots, I$, usando o algoritmo Newton-Raphson ou "Scoring" de Fisher através das expressões da Seção 5.2.

Estes passos compõem cada iteração do algoritmo EM, as quais serão repetidas até que algum critério de parada seja alcançado. Após a finalização do processo, os erros-padrão são obtidos com o uso de 5.29.

Devemos notar que no passo M as expressões para a maximização são um pouco modificadas, com relação às expressões da Seção 4.2, devido a introdução de novos grupos. Se $\widehat{\boldsymbol{\zeta}}_i^{(t)}$ é uma estimativa de $\boldsymbol{\zeta}_i$ na iteração t, o processo iterativo de Newton-Raphson para obtenção de $\widehat{\boldsymbol{\zeta}}_i^{(t+1)}$ é dado pela expressão (4.7), onde

$$\boldsymbol{h}(\boldsymbol{\zeta}_i) = \sum_{k=1}^{K} \sum_{l=1}^{q} (r_{kli} - f_{kli} P_{kli}) W_{kli} \boldsymbol{h}_{kli},$$

$$\boldsymbol{H}(\boldsymbol{\zeta}_i) = \sum_{k=1}^{K} \sum_{l=1}^{q} (r_{kli} - f_{kli} P_{kli}) W_{kli} \left\{ \boldsymbol{H}_{kli} - (r_{kli} - f_{kli} P_{kli}) W_{kli} \boldsymbol{h}_{kli} \boldsymbol{h}'_{kli} \right\},$$

com $P_{kli}, W_{kli}, \boldsymbol{H}_{kli}$ e $\boldsymbol{h}_{kli}$ similares à Seção 4.2, com θ_k substituída por $\overline{\theta}_{kl}$. Para a aplicação do método "Scoring" de Fisher, devemos substituir $H(\boldsymbol{\zeta}_i)$ pelo seu valor esperado, ou seja,

$$\boldsymbol{\Delta}(\boldsymbol{\zeta}_i) = -\sum_{k=1}^{K} \sum_{l=1}^{q} \left\{ P_{kli}^* Q_{kli}^* W_{kli} \boldsymbol{h}_{kli} \boldsymbol{h}'_{kli} \right\}. \tag{4.84}$$

4.9 Estimação das proficiências

Uma etapa que pode ser considerada opcional é a estimação das proficiências. Talvez o interesse da análise se concentre apenas na estimação dos parâmetros dos itens e populacionais, sem relevar a estimação das proficiências. Em caso contrário, as proficiências podem ser estimadas por MV, como descrito na Seção 4.3, ou de forma bayesiana, como descrito na Seção 4.4.2. Devido a presença de vários grupos na análise, as expressões para obtenção das proficiências são ligeiramente modificadas, e por isso as apresentaremos nesta seção.

Vale ressaltar que em todos os métodos de estimação descritos abaixo, consideraremos fixos os parâmetros dos itens e os populacionais.

4.9.1 Estimação por MV

Neste caso, a estimação das proficiências é feita iterativamente pelo algoritmo Newton-Raphson ou método "Scoring"de Fisher. Considerando $\widehat{\theta}_{kj}^{(t)}$ uma estimativa de θ_{kj} na iteração t, então na iteração $t+1$ teremos que

$$\widehat{\theta}_{kj}^{(t+1)} = \widehat{\theta}_{kj}^{(t)} - [H(\widehat{\theta}_{kj}^{(t)})]^{-1} h(\widehat{\theta}_{kj}^{(t)}), \tag{4.85}$$

onde

$$H(\theta_{kj}) \;=\; \sum_{i \in I_k} (u_{kji} - P_{kji}) W_{kji} \left\{ H_{kji} - (u_{kji} - P_{kji}) W_{kji} h_{kji}^2 \right\},$$

com $P_{kji}, W_{kji}, H_{kji}$ e h_{kji} similares à Seção 4.3, com θ_j substituída por θ_{kj}. Para aplicação do método "Scoring"de Fisher, devemos substituir $H(\theta_{kj})$ pelo seu valor esperado, ou seja,

$$\Delta(\theta_{kj}) \;=\; -\sum_{i \in I_k} P_{kji}^* Q_{kji}^* W_{kji} h_{kji}^2.$$

4.9.2 Estimação por EAP

A estimação de θ_{kj} pela média da posteriori (EAP) consiste em obter a esperança da posteriori, que pode ser escrita como

$$g(\theta | \boldsymbol{u}_{kj.}, \boldsymbol{\zeta}, \boldsymbol{\eta}_k) = \frac{P(\boldsymbol{u}_{kj.} | \theta, \boldsymbol{\zeta}) g(\theta | \boldsymbol{\eta}_k)}{P(\boldsymbol{u}_{kj.} | \boldsymbol{\zeta}, \boldsymbol{\eta}_k)}. \tag{4.86}$$

Segue que a esperança da posteriori é

$$\widehat{\theta}_{kj} \equiv E[\theta | \boldsymbol{u}_{kj.}, \boldsymbol{\zeta}, \boldsymbol{\eta}_k] = \frac{\int_{\Re} \theta g(\theta | \boldsymbol{\eta}_k) P(\boldsymbol{u}_{kj.} | \theta, \boldsymbol{\zeta}) d\theta}{\int_{\Re} g(\theta | \boldsymbol{\eta}_k) P(\boldsymbol{u}_{kj.} | \theta, \boldsymbol{\zeta}) d\theta}. \tag{4.87}$$

Esta forma de estimação tem a vantagem de ser calculada diretamente, não necessitando da aplicação de métodos iterativos. Além disso, as quantidadas necessárias para o seu cálculo são um produto final da etapa de estimação. Por conta disso alguns autores por exemplo, Mislevy & Stocking (1989) recomendam esta escolha para a estimação das proficiências.

Mais uma vez vale ressaltar que maiores detalhes para os desenvolvimentos aqui apresentados podem ser encontrados em Andrade, Tavares & Vale (2000).

Capítulo 5

Equalização

5.1 Introdução

Em primeiro lugar, é importante definir o conceito de *Equalização*, que é um dos mais importantes da TRI e um dos grandes objetivos das Avaliações Educacionais. Equalizar significa equiparar, tornar comparável, o que no caso da TRI significa colocar parâmetros de itens vindos de provas distintas ou proficiências de respondentes de diferentes grupos, na mesma métrica, isto é, numa escala comum, tornando os itens e/ou as proficiências comparáveis.

Existem dois tipos de equalização: a equalização *via população* e a equalização *via itens comuns*. Isto significa que há duas maneiras de colocar parâmetros, tanto de itens quanto de proficiências, numa mesma métrica: na primeira usamos o fato de que se um único grupo de respondentes é submetido a provas distintas, basta que todos os itens sejam calibrados conjuntamente para termos a garantia de que todos estarão na mesma métrica. Já na equalização via itens comuns, a garantia de que as populações envolvidas terão seus parâmetros em uma única escala será dada pelos itens comuns entre as populações, que servirão de ligação entre elas.

No capítulo anterior, apresentamos os métodos de estimação mais utilizados quando *todos* os parâmetros dos itens de uma *única* prova devem ser estimados. No entanto, esta é apenas uma das possíveis situações que na prática iremos encontrar. A seguir, listaremos os 6 casos possíveis, quanto ao número de grupos e de tipos de prova envolvidos.

1. Um único grupo fazendo uma única prova.

2. Um único grupo, dividido em dois subgrupos, fazendo duas provas, totalmente distintas (nenhum item comum).

3. Um único grupo, dividido em dois subgrupos, fazendo duas provas, apenas parcialmente distintas, ou seja, com alguns itens comuns.

4. Dois grupos fazendo uma única prova.

5. **Dois grupos fazendo duas provas, totalmente distintas (nenhum item comum).**

6. Dois grupos fazendo duas provas, apenas parcialmente distintas, ou seja, com alguns itens comuns.

Note que para simplificar, listamos os casos acima utilizando apenas duas provas e duas populações, mas as situações envolvendo um número maior de provas e/ou de populações são análogas.

Além disso, os problemas de estimação também podem diferir dependendo do conjunto de itens que necessita ser estimado, ou seja, se nosso conjunto de itens é composto de:

1. **apenas itens novos (ou seja, itens que ainda não foram calibrados)*;**

2. apenas itens já calibrados;

3. itens novos e itens calibrados.

5.2 Tipos de equalização

Uma vez listadas as diversas situações e casos que podemos ter, vamos fazer uma abordagem apenas à situação que envolve o nosso trabalho. Obviamente, pode-se ter as situações 1 a 6 combinadas com os casos (a) a (c). Para facilitar a explicação, trataremos aquí apenas da situação 5 considerando o caso mais simples, ou seja, o caso (a).

Cabe ressaltar que as análises e comentários deste capítulo serão feitas considerando-se o modelo logístico unidimensional de 3 parâmetros. Maiores detalhes podem ser vistos em Andrade, Tavares & Vale (2000).

5.2.1 Dois grupos fazendo duas provas totalmente distintas

Este é o único dos 6 casos que não pode ser resolvido pela TRI. Obviamente é possível calibrar separadamente os itens das duas provas, mas o problema é que *não* podemos

* Para este trabalho as situações em que nos encontramos estão em negrito.

fazer nenhum tipo de comparação entre os resultados obtidos, uma vez que eles estarão em métricas *diferentes*. Neste caso, não faz sentido comparar os resultados destes dois grupos, assim como não faz sentido comparar diretamente 40^oC com 40^oF . Assim como essas duas temperaturas estão em escalas de medida diferentes, os parâmetros obtidos nestas duas provas também estarão. A diferença é que, no caso das temperaturas, há uma relação conhecida entre as duas escalas, e assim, é possível colocarmos uma das temperaturas na mesma escala que a outra, possibilitando então, a comparação. Já no caso das provas, não existe nenhuma relação entre elas e nem entre os dois grupos, que torne possível a comparação.

Um exemplo que ilustra esta situação seria as provas do ENEM que são elaboradas em cada ano de forma distinta. Estas provas poderiam ser calibradas separadamente e seus resultados poderiam ser interpretados isoladamente dentro de cada ano, mas não poderíamos comparar os resultados dos itens e nem das proficiências estimadas para os indivíduos dos diferentes anos.

5.3 Diferentes problemas de estimação

Vamos agora considerar outro ponto bastante importante na TRI: o conjunto de itens a ser calibrado. Vamos comentar inicialmente os casos (a) a (c), considerando-se o caso 1, ou seja, o caso em que uma única prova foi aplicada a um único grupo de respondentes.

5.3.1 Quando todos os itens são novos

Neste caso, todos os itens são considerados “novos”, ou seja, deseja-se calibrar o conjunto completo de itens. Este é o caso trivial, que foi considerado até agora. Para resolver este problema basta utilizar alguma das técnicas de estimação descritas no capítulo anterior.

5.3.2 Quando todos os itens já estão calibrados

Este é o caso em que todos os itens já foram calibrados anteriormente, ou seja, quando não desejamos calibrar nenhum dos itens e estamos interessados apenas em estimar as proficiências dos respondentes. Este é um caso também bastante frequente na TRI, devido ao impulso que esta teoria deu na criação de *bancos de itens*. Tais bancos são formados por conjuntos de itens que já foram testados e calibrados a partir de um número significativo de indivíduos de uma dada população. Desta maneira, assumimos que os parâmetros desses itens já são "conhecidos", ou seja, assumimos que conhecemos os verdadeiros valores dos parâmetros desses itens e assim, sempre que desejarmos, podemos aplicar novamente alguns desses itens do banco a outros indivíduos (ou até mesmo a um *único* indivíduo) e poderemos então estimar apenas suas proficiências, que estarão sempre na mesma métrica dos parâmetros dos itens.

A questão da métrica é um ponto que deve ser considerado com bastante cuidado numa situação como esta. Quando se "constrói"um banco de itens, uma informação fundamental é a *escala* em que aqueles itens foram calibrados. Isto porque as proficiências de indivíduos que serão estimadas futuramente a partir daqueles itens estarão nesta mesma métrica e portanto, quaisquer comparações diretas só poderão ser feitas com outros sujeitos que também tenham suas proficiências nesta escala.

Assim, para resolver este problema, basta utilizar um dos processos de estimação das proficiências dos indivíduos quando os parâmetros dos itens já são conhecidos, que foram descritos na Seção 4.3 do Capítulo 4.

5.3.3 Quando alguns itens são novos e outros já estão calibrados

Neste caso, temos itens "novos"e itens já calibrados, ou seja, desejamos calibrar alguns itens e manter os parâmetros de outros, que já foram calibrados anteriormente. Este também é uma situação que está tipicamente ligada à criação de *bancos de itens*. Isto porque um banco de itens está continuamente em formação, ou seja, é bastante comum estarmos interessados em acrescentar novos itens ao conjunto que já se encontra no banco (assim como também é comum a retirada de itens do banco). Neste caso, o problema fundamental é garantir que os itens novos sejam calibrados na mesma métrica em que estão os outros itens do banco.

Na prática, este é um problema de solução mais complexa do que possa parecer em princípio. Isto porque é indispensável o uso de programas computacionais especificamente desenvolvidos para a análise de itens via TRI e esses programas ainda apresentam algumas dificuldades com relação a situações como essa.

5.4 Equalização a posteriori

Até aqui discutimos formas de equalização entre 2 ou mais populações feitas *durante* o próprio processo de estimação dos parâmetros. Mas, também é possível fazer a equalização *a posteriori*, isto é, depois de terminado o processo de calibração dos itens. Basicamente, a equalização a posteriori é feita da seguinte maneira: calibra-se separadamente os dois conjuntos de itens, que foram submetidos às duas populações de interesse. Obviamente, a condição necessária é que hajam itens comuns entre os dois conjuntos. Assim, para os itens comuns, teremos dois conjuntos de estimativas, cada uma na métrica de suas respectivas populações. Daí, através dessas duas estimativas para os itens comuns estabelece-se algum tipo de *relação* que permita colocarmos os parâmetros de um dos conjuntos de itens na escala do outro. Com todos os itens na mesma métrica, pode-se então estimar as proficiências de todos os respondentes, que então estarão também na mesma escala.

Pela propriedade de invariância, já discutida no Capítulo 4, dado que o modelo é adequado aos dados, os parâmetros a e b de um certo item apresentado a 2 grupos de respondentes devem satisfazer, a menos de flutuações amostrais, as seguintes relações lineares:

$$b_{G1} = \alpha b_{G2} + \beta \quad \text{e} \quad a_{G1} = \frac{1}{\alpha} a_{G2}, \tag{5.1}$$

onde b_{G1} e b_{G2} são os valores do parâmetro de dificuldade e a_{G1} e a_{G2} são os valores do parâmetro de discriminação nos grupos 1 e 2, respectivamente. Uma vez determinados os coeficientes α e β , as estimativas dos parâmetros dos itens do grupo 2 podem facilmente ser colocados na mesma escala das estimativas do grupo 1.

Vários métodos, que se baseiam nessas relações lineares existentes entre os parâmetros de um mesmo item medidos em escalas diferentes, poderiam ser então utilizados para determinar os coeficientes α e β. A solução mais natural — pelo próprio tipo de relação existente entre os parâmetros — seria determinar esses coeficientes através de uma regressão linear simples. No entanto, a crítica feita à utilização desse método é que ele *não é simétrico*, ou seja, uma regressão de x por y é diferente de uma regressão de y por x.

Um dos métodos de equalização a posteriori existentes que não apresenta esse problema, ou seja, é invariante (simétrico) em relação às variáveis utilizadas, é denominado *Média-Desvio (Mean-Sigma)*. O método Média-Desvio utiliza:

$$\alpha = \frac{S_{G1}}{S_{G2}} \quad \text{e} \quad \beta = M_{G1} - \alpha M_{G2}, \tag{5.2}$$

onde S_{G1} e S_{G2} são os desvios-padrão e M_{G1} e M_{G2} as médias amostrais das estimativas dos parâmetros de dificuldade dos itens comuns nos grupos 1 e 2, respectivamente. Da mesma forma, as proficiências dos respondentes do grupo 2 podem ser colocadas na mesma escala das proficiências dos respondentes do grupo 1 a partir da relação

$$\theta^1_{G2} = \alpha\theta_{G2} + \beta, \tag{5.3}$$

onde θ^1_{G2} é o valor da proficiência θ_{G2} na escala do grupo 1. Maiores detalhes sobre este e outros métodos de equalização, como por exemplo *Média-Desvio Robusto* e *Curva Característica*, podem ser encontrados em Kolen & Brennan (1995).

Uma última observação sobre equalização deve ser feita com relação à *quantidade* de itens comuns. Certamente, quanto maior o número de itens comuns, melhor será a qualidade da equalização. Assim, o melhor caso de equalização entre dois grupos distintos é a situação da Seção 6.2.4, ou seja, quando trata-se exatamente da mesma prova. No entanto, já sabemos que não é necessário que todos os itens sejam comuns. O número mínimo de itens comuns necessário para uma boa equalização entre duas populações depende basicamente de dois fatores: do tipo de equalização que será feita e da "qualidade"desses itens comuns.

Equalizações feitas durante o processo de calibração, com os modelos para duas ou mais populações, são mais "eficazes"e portanto, exigem um número menor de itens comuns do que equalizações feitas a posteriori. Além disso, se os itens comuns utilizados na equalização tiverem níveis de dificuldade baixos ou altos demais com relação às populações envolvidas, ou então se apresentarem baixo poder de discriminação, haverá necessidade de um número maior de itens.

Alguns autores têm sugerido pelo menos 6 itens comuns entre 2 provas de 30 itens, quando a equalização é feita durante a calibração. Um estudo de simulação considerando diferentes situações de equalização pode ser encontrado em Andrade (1999).

5.5 A Escala de proficiência

Diferentemente da medida escore em um teste com I questões do tipo certo/errado, que assume valores inteiros entre 0 e I, na TRI a proficiência pode teoricamente assumir qualquer valor real entre $-\infty$ e $+\infty$. Assim, precisa-se estabelecer uma origem e uma unidade de medida para a definição da escala. Esses valores são escolhidos de modo a representar, respectivamente, o valor médio e o desvio-padrão das proficiências dos indivíduos da população em estudo. Para o gráfico que será mostrado mais adiante, utilizou-se a escala com média igual a 0 e desvio-padrão igual a 1, que será representada por escala (0,1). Essa escala é bastante utilizada pela TRI, e nesse caso, os valores do parâmetro b variam (tipicamente) entre -4 e +4. Com relação ao parâmetro a, espera-se valores entre 0 e +3, sendo que os valores mais apropriados de a seriam aqueles maiores do que 0,45 para o modelo normal e 0,75 para o modelo logístico.

Apesar da frequente utilização da escala (0,1), em termos práticos, não faz a menor diferença estabelecer-se estes valores ou outros quaisquer. O importante são as relações de ordem existentes entre seus pontos. Por exemplo, na escala (0,1) um indivíduo com proficiência 1,20 está 1,20 desvios-padrão acima da proficiência média. Este mesmo indivíduo teria a proficiência 620, e consequentemente estaria também 1,20 desvios-padrão acima da proficiência média, se a escala utilizada para esta população fosse a escala(500,100). Isto pode ser visto a partir da transformação de escala:

$$a(\theta - b) = (a/100)[(100 \times \theta + 500) - (100 \times b + 500)] = a^*(\theta^* - b^*),$$

onde $a(\theta - b)$ é a parte do modelo probabilístico proposto envolvida na transformação. Assim, tem-se que:

1. $\theta^* = 100 \times \theta + 500$,
2. $b^* = 100 \times b + 500$,
3. $a^* = a/100$,
4. $P(U_i = 1|\theta) = P(U_i = 1|\theta^*)$.

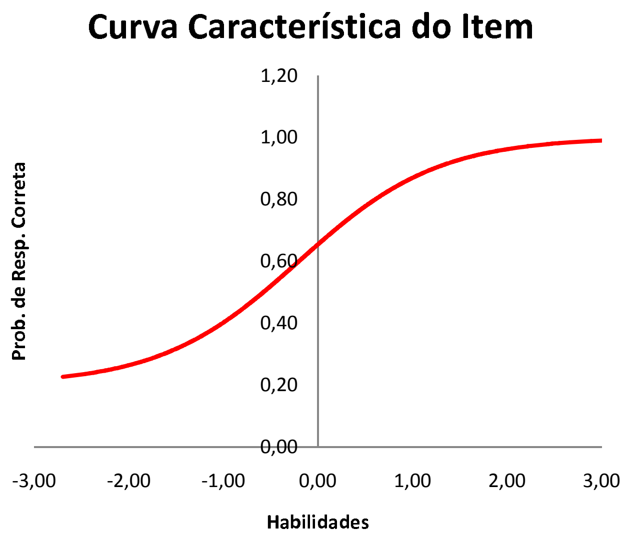

Figura 5.1 *Curvas características do item*

Por exemplo, os valores dos parâmetros a e b da Figura 5.1, na escala (0,1) são, respectivamente, 0,80 e -0,20 e seus correspondentes na escala(500;100) são, respectivamente, 0,008 = 0,80 / 100 e 480 = 100 × (-0,20) + 500. Além disso, um indivíduo com proficiência $\theta = 1,00$ medida na escala (0,1) tem sua proficiência representada por $\theta^* =$ 100 × 1,00 + 500 = 600 na escala(500;100) e

$$\begin{aligned} P(U_1 = 1|\theta = 1) &= 0,20 + (1 - 0,20)\frac{1}{1 + e^{-1,7 \times 0,80 \times (1 - (-0,20))}} \\ &= 0,20 + (1 - 0,20)\frac{1}{1 + e^{-1,7 \times 0,008 \times (600 - 480)}} \\ &= P(U_1 = 1|\theta^* = 600) = 0,87, \end{aligned}$$

ou seja, a probabilidade de um indivíduo responder corretamente a um certo item é sempre a mesma, independentemente da escala utilizada para medir a sua proficiência, ou ainda, a proficiência de um indivíduo é *invariante* à escala de medida. Assim, não faz qualquer sentido querermos analisar itens a partir dos valores de seus parâmetros a e b sem conhecer a escala na qual eles foram determinados.

O modelo proposto pressupõe a unidimensionalidade do teste, isto é, a homogeneidade do conjunto de itens que supostamente devem estar medindo um *único* traço latente. Em outras palavras, deve haver apenas uma proficiência responsável pela realização de todos os itens da prova. Parece claro que qualquer desempenho humano é sempre multideterminado ou multimotivado, dado que mais de um traço latente entra na execução de qualquer tarefa. Contudo, para satisfazer o postulado da unidimensionalidade, é suficiente admitir que haja uma proficiência *dominante* (um fator dominante) responsável pelo conjunto de itens. Este fator é o que se supõe estar sendo medido pelo teste Andrade, Tavares & Valle (2000).

Uma outra suposição do modelo é a chamada independência local ou independência condicional, a qual assume que para uma dada proficiência as respostas aos diferentes itens da prova são independentes. Esta suposição é fundamental para o processo de estimação dos parâmetros do modelo. Na realidade, como a unidimensionalidade implica independência local veja Hambleton & Swaminathan (1991), tem-se somente uma e não duas suposições a serem verificadas. Assim, itens devem ser elaborados de modo a satisfazer a suposição de unidimensionalidade.

Capítulo 6

Análise Fatorial

6.1 Introdução

A Análise Fatorial trata do relacionamento interno de um conjunto de variáveis. Suas idéias básicas se devem, principalmente, a psicólogos como Charles Spearman, Thomson, Thurstone e Burt, que buscavam obter uma melhor compreensão para a “inteligência”Pasquali (2004).

Os testes de inteligência eram - e ainda são - montados com uma grande variedade de itens que variam em graus de memorização, proficiência verbal, proficiência matemática, entre outras. A Análise Fatorial foi desenvolvida para analisar esses testes e averiguar se a “inteligência”era determinada por um único fator geral ou por vários fatores.

Na Análise Fatorial convencional, p variáveis observadas são modeladas como funções lineares de um número menor, m, de outras variáveis contínuas, denominadas variáveis latentes ou fatores. Essas variáveis são empregadas na estimação da correlação entre as variáveis observadas.

Os objetivos principais da Análise Fatorial são:

- determinar o número de fatores que fornecem um ajuste satisfatório à matriz de correlação observada;
- estimar os coeficientes de regressão das variáveis observadas nos fatores.

Com isso, almeja-se uma explicação parcimoniosa do relacionamento entre as variáveis observadas.

Basicamente, o modelo fatorial é motivado pelo seguinte argumento: suponha variáveis que possam ser agrupadas por suas correlações, isto é, suponha que todas as variáveis de um grupo particular de variáveis são fortemente correlacionadas entre si, mas têm, relativamente, baixa correlação com variáveis de outros grupos diferentes. Então, é conceitual que cada grupo de variáveis representa uma única construção básica, ou fator, que é responsável pelas correlações observadas. Por exemplo, correlações de testes de Inglês, Francês e Português sugerem um fator básico: “domínio verbal ou em línguas”. Um segundo grupo de variáveis, representando escores em ciências exatas: Matemática e Física, por exempo, se avaliado, corresponderia a outro fator. É esse tipo de estrutura que a Análise Fatorial procura determinar. Em Nojosa (2003) maiores detalhes poderão ser vistos com relação a Análise Fatorial Convencional.

6.2 Dimensionalidade baseada na Análise Fatorial de Informação Plena

Bock & Aitkin (1981), Bock, Gibbons & Muraki (1988) introduziram um novo método de análise fatorial de itens, baseado na TRI, o qual não requer o cálculo das intercorrelações entre os itens. Deram-lhe o nome de Full-Information Factor Analysis - FIFA, porque trabalha com informações completas em lugar dos métodos de informação limitada ou sumarizada, tais como as correlações. Embora laborioso de um ponto de vista computaciional, particularmente quando o número de itens for grande, este método evita a série de problema encontrados na Análise Fatorial tradicional e no presente, parece ser o melhor método para decidir a unidimensionalidade de uma série de itens, tanto dicotômicos quanto politômicos Bartholomew (1980).

6.2.1 Casos Heywood

É possível que na construção das matrizes de correlação uma ou mais variâncias específicas (unicidades) sejam iguais a zero Heywood, (1931). Unicidades iguais a zero correspondem a variáveis observadas completamente inseridas no espaço dos fatores, ou modeladas perfeitamente pelas variáveis latentes, sem erros de medida. No caso de unicidades não-positivas, Jöreskog & Sörbom (1980) advertem para um mal ajuste do modelo, que pode ocorrer quando um ou mais fatores são pobremente identificados pelo conjunto de variáveis observadas, ou por flutuações desfavoráveis na amostra quando esta é pequena. Os chamados casos Heywoods também ocorrem quando a freqüência observada numa tabela de contigência de um par de itens é zero, e como conseqüência, o valor absoluto da correlação torna-se um.

Na análise fatorial de informação plena, aqui discutida, os casos *Heywood* caracterizam-se quando um ou mais parâmetros "a"aumentam continuamente com o aumento do número de ciclos do algoritmo EM. Dependendo do caso, sugere-se a exclusão dos itens envolvidos ou a aplicação de procedimentos bayesianos.

Nesta seção será discutida a análise fatorial de informação plena introduzida por Bock & Aitkin (1981). O método não requer o cálculo de coeficientes de correlação inter-itens e supera vários dos problemas presentes na análise fatorial da matriz tetracórica [ver Nojosa (2003)].

Será assumido o modelo fatorial para variáveis dicotômicas apresentado através da expressão.

$$\begin{aligned} Y_1 &= \lambda_{11}\theta_1 + \lambda_{12}\theta_2 + \ldots + \lambda_{1m}\theta_m + e_1 \\ Y_2 &= \lambda_{21}\theta_1 + \lambda_{22}\theta_2 + \ldots + \lambda_{2m}\theta_m + e_2 \\ Y_p &= \lambda_{p1}\theta_1 + \lambda_{p2}\theta_2 + \ldots + \lambda_{pm}\theta_m + e_p \end{aligned}$$

As suposições são as seguintes:

1. os resíduos e_i seguem uma distribuição normal de média 0 e variância σ^2;
2. os resíduos são independentes entre si e dos $\theta^{'}s$;
3. $\boldsymbol{\theta}^{'} = (\theta_1, \theta_2, \ldots, \theta_m)$, vetor dos fatores, segue uma distribuição normal multivariada com vetor de média 0 e matriz de covariâncias I_m;
4. $\mathbf{Y} = (Y_1, Y_2, ..., Y_p)$, vetor de variáveis pseudo-observadas, seguem uma distribuição normal multivariada com vetor de médias 0 e matriz de covariância $\boldsymbol{\Sigma}$, onde $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^{'} + \Psi$.

Com o intuito de aproximar a teoria da resposta ao item e a análise fatorial, o leitor pode interpretar a variável $\mathbf{Y}$ como proficiência geral do indivíduo e $\boldsymbol{\theta}$ como o vetor de habilidades específicas. Contudo, neste capítulo, $\boldsymbol{\theta}$ será referido simplesmente como vetor de fatores.

A probabilidade de um indivíduo j responder corretamente ao item i, condicionado ao vetor de fatores θ_j , será dada por

$$\begin{aligned} \Phi_i(\theta_j) &= P(U_{ji} = 1 \mid \theta_j) = \frac{1}{\sqrt{2\pi}\sigma_i} \int_{-\infty}^{\gamma_i} \exp\left[-\frac{1}{2}\left(\frac{t - \sum_{k=1}^{m} \lambda_{ik}\theta_{kj}}{\sigma_i}\right)^2\right] dt \\ &= F\left(\frac{\gamma_i - \sum_{k=1}^{m} \lambda_{ik}\theta_{kj}}{\sigma_i}\right) \end{aligned} \quad (6.1)$$

ou seja, $\Phi_i(\theta_j)$ á a distribuição acumulada normal padrão. Assim, a probabilidade do j-ésimo indivíduo apresentar o padrão de respostas $\mathrm{u}_{.j} = [u_{1j}, u_{2j}, \ldots, u_{pj}]$, condicionado a θ_j, será dada por

$$P(U_{ji} = \mathrm{u}_{ji} \mid \theta_j) = \prod_{i=1}^{p} [\Phi_i(\theta_j)]^{u_{ji}} [1 - \Phi_i(\theta_j)]^{1-u_{ji}},$$

com $\Phi_i(\theta_j) = P(U_{ji} = 1 \mid \theta_j)$, e a probabilidade marginal será dada por

$$P(U_{ji} = \mathrm{u}_{ji}) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(U_{ji} = \mathrm{u}_{ji} \mid \theta_j) f(\theta) d\theta_1 d\theta_2 \ldots d\theta_m = \int_{R^m} L_j(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\theta,$$

onde f é a função densidade de uma distribuição $N_m(\mathbf{0}, I_m)$. Essa integral pode ser aproximada através do uso de pontos de quadratura de Gauss-Hermite Stroud & Sechrest, (1966). A expressão resultante é:

$$P_{ji} = P(U_{ji} = \mathrm{u}_{ji}) \approx \sum_{\kappa_m}^{q} \cdots \sum_{\kappa_2}^{q} \sum_{\kappa_1}^{q} L_j(U_\kappa) A(U_{\kappa_1}) A(U_{\kappa_2}) \ldots A(U_{\kappa_m}),$$

onde $L(U_\kappa) = P(U_{ji} = \mathrm{u}_{ji} \mid \theta_j = U_\kappa)$. Assim, a integral no espaço m-dimensional está sendo substituída pelo somatório em um "grid" de q^m pontos de quadratura $U_\kappa = (U_{\kappa_1}, U_{\kappa_2}, \ldots, U_{\kappa_m})$, $\kappa = 1, 2, \ldots, q$. Como as dimensões do vetor $\boldsymbol{\theta}$ foram assumidas ortogonais, os pesos $A(U_\kappa)$ avaliados em cada ponto são dados pelo produto dos pesos associados a cada coordenada.

Estimação

No processo de estimação da matriz $\mathbf{\Sigma}$, a matriz das cargas fatoriais $\mathbf{\Lambda}$ é responsável pela determinação da dimensão do vetor de fatores $\boldsymbol{\theta}$. Em outras palavras, e de um modo geral, as magnitudes das cargas fatoriais indicam quantos fatores devem ser contemplados pelo modelo e quanto da variância amostral de cada variável é devida a cada fator [ver Nojosa (2003)].

A dimensionalidade na Teoria da Resposta ao Item é vista como o espaço referente ao vetor $\boldsymbol{\theta}$. Portanto, a Análise Fatorial de Informação Plena será utilizada nesse trabalho como método para a verificação da dimensionalidade do espaço do vetor $\boldsymbol{\theta}$.

Um dos critérios utilizado para a determinação da dimensionalidade foi o critério de percentagem de variância. Esse critério é uma abordagem baseada na conquista de um percentual de variancia acumulada especificado da variância total extráda por fatores sucessivos. O objetivo é garantir significância prática para os fatores determinados, garantindo que expliquem pelo menos um montante especificado de variância. Em ciências naturais, o procedimento de obtenção de fatores não deveria ser parado até os fatores extraídos explicarem pelo menos 95% da variância ou até o último fator explicar apenas uma pequena parcela (menos de 5%). Em contraste em ciências sociais, na qual as informações geralmente são menos precisas, não é raro considerar uma solução que explique 60% da variância total (e alguns casos até menos) como satisfatória Hair, Anderson, Tatham & Black (2005). Outro critério também utilizado para determinar a dimensionalidade é o da observação feita no scree-plot Cattel (1966), no qual se dispõe dos autovalores colocados em ordem decrescente. Por este critério procura-se no gráfico um "ponto de salto"que estaria representando um decréscimo de importância em relação à variância total. O número de fatores seria, então, igual ao número de autovalores anteriores ao "ponto de salto"Mingotti (2007). Ou seja, quando o autovalor no primeiro fator for muito maior que nos demais, isso significa que o conjunto de variáveis (itens) podem ser representadas por apenas um fator.

No programa TESTFACT encontram-se implementadas a análise fatorial de informação plena e a análise fatorial através da matriz de correlações tetracóricas.

Capítulo 7

Aplicação

7.1 Introdução

O Exame Nacional do Ensino Médio (ENEM) introduziu um novo conceito de avaliação educacional no Brasil. Diferentemente de avaliações disciplinares como, por exemplo, o Sistema Nacional de Avaliação do Ensino Básico (SAEB) e o VESTIBULAR, onde cada conhecimento é medido em testes individuais, o ENEM é um exame interdisciplinar, onde os vários conhecimentos associados aos conteúdos do ensino fundamental e médio são avaliados de uma só vez por um único teste. O grande diferencial do exame pode ser atribuído aos itens que o compõem. Cada um deles é elaborado de modo a avaliar até 5 competências, mesclando os conhecimentos de diferentes disciplinas. O ENEM caracteriza-se como o exame do perfil de saída da escolaridade básica e tem como um dos principais objetivos fornecer ao participante subsídios para a sua autoavaliação (INEP, 1999; INEP 2000). Entretanto, com o aumento do número de instituições de ensino superior que vêm aderindo à utilização dos resultados do exame como parte de processos seletivos, o número de participantes tem aumentado significativamente. Este capítulo apresenta a análise da dimensionalidade e a modelagem dos dados do Exame Nacional do Ensino Médio dos anos de 1998 a 2008 culminando em uma escala, que é a grande finalidade deste trabalho. A verificação da dimensionalidade foi feita através da Análise Fatorial de Informação Plena introduzida no Capítulo 6. Na modelagem dos dados foi adotado o modelo logístico unidimensional de 3 parâmetros (ML3) e, principalmente o processo de estimação das proficiências apresentado no Capítulo 4. O Exame Nacional do Ensino Médio é aplicado desde 1998 em todo território nacional. É constituído de um teste único contendo 63 itens de múltipla escolha e uma proposta para redação. Os itens objetivos e a redação destinam-se a avaliar as competências desenvolvidas pelos participantes ao longo da escolaridade básica. O exame tem caráter voluntário e dele podem participar, mediante inscrição, os concluintes do ensino médio, no ano de realização do exame, e também os que já o concluíram em anos anteriores, em qualquer de suas modalidades. O ENEM é estruturado por uma matriz de competências que define claramente os pressupostos do exame e delineia suas características operacionais. O modelo da matriz contempla a indicação das competências gerais próprias do aluno, na fase de desenvolvimento cognitivo (*conhecimento*) correspondente ao término da escolaridade básica, associadas aos conteúdos do ensino fundamental e médio. As cinco competências e habilidades tratadas pelo ENEM estão listadas no Capítulo 1. A parte objetiva do exame é composta de 63 itens cujas ordens são alteradas com o intuito de formar 4 diferentes cadernos de prova contendo os mesmos itens. Esses testes são identificados por cores distintas que variam no decorrer dos anos, e a cada um deles corresponde aproximadamente 25,0 % do total de inscritos. A distribuição dos testes segue uma sequência alternada de cores de modo que todos eles estão presentes em todas as localidades de realização do exame. O número de inscritos em todo Brasil em cada ano está na Tabela A.1 do Apêndice A. Nesse trabalho, serão utilizados os dados referentes à parte objetiva do teste de cor amarela (teste amarelo) de 1998 a 2008. Como a distribuição dos testes segundo as cores é bem heterogênea em cada uma das localidades de aplicação do exame, tem-se uma amostra representativa da população de examinados e de tamanho significativo para a estimação dos parâmetros necessários à determinação da dimensionalidade dos dados e também para a estimação dos parâmetros dos itens. Para a estimação das proficiências dos participantes dispõe-se inicialmente de 640 itens, visto que uns poucos foram retirados da análise.

7.2 Resultados

O estudo da dimensionalidade do ENEM foi precedido por uma análise preliminar dos 693 itens que o compõem em todos os anos. As estatísticas descritivas e algumas medidas da teoria clássica foram obtidas. Alguns itens foram retirados da análise por apresentarem correlação bisserial negativa ou inferior a 0,15 ou tiveram seus parâmetros fora das especificações adotadas neste trabalho (ver Capítulo 5).

Tabela 7.1 *Itens retirados por apresentarem correlação bisserial inferior a* 0.15.

Ano	Itens retirados
1998	12, 23, 29, 34, 49,56 e 60
1999	9, 21 e 45
2000	11, 36 e 53
2001	7, 37 e 47
2002	8, 22, 32, 38, 40, 42, 52 e 60
2003	9
2004	8, 14 e 44
2005	3, 15, 29, 34, 46, 63
2006	3, 24, 28, 34, 35, 56, 59, 61, 62
2007	4, 14, 24, 51, 55
2008	09, 21, 38, 46, 63

Isto significa que dos indivíduos que o responderam corretamente, a maior parte foi do grupo de indivíduos de pior desempenho no exame. É provável que a formulação do item tenha induzido os candidatos a não optar pela alternativa correta. Isto poderia comprometer as estimativas dos parâmetros necessários para a determinação da dimensionalidade e também as estimativas dos parâmetro dos itens e das proficiências. Estes resultados levaram à exclusão desses itens do estudo. Da amostra em estudo, foram excluídos também 283 indivíduos que entregaram o exame sem responder nenhum item. Dessa forma, a base de dados do ENEM utilizada neste trabalho será composta por uma matriz de respostas referente a 640 itens e 325830 indivíduos. Na Tabela A.1 e A.5 do Apêndice A essas informações são dadas com maior precisão. Continuando a análise, foram obtidos os percentuais de resposta em branco para cada item de cada ano na qual a (Tabela 7.2) teremos maiores detalhes. O objetivo foi verificar a taxa de não-resposta aos últimos itens do exame. Um percentual grande - considera-se grande qualquer valor acima de 20,0% - poderia indicar que o tempo especificado para a realização do teste foi insuficiente Nojosa (2003). O menor percentual de respostas a cada item de seus respectivos anos estão expostos na tabela anteriormente citada, o que sugere não ter ocorrido problema quanto ao controle de tempo no exame. Os passos seguintes referem-se propriamente à verificação da dimensionalidade.

Tabela 7.2 *Valores em percentuais de itens deixados em branco para os últimos 5 de cada ano (1998 a 2008).*

ANO	item 59	item 60	item 61	item 62	item 63
1998	0,7	0,7	0,7	0,7	0,8
1999	0,5	0,5	0,6	0,5	0,5
2000	0,1	0,2	0,2	0,2	0,2
2001	0,5	0,5	0,5	0,6	0,3
2002	0,6	0,7	0,7	0,5	0,5
2003	0,3	0,3	0,3	0,3	0,4
2004	0,3	0,5	0,4	0,5	0,4
2005	0,3	0,3	0,4	0,4	0,3
2006	0,2	0,3	0,2	0,2	0,2
2007	0,3	0,3	0,3	0,3	0,3
2008	0,6	0,6	0,6	0,5	0,6

Na análise da dimensionalidade dos itens o primeiro passo foi a organização do banco de dados. O programa Statistical Package for the Social Sciences - SPSS, versão 16.0, foi utilizado para esse fim. Com ele procedeu-se à formatação da base de dados de forma conveniente à utilizaçãoo dos programas BilogMG Mislevy & Bock (1990) e TESTFACT, versão 4.0. O segundo passo foi a utilização do programa TESTFACT para determinar a dimensionalidade dos dados. Este programa tem como base teórica os artigos de Bock & Aitkin (1981) e Dempster, Laird & Rubin (1977). O TESTFACT vem sendo utilizado, principalmente, na verificação da dimensionalidade de testes. Nele se encontram implementadas a análise fatorial através da matriz de correlações tetracóricas e a análise fatorial de informação plena. O programa fornece ainda estatísticas descritivas dos itens e algumas medidas utilizadas na teoria clássica.

Como dito anteriormente a Análise Fatorial de Informação Plena será utilizada para determinar a dimensionalidade dos dados do ENEM de todos os anos (caderno amarelo), ou seja, para a determinação do número de fatores necessários ou adequados para a explicação desses testes. Segundo Nojosa (2003), esse método utiliza as cargas fatoriais obtidas a partir da análise fatorial principal sobre a matriz de correlações tetracóricas Divgi (1979) como valores iniciais para o algoritmo EM na análise fatorial de informação plena. Este programa permite soluções para os principais problemas encontrados na construção da matriz de correlações tetracóricas. Essas soluções dizem respeito a possíveis correções antes da aplicação da análise fatorial que são: (i) correção para o acerto casual; (ii) correção para respostas omitidas; (iii) correção para obter uma matriz positiva definida e (iv) correção para evitar casos Heywood. Maiores detalhes podem ser vistos em Wilson et al. (1998). Em relação aos casos Heywood, caracterizados pelo aumento dos parâmetros a's (veja Subseção 3.4.3), o programa permite atribuir distribuições a priori restrita em relação a alguns dos parâmetros dos itens. Como fase final, a matriz de correlações tetracóricas corrigida é submetida à análise fatorial principal com iterações de comunalidade. Esta análise é equivalente à análise fatorial MINRES Harman, (1976).

Os resultados da análise para a verificação da dimensionalidade do Exame Nacional do Ensino Médio dos anos de 1998 a 2008 estão representados na Tabela 7.3 a seguir e também na Figura C.1 que encontra-se no Apêndice C. A tabela mostra os valores da variancia explicada por cada fator Hair, Anderson, Tatham & Black (2005), enquanto a figura supra citada refere-se ao critério do gráfico scree-plot Cattel (1966)(ver capítulo 6). Portanto, como observado em cada ano o número de fatores necessários para explicação de cada um dos conjuntos de variaveis é apenas um, o que significa que o ENEM pode ser considerado um teste unidimensional, logo, o modelo logítico de 3 parâmetros unidimensional pode ser utilizado para a colocação desse teste numa mesma métrica. No Apêndice B pode ser visto a sintaxe do TESTFACT 4.0 utilizada para a geração desses dados.

Tabela 7.3 *Variância total explicada por cada fator*

ANO		1998		1999		2000		2001
Fatores	autovalor	% variância	autovalor	% variância	autovalor	% variância	autovalor	% variância
1	30,21	80,5%	23,47	80,5%	35,87	91,0%	43,73	87,4%
2	4,20	11,2%	2,06	7,1%	1,74	4,4%	3,20	6,4%
3	1,32	3,5%	1,4	4,8%	0,69	1,8%	1,69	3,4%
4	0,95	2,5%	1,25	4,3%	0,61	1,5%	0,88	1,8%
5	0,87	2,3%	0,99	3,4%	0,49	1,2%	0,56	1,1%
total	37,55	100,0%	29,17	100,0%	39,4	100,0%	50,06	100,0%
ANO		2002		2003		2004		2005
	autovalor	% variância	autovalor	% variância	autovalor	% variância	autovalor	% variância
1	39,32	85,1%	27,31	84,9%	31,22	83,3%	36,68	81,7%
2	3,18	6,9%	1,57	4,9%	2,9	7,7%	4,56	10,2%
3	1,3	2,8%	1,33	4,1%	1,34	3,6%	1,54	3,4%
4	1,28	2,8%	1,06	3,3%	1,05	2,8%	1,24	2,8%
5	1,11	2,4%	0,89	2,8%	0,95	2,5%	0,86	1,9%
total	46,19	100,0%	32,16	100,0%	37,46	100,0%	44,88	100,0%
ANO		2006		2007		2008		
Fatores	autovalor	% variância	autovalor	% variância	autovalor	% variância		
1	20,83	75,3%	23,01	82,4%	20,56	71,4%		
2	2,89	10,4%	1,76	6,3%	3,09	10,7%		
3	1,67	6,0%	1,28	4,6%	2,02	7,0%		
4	1,33	4,8%	0,99	3,5%	1,75	6,1%		
5	0,94	3,4%	0,89	3,2%	1,36	4,7%		
total	27,66	100,0%	27,93	100,0%	28,78	100,0%		

Para a estimação dos parâmetros foi utilizado o BilogMG, que é um programa computacional específico para análise de itens dicotômicos ou dicotomizados via TRI. Neste programa estão implementados os modelos unidimensionais logístico e ogiva-normal de 1, 2 e 3 parâmetros para uma ou mais populações de respondentes.

O método de estimação por Máxima Verossimilhança Marginal (ver Seção 4.6) foi utilizado para estimar os parâmetros dos itens na qual a populaçao de referência foi o ENEM de 1998. Para esse grupo considerou-se que as proficiências dos mesmos tem média 0 (zero) e desvio padrão 1 (um). Na Tabela 4.1 do Apêndice A estão expostas as estimativas das mesmas.

Para a estimação das proficiências utilizou-se do metodo de estimação pela Média da Posteriori - EAP (ver Seção 5.4.2).

As Figuras 7.1 a 7.2 mostram as estimativas dos parâmetros dos itens do ENEM para os anos de 1998 a 2008, na qual o modelo adotado foi o Logístico Unidimensional de 3 parâmetros. Os itens do ENEM foram caracterizados pelos parâmetros (a) discriminação, (b) dificuldade e (c) acerto ao acaso. No Apêndice A as estimativas destes parâmetros são apresentadas para os 640 itens considerados nos testes decorridos em todos os anos anteriormente citados. Essas figuras mostram que os itens ofereceram boas estatísticas, pois os mesmos estão com seus parâmetros dentro das especificações adotadas (ver Capítulo 4).

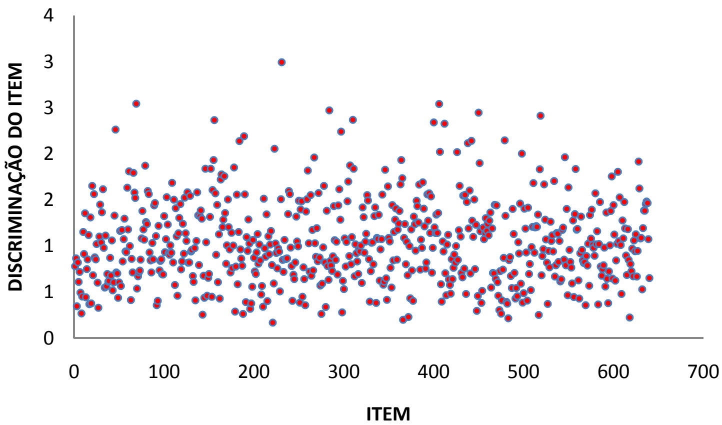

Figura 7.1 *Parâmetro de Discriminação.*

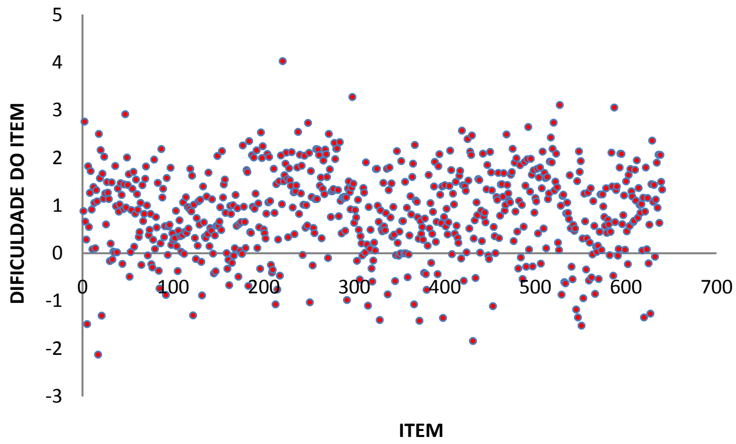

Figura 7.2 *Parâmetro de Dificuldade.*

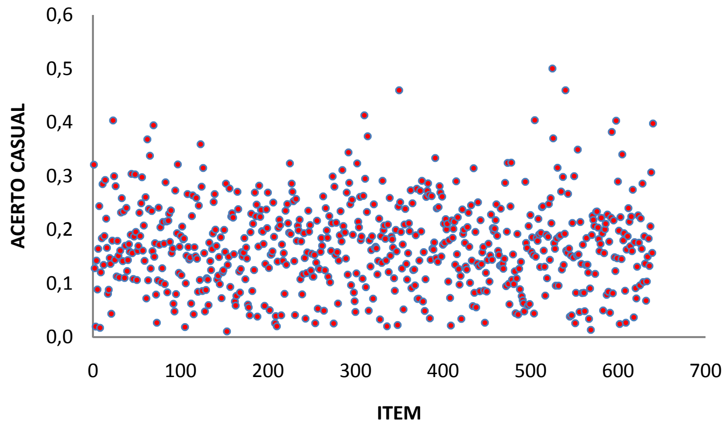

Figura 7.3 *Parâmetro de Acerto Casual.*

As Tabelas 7.2 e 7.3 mostram uma comparação entre as proporções de acertos e estimativas do parâmetro de dificuldade da TRI. Pode-se notar uma relação inversa entre proporção de acerto e o parâmetro de dificuldade do item.

Tabela 7.4 *Estimativas das Proporções de Acerto por Ano*

ANO	MÉDIA POR ANO
1998	42,77
1999	53,18
2000	52,60
2001	41,24
2002	34,61
2003	49,47
2004	46,36
2005	40,61
2006	38,82
2007	52,96
2008	43,23

Tabela 7.5 *Estimativas dos Parâmetros de Dificuldade por Ano*

ANO	MÉDIA
1998	594,18
1999	560,60
2000	556,25
2001	611,33
2002	634,51
2003	564,70
2004	568,97
2005	598,46
2006	626,34
2007	545,43
2008	601,32

A Figura 7.4 e a Figura 7.5 a seguir procuram visualizar melhor essa comparação entre a média da proporção de acerto em cada ano e a média extraída do parâmetro de dificuldade da TRI, também para cada ano. Ambos têm significados parecidos, porém, o parâmetro de dificuldade da TRI para todos os anos estão colocados numa mesma métrica, enquanto, as proporções de acertos estão em métricas diferentes. Observa-se também nessas mesmas figuras, que a medida que as proporções de acertos crescem ou decrescem de um ano para outro, o parâmetro de dificuldade tem comportamento inverso, porém, algumas considerações importantes podem sem feitas, pois, mesmo tendo as proporções de acertos em métricas diferentes, isoladamente pode-se notar que em alguns anos o desempenho dos alunos foram melhor do que em outros. Ainda sobre as figuras supra citadas, considerando o ano de 1998 como ponto de partida para a avaliação do ENEM e, tendo como sequência logicamente, o ano de 1999, pôde-se observar que nesse ano obteve-se uma proporção média de acertos de 53,18%, enquanto, o parâmetro de dificuldade obteve-se média de 560,60. Já em 2000 vê-se que a proporção média de acerto manteve-se praticamente a mesma, 52,60%, enquanto a média do parâmetro de dificuldade decresceu para 556,25. Esse fato mostra que as proporções médias de acertos não estão na mesma métrica, pois, o que se observa é que esses valores se mantem praticamente iguais para os dois anos, mas, isso não acontece para a média do parâmetro de dificuldade, implicando que para esse dois anos a métrica para a proporção média de acertos não é a mesma. Portanto, o mesmo pode-se dizer para os demais anos, pois a metodologia usada para a elabaração dos testes no decorrer dos anos não muda, ou seja, para cada ano um teste é elaborado sendo composto por 63 itens toltalmente diferentes dos anos anteriores, o que impossibilita a comparação no decorrer dos anos. Outra observação que também pode ser feita com relação às médias dos parâmetros de dificuldade, diz respeito à questão da elaboração de testes paralelos, portanto, a Figura 7.5 mostra que os mesmos não os são.

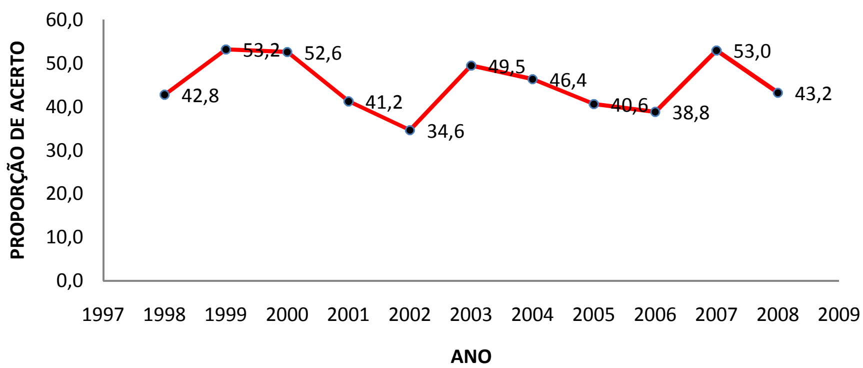

Figura 7.4 *Proporções de Acertos.*

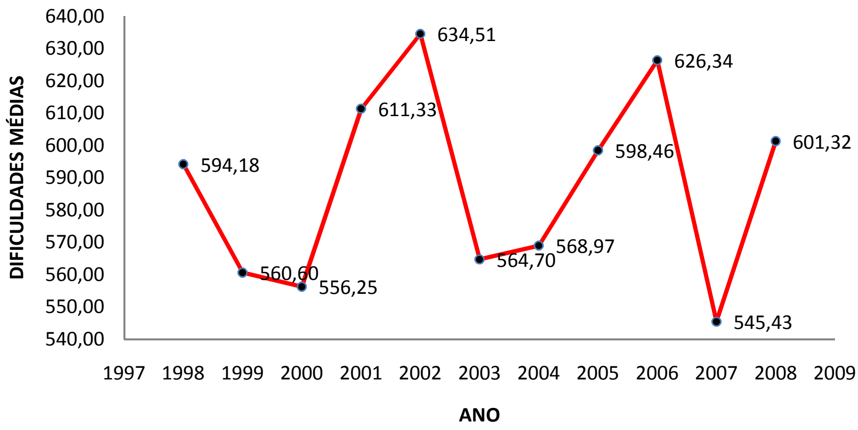

Figura 7.5 *Média das Dificuldades Anuais.*

Vale ressaltar que os resultados, tanto para as tabelas quanto para as figuras, a que nos referimos nesta seção, foram feitas transformações lineares como exposto no Capítulo 5, a fim de melhorar a compreensão do leitor.

Na Tabela 7.6 a seguir estão expostos os resultados obtidos para as proficiências médias anuais e, na Figura 7.6, tem-se uma visualização desses resultados em relação ao que está acontecendo com os grupos estudados.

Observando o comportamento dos dados, pode-se ver uma certa instabilidade para os primeiros cinco anos. Já para os anos seguintes, os mesmos apresentam uma leve tendência de crescimento, o que significa que o desempenho dos grupos em estudo pode estar melhorando.

Tabela 7.6 *Médias das Proficiências por Ano.*

ANO	MÉDIA
1998	500,00
1999	535,58
2000	532,85
2001	509,08
2002	503,48
2003	512,37
2004	512,52
2005	512,20
2006	511,73
2007	515,15
2008	521,97

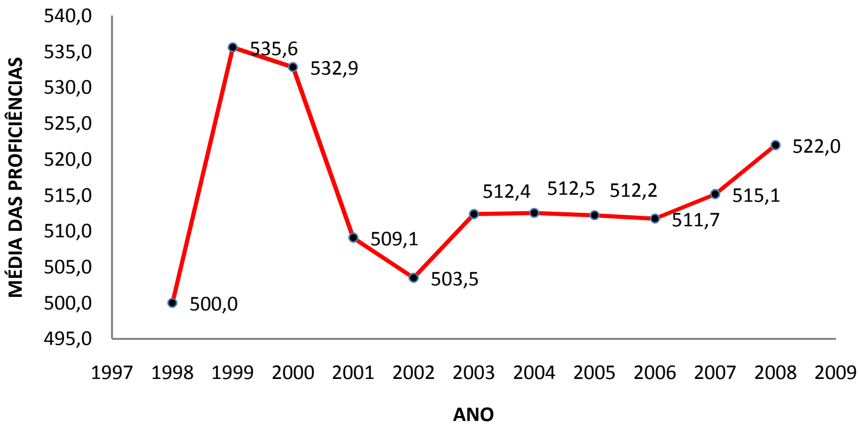

Figura 7.6 *Médias das Proficiências por Ano.*

Capítulo 8

Conclusão e Recomendação

8.1 Considerações Finais

Este trabalho teve como objetivo criar uma escala de proficiências para o Enem e colocá-lo numa mesma métrica com o intuito de tornar comparável os grupos em estudo (ENEM 1998 -2008). Para a verificação da dimensionalidade de cada grupo, foi utilizada como feramenta a Análise Fatorial de Informação Plena, que é especifica para dados dicotômicos, na qual através da mesma pôde-se mostrar que o Enem pode ser ajustado por uma modelo Unidimensional. Para a estimação dos parâmetros dos itens, o método de máxima verossimilhança marginal foi empregado, pois, é o mais apropriado para o caso de avaliações educacionais. Na estimação das proficiências dos testandos, foi utilizado o método da esperança da posteriori (EAP), pois o mesmo tem a vantagem de ser calculada diretamente, não necessitando da aplicação de métodos iterativos e é o mais recomendado por alguns autores, como por exemplo, Mislevy & Stocking (1989).

Uma prova de ligação contendo itens em comum a todos os grupos, aplicada a um grupo de alunos do Estado da Bahia, foi o que possibilitou a realização desse trabalho ao qual os resultados estão expostos no capítulo anterior, mostrando todos os grupos numa mesma métrica e, assim tornando-os comparáveis no decorrer dos anos.

Conjecturas* podem ser feitas com relação ao comportamento do gráfico das proficiências apresentados no capítulo anterior, pois o mesmo apresenta uma subida brusca no rendimento dos alunos de 1998 para 1999. Nenhuma afirmação pode ser feita, mas, pode-se especular† com relação a uma preparação dos mesmos para a realização da prova e também no tocante ao número de alunos em sala de aula para os primeiros anos de realização do ENEM. Para o ano de 1998 pode-se dizer que o hábito de fazer esse tipo de prova tenha

* Uma idéia, fórmula ou frase, a qual não foi provada ser verdadeira, baseada em suposições ou idéias com fundamento não verificado.

† Investigação teórica, esp. de natureza exploratória, sem apoio de evidência sólida (Aurélio)

influenciado no baixo rendimento. Já em 1999 com um número de alunos em sala de aula aproximadamente o mesmo de 1998 e, com a experiência já adquerida pelos professores, é bem provável que esse fato tenha influenciado para o bom resultado daquele ano. Para os 3 anos seguintes, ocorre uma queda no rendimento que pode estar relacionada ao processo de inclusão de alunos em sala de aula e provalvelmente pela falta de certas habilidades dos professores no que diz respeito a receber um número muito grande de alunos ao qual os mesmos não estavam habituados. Concomitantemente ao que foi dito, pode-se dizer também que, quando se aumenta número de alunos em sala e também professores, existe a possibilidade de nesse meio estarem entrando indivíduos com baixo nível de proficiência, o que pode vir a prejudicar o ensino.

Para os anos seguintes uma leve tendência de crescimento pode estar relacionada à questão da manutenção do número de alunos em sala e, pela habilidade adquerida pelos professores em atender uma demanda em número maior. Outro fato que também pode-se levar em consideração, são as políticas educacional implantadas nos últimos anos, que podem estar surtindo um efeito positivo maior que as implantadas nos anos iniciais do ENEM, para os quais nota-se uma certa instabilidade nos resultados, que talvez sejam os reflexos das políticas educacionais adotadas na época. O que foi dito nos últimos dois parágrafos, são apenas conjecturas, porém, fica como sugestão que um estudo seja feito para verifivar se o acréscimo do número de alunos em sala de aula influencia ou não no rendimento dos mesmos.

8.2 Recomendações para trabalhos futuros

- Equalizar o ENEM utilizando modelos multidimensionais com o intuito de comparar os resultados obtidos com o modelo unidimensional;
- Fazer um estudo com o intuito de vericar se o acréscimo de alunos em sala de aula influência no rendimento dos mesmos.

Apêndice A

TABELAS

Tabela A.1 *Quantidades: total de inscritos, caderno amarelo, amostra do caderno amarelo e valores missing na amostra de cada caderno do ENEM de* 1998 *a* 2008.

Ano	Nº Insc. Total	Nº Insc. Parcial *	Amostra	Valores Missing
1998	115575	29978	29878	34
1999	315960	80251	25000	18
2000	352487	89582	25000	11
2001	1200883	311494	22864	27
2002	1318820	341094	30000	27
2003	1322645	343396	30000	17
2004	1035642	269172	25000	29
2005	2200618	569761	25000	18
2006	2783968	704623	25000	25
2007	2737886	694368	40000	16
2008	2262108	740704	40000	59
Bahia	8269	-	8269	-
Total	15654861	4174423	326011	265

Fonte: INEP, 2008

* Números de inscritos referentes aos que realizaram a prova do tipo amarela (tipo 1)

A Tabela A.2 a seguir contém um BIB com as especificações da montagem da prova de ligação que em conjunto com a Tabela A.3 mostram um esquema de como os itens foram escolhidos para a montagem dos cadernos de provas, que seviram como link entre os testes do ENEM ocorridos de 1998 a 2008. Esta prova foi aplicada no dia 1 setembro de 2008, sendo que as mesmas ocorreram no Estado da Bahia tendo como público alvo os alunos de 1º's e 2º's anos.

Tabela A.2 *Especificações para a montagem da prova de ligação.*

BIB DOS CADERNOS DA PROVA DE LIGAÇÃO		
CADERNOS	BLOCO 1	BLOCO 2
1	1	2
2	2	3
3	3	4
4	4	5
5	5	1
6	1	3
7	2	4
8	3	5
9	4	1
10	5	2

Fonte: INEP, 2008

Tabela A.3 *Cadernos de prova de ligação.* *continua*

Bloco 1			Bloco 2		
Ordem no bloco	Nome do item	Número do Item	Ordem no bloco	Nome do item	Número do Item
1	E9804	6	1	E9813	7
2	E9815	2	2	E9833	5
3	E9924	15	3	E9859	2
4	E9937	16	4	E9909	17
5	E9940	21	5	E9915	4
6	E9956	13	6	E9810	14
7	E9961	3	7	E9919	3
8	E0029	18	8	E9942	5
9	E0031	12	9	E9931	19
10	E0045	13	10	E9957	9
11	E0102	14	11	E0108	6
12	E0054	3	12	E0120	4
13	E0055	3	13	E0257	3
14	E0163	4	14	E0263	18
15	E0154	19	15	E0327	13
16	E0243	7	16	E0325	1
17	E0261	3	17	E9960	20
18	E0404	14	18	E0315	15
19	E0419	19	19	E0331	8
20	E0510	19	20	E0344	8
21	E0451	19	21	E0414	14
22	E0513	17	22	E0531	16
23	E0515	1	23	E0621	20
24	E0605	6	24	E0662	14
25	E0636	11	25	E0709	13
26	E0644	13	26	E0717	20
27	E0863	10	27	E0328	13
28	E0747	13	28	E0757	7
29	E0739	16	29	E0834	10
30	E0812	5	30	E0848	12
31	E0809	9	31	E0849	12

Tabela A.4 *Cadernos de prova de ligação.* *continuação*

Bloco 3			Bloco 4		
Ordem no bloco	Nome do item	Número do Item	Ordem no bloco	Nome do item	Número do Item
1	E9807	1	1	E9805	6
2	E9842	13	2	E9837	12
3	E9844	21	3	E9926	1
4	E9858	20	4	E9811	7
5	E9916	17	5	E9928	11
6	E9950	13	6	E9945	7
7	E0304	2	7	E0123	17
8	E0516	1	8	E9946	21
9	E9947	2	9	E0021	20
10	E9949	20	10	E0107	17
11	E9953	20	11	E0130	19
12	E0009	4	12	E0118	18
13	E0161	20	13	E0332	9
14	E0343	16	14	E0144	3
15	E0209	6	15	E0238	15
16	E0313	4	16	E0230	6
17	E0260	10	17	E0318	18
18	E0319	11	18	E0336	8
19	E0316	1	19	E0509	6
20	E0407	13	20	E0356	21
21	E0408	1	21	E0611	19
22	E0453	20	22	E0401	3
23	E0461	10	23	E0804	2
24	E0523	3	24	E0422	5
25	E0529	8	25	E0760	8
26	E0846	5	26	E0428	12
27	E0761	17	27	E0615	21
28	E0763	9	28	E0840	4
29	E0726	6	29	E0638	8
30	E0811	15	30	E0821	14
31	E0838	18	31	E0758	17

Tabela A.5 *Cadernos de prova de ligação.*

Bloco 5		
Ordem no bloco	Nome do item	Número do Item
1	E9828	8
2	E9853	16
3	E9860	2
4	E9922	19
5	E0037	17
6	E0002	11
7	E0008	15
8	E0047	16
9	E0301	6
10	E0052	21
11	E0534	14
12	E0059	4
13	E0333	17
14	E0063	13
15	E0104	20
16	E0115	11
17	E0220	18
18	E0252	10
19	E0337	17
20	E0352	18
21	E0424	18
22	E0430	21
23	E0537	11
24	E0855	15
25	E0449	16
26	E0548	16
27	E0553	20
28	E0661	1
29	E0746	19
30	E0805	9
31	E0825	7

Tabela A.6 *Estimativas dos parâmetros dos itens* *continua*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
1	E9801	0,7774	0,8849	0,3209	31	E9835	1,3646	0,2022	0,1609
2	E9802	0,8651	2,7598	0,1281	32	E9836	1,6148	1,4898	0,2317
3	E9803	0,3440	0,6635	0,0195	33	E9837	0,9790	-0,1283	0,2585
4	E9804	0,8199	0,2890	0,1422	34	E9838	0,5463	0,0371	0,1413
5	E9805	0,6091	-1,4818	0,0886	35	E9839	1,1119	0,0133	0,2340
6	E9806	0,7175	1,8276	0,1643	36	E9840	0,5613	1,8304	0,1100
7	E9807	0,4907	0,5509	0,2436	37	E9841	0,6949	0,9853	0,1701
8	E9808	0,2681	1,2656	0,0169	38	E9842	0,6109	0,0212	0,2403
9	E9809	0,4515	1,7213	0,1203	39	E9843	0,6189	1,3589	0,1230
10	E9810	1,1516	0,9155	0,1839	40	E9844	0,8689	0,8755	0,1431
11	E9811	0,9120	0,0842	0,2846	41	E9845	0,6780	1,0252	0,1728
12	E9813	1,3544	1,3901	0,1340	42	E9846	1,0583	1,4721	0,1591
13	E9814	0,4416	1,0550	0,1882	43	E9847	0,5179	1,2245	0,1902
14	E9815	0,7479	0,1042	0,2923	44	E9848	0,6036	-0,2191	0,3039
15	E9816	0,8520	1,3244	0,2205	45	E9850	1,3284	0,9293	0,1563
16	E9817	1,1166	1,0875	0,1664	46	E9851	2,2613	1,4451	0,1081
17	E9818	0,3664	-2,1239	0,0805	47	E9852	0,7118	2,9183	0,1869
18	E9819	1,3092	2,5039	0,0880	48	E9853	0,7015	0,9589	0,3028
19	E9820	0,3761	1,5742	0,1473	49	E9854	0,4373	2,0089	0,1957
20	E9821	1,6509	2,1641	0,1361	50	E9855	0,6043	1,4359	0,1061
21	E9822	0,6823	-1,3081	0,0430	51	E9857	1,1606	1,6575	0,1605
22	E9824	1,5605	1,6707	0,1795	52	E9858	0,5656	-0,4896	0,1697
23	E9825	0,8768	1,1422	0,4035	53	E9859	1,1714	0,0267	0,2314
24	E9826	0,8541	2,0270	0,2996	54	E9861	0,8313	0,9061	0,1879
25	E9827	1,0197	1,2583	0,1436	55	E9862	1,0722	1,3508	0,2489
26	E9828	0,6004	0,6075	0,2808	56	E9863	1,2948	1,7081	0,2977
27	E9830	0,3270	1,4903	0,1129	57	E9901	0,9396	0,1375	0,1607
28	E9831	1,1015	1,2724	0,1792	58	E9902	0,9315	0,8311	0,2077
29	E9832	1,4470	1,1365	0,1513	59	E9903	1,6312	1,5477	0,1417
30	E9833	1,0354	-0,1626	0,1108	60	E9904	1,1793	1,2399	0,2604

Tabela A.7 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
61	E9905	1,8083	0,9144	0,0720	91	E9937	1,1056	0,5703	0,1645
62	E9906	0,9979	0,3412	0,3679	92	E9938	0,3555	-0,8710	0,0618
63	E9907	0,4120	0,4442	0,1057	93	E9939	0,4042	1,5811	0,0477
64	E9908	0,8587	1,0573	0,2387	94	E9940	1,2176	-0,0343	0,2727
65	E9910	0,7103	-0,2445	0,3378	95	E9941	0,7343	0,5605	0,0801
66	E9911	1,7926	1,4308	0,1276	96	E9942	0,8357	0,6022	0,1934
67	E9912	1,5792	0,6760	0,2347	97	E9943	0,8537	1,7924	0,3216
68	E9913	1,5187	0,8248	0,1599	98	E9944	0,6928	0,1740	0,1196
69	E9914	2,5437	1,5646	0,3943	99	E9946	0,6891	0,4153	0,1683
70	E9915	0,5358	1,8195	0,1495	100	E9947	1,3215	0,2728	0,1885
71	E9916	0,6685	1,0037	0,1745	101	E9948	0,9310	0,3219	0,1150
72	E9917	0,7120	-0,0421	0,0791	102	E9949	1,0730	0,3161	0,2057
73	E9918	0,7304	0,3823	0,0266	103	E9950	1,5271	0,8887	0,1818
74	E9919	0,8500	0,4917	0,1056	104	E9951	1,2298	0,1488	0,1834
75	E9920	1,0796	0,8237	0,2410	105	E9952	0,9186	-0,3697	0,0187
76	E9922	1,3793	-0,2261	0,1718	106	E9953	0,9032	0,4700	0,0998
77	E9923	1,1781	-0,2993	0,2041	107	E9954	1,0061	0,6895	0,1511
78	E9924	1,0674	0,2848	0,1017	108	E9955	0,9244	0,3586	0,2661
79	E9925	1,8704	1,9639	0,1283	109	E9956	1,6830	0,0303	0,1674
80	E9926	1,2577	0,0874	0,2144	110	E9957	0,8194	0,3810	0,1443
81	E9927	1,2372	0,7591	0,1739	111	E9958	1,5018	1,0693	0,1479
82	E9928	1,5953	0,5422	0,0985	112	E9959	0,5756	-0,0117	0,0624
83	E9929	1,5614	1,4755	0,2880	113	E9960	1,4079	0,4490	0,2250
84	E9930	0,9274	-0,3644	0,2153	114	E9961	1,1020	1,1731	0,2639
85	E9931	0,8417	-0,7378	0,0832	115	E9962	0,4600	0,4281	0,0423
86	E9932	0,7102	0,2076	0,2164	116	E9963	1,1361	0,9693	0,1673
87	E9933	1,0306	2,1841	0,2297	117	E0001	1,3019	0,5188	0,1503
88	E9934	0,8531	1,1717	0,1772	118	E0002	1,4535	0,9281	0,2589
89	E9935	1,4381	1,3526	0,2357	119	E0003	0,7446	1,7726	0,1459
90	E9936	1,4643	0,3413	0,1562	120	E0004	0,8308	0,8796	0,0843

Tabela A.8 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
121	E0005	1,2391	1,6284	0,2448	151	E0037	1,3139	0,6669	0,1584
122	E0006	0,9131	-1,3019	0,1061	152	E0038	1,8363	0,9644	0,2857
123	E0007	1,5217	1,0175	0,3588	153	E0039	0,4477	0,4884	0,0104
124	E0008	1,1376	0,3188	0,2799	154	E0040	1,6071	2,1446	0,1344
125	E0009	1,0525	-0,1204	0,0857	155	E0041	1,9319	1,1686	0,1496
126	E0010	1,5810	0,7396	0,3148	156	E0042	2,3636	1,4429	0,2769
127	E0012	0,8184	0,1596	0,1561	157	E0043	1,5687	1,5470	0,0931
128	E0013	0,8205	1,0876	0,0477	158	E0044	0,9045	0,8476	0,2242
129	E0014	1,5433	0,5986	0,0874	159	E0045	1,4463	0,0038	0,2304
130	E0015	0,9406	1,1232	0,1094	160	E0046	0,6049	-0,3234	0,2224
131	E0016	0,9571	-0,1798	0,1069	161	E0047	1,1245	-0,1338	0,1542
132	E0017	0,6731	-0,8800	0,0593	162	E0048	0,4275	-0,6651	0,0661
133	E0018	0,6087	1,4000	0,1758	163	E0049	1,7197	0,9944	0,0586
134	E0019	0,9743	0,6510	0,1413	164	E0050	1,7754	0,8266	0,0766
135	E0020	0,4086	0,3500	0,1245	165	E0051	1,2029	-0,1979	0,2706
136	E0021	1,0530	0,3233	0,1639	166	E0052	1,2827	1,0073	0,2379
137	E0022	1,5824	1,1895	0,1929	167	E0054	1,7575	-0,0506	0,1216
138	E0023	1,3321	0,4002	0,2512	168	E0055	1,3562	0,1045	0,0828
139	E0024	0,7920	1,6892	0,2660	169	E0056	0,8062	1,2263	0,1268
140	E0025	1,3201	0,1507	0,1999	170	E0057	1,0066	0,5730	0,1581
141	E0026	1,3835	0,5321	0,1695	171	E0058	1,5038	0,9462	0,1501
142	E0027	1,2936	0,4795	0,1069	172	E0059	1,2027	-0,4893	0,1571
143	E0028	0,2533	1,1351	0,1144	173	E0060	0,9619	0,6148	0,1844
144	E0029	0,6741	-0,0073	0,1490	174	E0061	0,7118	0,0542	0,1078
145	E0030	0,4352	-0,4229	0,0985	175	E0062	1,1434	0,6544	0,1664
146	E0031	1,8367	0,3727	0,1855	176	E0063	0,7597	-0,0129	0,1462
147	E0032	1,0407	-0,3793	0,0725	177	E0101	1,0032	2,2618	0,0604
148	E0033	0,4600	0,5774	0,1203	178	E0102	1,8502	0,9759	0,0553
149	E0034	0,6540	2,0384	0,1241	179	E0103	0,2851	0,6579	0,0406
150	E0035	0,6946	0,5791	0,1993	180	E0104	0,7758	0,1132	0,1952

Tabela A.9 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
181	E0105	1,5555	1,7426	0,2294	211	E0136	1,5123	1,7708	0,1405
182	E0106	0,5836	1,7045	0,0858	212	E0138	0,9721	0,8502	0,2032
183	E0108	1,1500	-0,6844	0,2110	213	E0139	0,8814	-1,0683	0,2184
184	E0109	2,1368	2,3546	0,1248	214	E0140	1,3418	1,4479	0,2079
185	E0110	0,9620	0,4436	0,2688	215	E0141	0,4014	-0,7597	0,0405
186	E0111	0,8627	0,4274	0,2247	216	E0142	1,1024	0,2865	0,1723
187	E0112	0,7835	2,0560	0,2464	217	E0143	0,9080	1,5070	0,1419
188	E0113	0,2623	0,9772	0,0381	218	E0144	1,0150	-0,4686	0,1564
189	E0114	2,1895	2,0060	0,1686	219	E0145	1,1615	2,0517	0,0802
190	E0115	1,5601	0,9316	0,2818	220	E0146	0,7932	1,0327	0,2353
191	E0116	0,3890	1,2488	0,2237	221	E0147	0,1670	4,0291	0,1348
192	E0117	1,0942	2,1581	0,2368	222	E0149	0,5924	1,5088	0,2475
193	E0118	0,9452	0,1119	0,1972	223	E0150	2,0530	1,6398	0,2117
194	E0119	1,2858	1,0473	0,1962	224	E0151	1,0254	1,8783	0,1942
195	E0120	1,0245	0,5444	0,1731	225	E0152	0,9216	2,1103	0,3235
196	E0121	0,3178	-0,3232	0,0581	226	E0153	0,7559	1,5589	0,1921
197	E0122	0,3754	2,5354	0,1363	227	E0154	0,9736	0,3390	0,2858
198	E0123	0,5556	1,9972	0,2374	228	E0155	0,9368	1,7848	0,2706
199	E0124	1,0649	2,2446	0,2027	229	E0156	0,5047	1,5287	0,2527
200	E0125	0,8693	0,4977	0,2689	230	E0157	1,5069	1,4174	0,1412
201	E0126	0,7217	0,4257	0,1286	231	E0158	2,9925	2,1210	0,0410
202	E0127	0,6949	0,3104	0,0500	232	E0159	0,8489	0,9606	0,2570
203	E0128	0,9279	0,8306	0,1534	233	E0160	0,7158	1,2767	0,1790
204	E0129	1,1286	2,0811	0,1538	234	E0161	1,0637	0,3997	0,2072
205	E0130	0,8059	1,0701	0,1597	235	E0162	0,8557	1,3012	0,1987
206	E0131	1,0130	2,0219	0,2503	236	E0163	0,8577	1,4209	0,2186
207	E0132	0,8416	1,0933	0,1866	237	E0201	0,8637	1,7975	0,1077
208	E0133	0,4733	-0,2411	0,0254	238	E0202	1,3452	2,5478	0,1781
209	E0134	0,5631	-0,4049	0,0388	239	E0203	1,5936	0,8277	0,0793
210	E0135	0,3396	-0,3442	0,0200	240	E0204	1,5609	2,0667	0,1167

Tabela A.10 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
241	E0205	0,9807	1,2578	0,1934	271	E0239	0,8348	-0,1007	0,1020
242	E0206	0,7823	1,8387	0,1473	272	E0241	1,5716	2,5045	0,1606
243	E0207	0,4026	-0,0359	0,0344	273	E0243	0,8740	1,7589	0,2120
244	E0209	0,6857	1,0201	0,1189	274	E0244	0,5687	0,7637	0,2991
245	E0210	0,6402	2,0465	0,1576	275	E0245	0,2624	1,2979	0,0253
246	E0211	0,8209	0,5344	0,2351	276	E0246	0,5888	1,3040	0,2799
247	E0212	1,3160	1,0100	0,1963	277	E0247	0,9848	1,3455	0,1530
248	E0213	1,0474	0,5825	0,2087	278	E0248	1,6511	2,3124	0,2126
249	E0214	0,9940	2,7341	0,1238	279	E0249	0,8195	1,9756	0,1582
250	E0215	1,4900	2,1073	0,1554	280	E0250	0,3361	2,1914	0,0625
251	E0216	0,3840	-1,0262	0,1124	281	E0251	1,3803	2,1988	0,1425
252	E0217	1,3950	1,7241	0,1125	282	E0253	0,7823	1,3238	0,1893
253	E0218	1,4850	1,1822	0,1530	283	E0254	0,6298	0,7790	0,2344
254	E0219	0,4464	-0,2536	0,0259	284	E0255	2,4705	2,3291	0,1781
255	E0220	0,7660	0,5315	0,1386	285	E0256	0,8363	1,0983	0,3108
256	E0221	0,6382	0,4244	0,2166	286	E0257	0,7865	0,4183	0,2067
257	E0223	0,3556	1,1367	0,0530	287	E0258	0,7268	1,1724	0,1958
258	E0224	1,3723	2,1853	0,1276	288	E0259	0,8864	1,1533	0,1947
259	E0225	1,5055	1,1357	0,1259	289	E0261	1,4313	1,1871	0,1460
260	E0226	1,8177	2,1739	0,1588	290	E0262	1,0613	0,4729	0,2691
261	E0227	1,0295	2,0617	0,1844	291	E0263	0,8479	1,1708	0,1159
262	E0228	0,7198	1,4196	0,1701	292	E0301	1,0435	-0,9768	0,3436
263	E0229	1,5507	1,5905	0,2618	293	E0302	0,6883	1,3574	0,2871
264	E0230	0,6740	0,3159	0,0490	294	E0303	1,4401	1,5078	0,2469
265	E0231	1,1729	1,3505	0,1241	295	E0304	1,6155	1,2813	0,0968
266	E0233	0,7346	1,9421	0,2399	296	E0305	1,2836	0,9212	0,2558
267	E0234	1,9591	1,2410	0,1984	297	E0306	2,2406	1,3758	0,1906
268	E0235	1,0269	2,1726	0,1845	298	E0307	0,2773	3,2791	0,0829
269	E0236	0,8804	2,0794	0,1122	299	E0308	0,6200	1,4609	0,0664
270	E0237	1,1935	1,5996	0,2250	300	E0310	0,5322	0,6394	0,0468

Tabela A.11 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
301	E0311	1,0818	0,9063	0,1186	331	E0341	1,0344	0,4621	0,2907
302	E0312	0,7749	0,7492	0,3233	332	E0342	1,0381	1,4788	0,1989
303	E0313	1,0421	-0,1561	0,2146	333	E0343	0,8005	0,5635	0,2075
304	E0314	1,4832	1,1209	0,2042	334	E0344	1,3224	0,4809	0,1829
305	E0315	1,6986	0,4916	0,2611	335	E0345	1,4166	0,8441	0,1422
306	E0316	0,8883	-0,5519	0,1806	336	E0346	0,3807	1,7718	0,0200
307	E0317	1,8741	0,3579	0,1871	337	E0347	1,0874	-0,8628	0,2095
308	E0318	0,8009	0,2062	0,1083	338	E0348	1,0741	1,3241	0,1207
309	E0319	1,0962	-0,0432	0,2630	339	E0349	1,3369	1,8015	0,2588
310	E0320	2,3668	1,3674	0,4127	340	E0350	0,6228	0,7512	0,1514
311	E0321	1,8360	1,2716	0,2946	341	E0351	0,9622	1,4676	0,1872
312	E0322	0,6694	0,0143	0,0927	342	E0352	1,0734	0,3490	0,1818
313	E0323	1,0158	1,9084	0,2329	343	E0353	1,5550	0,9701	0,1667
314	E0324	0,9188	0,1894	0,3739	344	E0354	0,5775	0,3896	0,1093
315	E0325	0,8323	-1,0977	0,0492	345	E0355	0,6108	-0,5800	0,0859
316	E0326	0,6942	0,9727	0,1800	346	E0356	1,8249	-0,0396	0,1993
317	E0327	0,9619	0,0968	0,1315	347	E0357	0,6464	2,1410	0,1595
318	E0328	0,9403	0,4936	0,1617	348	E0358	0,4141	0,2202	0,0221
319	E0329	1,0034	-0,3205	0,1876	349	E0359	1,6463	-0,0493	0,2430
320	E0330	0,5968	-0,5950	0,0715	350	E0360	1,0862	0,4550	0,4597
321	E0331	1,4179	-0,1498	0,2473	351	E0361	1,0479	-0,0039	0,2507
322	E0332	1,3129	0,2249	0,1661	352	E0362	0,7620	1,9333	0,1955
323	E0333	0,9256	0,0586	0,1682	353	E0363	0,5483	0,6761	0,1277
324	E0334	1,0446	1,7658	0,1651	354	E0401	1,4432	0,6772	0,1725
325	E0335	1,1312	1,7717	0,1166	355	E0402	0,8422	-0,0022	0,0515
326	E0336	1,5492	0,9995	0,1420	356	E0403	1,1511	1,4987	0,2095
327	E0337	1,4910	0,6595	0,2205	357	E0404	1,2835	0,0008	0,2386
328	E0338	0,8558	-1,3980	0,0328	358	E0405	1,4072	0,9576	0,1362
329	E0339	0,3947	0,9020	0,1876	359	E0406	0,8499	-0,4993	0,1073
330	E0340	0,7268	0,5051	0,1379	360	E0407	1,1748	-0,0203	0,1563

Tabela A.12 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
361	E0409	1,2431	0,3773	0,2121	391	E0440	0,7532	0,6579	0,3332
362	E0410	1,3789	1,1957	0,1971	392	E0441	0,8197	-0,4392	0,1467
363	E0411	1,6676	0,8140	0,2734	393	E0442	1,5576	1,8283	0,2394
364	E0412	1,9317	1,6015	0,0911	394	E0443	1,2734	2,0759	0,1400
365	E0413	1,7327	1,8800	0,2491	395	E0445	1,5704	1,2278	0,2417
366	E0415	0,1982	-1,0728	0,1247	396	E0446	1,1397	0,6330	0,2808
367	E0416	1,0871	2,2731	0,1258	397	E0447	1,5376	1,4014	0,2691
368	E0417	1,2365	1,1068	0,0959	398	E0448	0,7093	-1,3530	0,1501
369	E0418	0,7993	1,0785	0,1893	399	E0449	0,9783	0,7610	0,2612
370	E0419	1,0449	0,7573	0,2761	400	E0450	2,3390	1,4265	0,1720
371	E0420	0,6806	0,3045	0,0707	401	E0451	1,1999	0,5632	0,2107
372	E0421	0,2271	-1,4144	0,0712	402	E0452	1,1383	0,9339	0,1759
373	E0422	0,9976	0,6759	0,1421	403	E0453	1,3391	0,3945	0,2194
374	E0423	0,4311	0,4052	0,2724	404	E0454	1,1398	0,6465	0,2004
375	E0424	1,1777	0,5497	0,1967	405	E0455	1,3566	1,4104	0,2111
376	E0425	0,7334	0,2821	0,2911	406	E0456	2,5389	2,1545	0,1030
377	E0426	0,4041	-0,4118	0,0659	407	E0457	2,0194	1,7440	0,1798
378	E0427	1,3174	0,7317	0,1583	408	E0458	0,5835	-0,0180	0,0726
379	E0428	1,2347	-0,4348	0,0485	409	E0459	0,3944	1,2501	0,0212
380	E0429	0,9474	-0,7667	0,1077	410	E0460	1,0466	1,0273	0,2169
381	E0430	1,4883	0,1575	0,2611	411	E0461	0,9394	0,4131	0,1847
382	E0431	1,4309	1,6436	0,2723	412	E0462	2,3265	1,4975	0,2013
383	E0432	1,2521	0,5227	0,2860	413	E0463	0,7020	0,5793	0,2149
384	E0433	1,6638	1,0522	0,2649	414	E0501	0,4631	0,3276	0,1347
385	E0434	0,7415	-0,1992	0,0345	415	E0502	1,0144	0,2187	0,2899
386	E0435	1,0659	0,4097	0,1438	416	E0504	1,1192	1,6294	0,1543
387	E0436	0,9791	0,8719	0,2623	417	E0505	0,6358	1,7269	0,2227
388	E0437	1,6891	1,4034	0,1647	418	E0506	0,5764	-0,1111	0,1268
389	E0438	1,4042	1,9185	0,1626	419	E0507	0,4806	2,5709	0,0973
390	E0439	1,2103	0,2405	0,1762	420	E0508	0,8557	1,4280	0,1136

Tabela A.13 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
421	E0509	0,6401	-0,5775	0,0336	451	E0542	1,8966	1,6947	0,1432
422	E0510	0,7617	1,4055	0,2377	452	E0543	1,1592	1,3209	0,1737
423	E0511	0,9071	1,3097	0,1920	453	E0544	0,3814	-1,1076	0,1032
424	E0512	0,8171	0,0217	0,1995	454	E0545	0,5112	0,5527	0,1520
425	E0513	0,9986	2,3945	0,1752	455	E0547	1,1964	0,3569	0,2281
426	E0514	2,0163	1,5268	0,1315	456	E0548	1,0740	0,0035	0,1438
427	E0516	0,8481	0,5181	0,1257	457	E0549	1,1398	1,1786	0,2026
428	E0517	1,1697	2,1319	0,2047	458	E0550	1,3009	1,1219	0,1969
429	E0518	1,6494	2,0629	0,2314	459	E0551	1,1623	1,6865	0,2017
430	E0519	1,1006	2,4688	0,1679	460	E0552	1,2376	0,3185	0,2174
431	E0520	0,7049	-1,8413	0,1008	461	E0553	1,1150	0,7887	0,2472
432	E0521	0,6591	0,3106	0,2451	462	E0554	1,2715	1,4441	0,1884
433	E0522	1,5449	0,6918	0,1422	463	E0555	1,3626	1,4477	0,1605
434	E0523	0,7539	-0,5166	0,0577	464	E0556	0,5472	1,1840	0,1527
435	E0524	1,5544	1,0695	0,3142	465	E0557	1,1891	1,6811	0,1330
436	E0525	0,4768	0,7971	0,0822	466	E0558	0,7062	0,8972	0,2390
437	E0526	1,4764	1,5591	0,1955	467	E0559	0,7007	1,0170	0,1342
438	E0527	2,1153	1,5380	0,0552	468	E0560	1,4461	2,4911	0,1365
439	E0528	1,1896	0,0174	0,1369	469	E0561	0,6428	1,4358	0,1273
440	E0530	1,5997	0,8974	0,2511	470	E0562	1,4383	1,2571	0,1456
441	E0531	0,7018	0,2556	0,1320	471	E0601	0,6049	1,1036	0,2873
442	E0532	2,1398	2,0877	0,0852	472	E0602	0,4582	1,0672	0,0949
443	E0533	1,2738	0,7726	0,2178	473	E0604	0,3032	1,6465	0,1961
444	E0535	1,0983	0,8553	0,1357	474	E0605	0,7260	1,7035	0,3242
445	E0536	1,5024	0,6507	0,1575	475	E0606	0,2647	2,1910	0,0990
446	E0537	0,7484	1,3397	0,1676	476	E0607	0,4099	0,2574	0,0609
447	E0538	1,2286	1,8506	0,0867	477	E0608	0,7411	1,9030	0,2435
448	E0539	0,2938	-0,0422	0,0266	478	E0609	0,8710	0,8746	0,3247
449	E0540	0,4558	-0,1174	0,1197	479	E0610	2,1450	1,9919	0,1066
450	E0541	2,4462	2,1258	0,0845	480	E0611	1,3264	0,4537	0,1537

Tabela A.14 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
481	E0612	0,3915	-0,3202	0,0985	511	E0646	1,6819	1,5791	0,1305
482	E0613	0,3727	1,0980	0,1225	512	E0647	0,9241	0,1083	0,1940
483	E0614	0,2135	1,8455	0,0443	513	E0648	1,2001	1,5982	0,2418
484	E0615	0,8452	-0,0888	0,1091	514	E0649	1,2149	1,1378	0,2435
485	E0616	0,8774	1,0148	0,1916	515	E0650	1,2074	1,2544	0,1608
486	E0617	0,6740	-0,5380	0,0420	516	E0651	1,3407	1,9229	0,1309
487	E0618	0,5432	0,5583	0,1273	517	E0652	0,2510	2,4278	0,0773
488	E0619	1,4000	1,9102	0,0891	518	E0653	1,8320	1,9084	0,1905
489	E0620	1,1286	1,4189	0,1399	519	E0654	2,4108	2,1934	0,1092
490	E0621	0,7062	-0,2784	0,0763	520	E0655	0,6822	2,7382	0,1864
491	E0622	0,4275	1,9730	0,0686	521	E0657	0,8892	1,3145	0,2476
492	E0623	0,5410	2,6470	0,0628	522	E0658	0,9638	1,5016	0,1439
493	E0625	0,4521	0,0811	0,0476	523	E0660	1,6679	2,0889	0,2586
494	E0626	1,0399	1,2913	0,2891	524	E0663	1,2831	1,3845	0,2121
495	E0627	0,9178	1,9795	0,1754	525	E0701	1,0687	0,2241	0,5000
496	E0629	1,4602	1,6472	0,1656	526	E0702	0,8160	0,0680	0,3701
497	E0630	0,5906	-0,2746	0,0585	527	E0703	0,3702	3,1094	0,1747
498	E0631	1,9998	1,7382	0,2067	528	E0705	0,9441	0,9959	0,1642
499	E0632	0,3846	0,8510	0,0614	529	E0706	0,7670	-0,8671	0,0955
500	E0633	0,4860	1,1588	0,2269	530	E0707	1,2607	1,2004	0,1834
501	E0636	1,4036	1,0917	0,1585	531	E0708	1,0543	1,2320	0,3152
502	E0637	1,1865	1,7733	0,1859	532	E0709	1,1143	-0,6027	0,0955
503	E0638	0,9292	0,4199	0,1526	533	E0710	0,9382	-0,6287	0,1878
504	E0639	0,4069	1,5870	0,0438	534	E0711	1,7040	0,8380	0,1892
505	E0640	1,4148	1,7989	0,4039	535	E0712	0,6659	1,0618	0,2708
506	E0641	0,6936	-0,2138	0,1495	536	E0713	0,7895	0,8456	0,1257
507	E0642	0,8042	1,3656	0,1842	537	E0715	1,1403	0,7175	0,2983
508	E0643	0,5209	2,1325	0,1790	538	E0716	1,1304	1,6367	0,1345
509	E0644	0,9879	0,5212	0,2200	539	E0717	0,8241	0,5252	0,1922
510	E0645	1,2069	1,6868	0,1909	540	E0718	1,0528	-0,2833	0,4596

Tabela A.15 *Estimativas dos parâmetros dos itens* *continuação*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
541	E0719	0,6145	-0,5426	0,1611	571	E0750	0,7875	0,4505	0,2257
542	E0720	0,6361	0,5950	0,1570	572	E0752	1,0221	1,2328	0,2180
543	E0721	1,1724	0,4691	0,2668	573	E0753	0,5680	0,0661	0,2067
544	E0722	1,1965	0,5190	0,1645	574	E0754	1,3333	1,2372	0,1176
545	E0723	0,8493	-1,1780	0,0380	575	E0756	1,0791	0,9830	0,2334
546	E0725	1,9625	1,5125	0,1518	576	E0757	1,3915	0,7598	0,2336
547	E0726	0,7880	-1,3433	0,0403	577	E0758	1,0125	0,5712	0,2040
548	E0727	1,3993	2,1382	0,2123	578	E0759	0,9790	1,3613	0,1821
549	E0728	0,4637	1,7089	0,1482	579	E0760	0,8924	0,4272	0,1602
550	E0729	1,6339	1,9329	0,2993	580	E0761	1,5704	0,7016	0,1732
551	E0730	0,9552	-1,5171	0,0254	581	E0762	1,0182	0,6079	0,2220
552	E0731	0,8262	0,3079	0,0823	582	E0763	1,1577	-0,0399	0,1926
553	E0732	0,7655	-0,9417	0,1516	583	E0801	0,3629	0,4528	0,0493
554	E0733	0,5910	1,2500	0,3491	584	E0802	1,0210	2,1091	0,1563
555	E0734	1,2750	0,6700	0,1537	585	E0803	1,4762	1,4036	0,2003
556	E0735	0,4444	-0,3414	0,0466	586	E0804	0,7768	-0,4713	0,2108
557	E0736	1,3444	0,3082	0,2141	587	E0805	0,6741	3,0560	0,0784
558	E0737	1,8340	1,2732	0,1360	588	E0806	0,7489	0,6568	0,2280
559	E0738	1,1752	-0,5690	0,0819	589	E0807	0,7891	1,3254	0,0842
560	E0739	1,1852	0,2232	0,1718	590	E0808	0,9808	1,4019	0,1776
561	E0740	0,6243	1,3618	0,0823	591	E0809	0,5577	-0,0473	0,0451
562	E0741	0,4283	-0,5062	0,0480	592	E0810	1,3620	0,0916	0,2224
563	E0742	0,9199	-0,0226	0,1343	593	E0811	0,6408	1,7976	0,3821
564	E0743	1,3162	1,0964	0,1842	594	E0812	0,3785	2,0882	0,0814
565	E0744	0,9548	0,5757	0,1348	595	E0813	0,7237	2,0859	0,2172
566	E0745	0,8419	-0,8521	0,1199	596	E0814	1,0052	0,6091	0,1483
567	E0746	0,8144	0,1378	0,0339	597	E0815	1,3792	0,8010	0,1229
568	E0747	0,8216	0,0325	0,0906	598	E0816	1,7528	1,2148	0,4028
569	E0748	0,3565	0,1610	0,0134	599	E0817	0,6441	0,0724	0,1467
570	E0749	0,8902	-0,5405	0,1901	600	E0818	1,4427	1,5098	0,2894

Tabela A.16 *Estimativas dos parâmetros dos itens*

ORDEM	NOME	a	b	c	ORDEM	NOME	a	b	c
601	E0819	1,1224	0,4600	0,1980	631	E0853	1,0869	1,4313	0,1365
602	E0820	0,6223	-0,2292	0,0241	632	E0854	0,5269	-0,0726	0,0677
603	E0822	0,8420	1,6436	0,2234	633	E0855	1,1931	1,1292	0,1035
604	E0823	1,0221	0,5760	0,1453	634	E0856	1,3851	1,8986	0,1493
605	E0824	1,8029	1,7687	0,3400	635	E0857	1,3806	0,9552	0,1833
606	E0825	1,0517	1,3007	0,1798	636	E0858	1,4529	2,0661	0,1323
607	E0826	1,5258	1,1917	0,1932	637	E0859	1,4824	0,6398	0,2061
608	E0827	1,0077	0,6516	0,0861	638	E0860	1,4648	2,0625	0,3065
609	E0828	0,4793	0,6817	0,0266	639	E0861	1,0725	1,4965	0,1565
610	E0829	1,2751	1,7497	0,1630	640	E0862	0,6514	1,3356	0,3974
611	E0830	1,0799	1,1049	0,1877	-	-	-	-	-
612	E0831	1,4756	1,9483	0,2401	-	-	-	-	-
613	E0832	1,1907	0,7431	0,2145	-	-	-	-	-
614	E0833	0,8952	1,3695	0,1690	-	-	-	-	-
615	E0834	0,6843	1,0137	0,2242	-	-	-	-	-
616	E0835	1,1806	0,8498	0,1599	-	-	-	-	-
617	E0836	0,8472	1,0318	0,2745	-	-	-	-	-
618	E0837	0,2217	0,0600	0,0365	-	-	-	-	-
619	E0839	0,8020	1,1663	0,1768	-	-	-	-	-
620	E0840	0,7245	-1,3445	0,0909	-	-	-	-	-
621	E0841	0,9688	0,8764	0,0716	-	-	-	-	-
622	E0843	0,6707	1,7965	0,1311	-	-	-	-	-
623	E0844	1,0577	0,0833	0,2264	-	-	-	-	-
624	E0845	0,9411	1,1607	0,1758	-	-	-	-	-
625	E0847	1,1257	-0,2125	0,0961	-	-	-	-	-
626	E0848	0,6660	0,6001	0,2209	-	-	-	-	-
627	E0849	0,9436	-1,2659	0,1644	-	-	-	-	-
628	E0850	1,9152	1,4498	0,2857	-	-	-	-	-
629	E0851	1,6214	2,3614	0,1025	-	-	-	-	-
630	E0852	1,0822	1,0788	0,1875	-	-	-	-	-

Apêndice B

SINTAXE

B.1 TESTFACT

```
>TITLE
LSAT DATA FULL-INFORMATION FACTOR ANALYSIS
>PROBLEM NITEM=63,RESPONSE=4,SELECT=61
>COMMENT;
BOCK AND LIEBERMAN LSAT DATA SECTION 7
STEP-WISE FULL INFORMATION FACTOR ANALYSIS,
VARIMAX ROTATION AND PROMAX ROTATION ARE PERFORMED.
THE DATA CARDS ARE LAYED OUT AS BELOW.
COLUMNS 1 TO 7 . . . . IDENTITY
COLUMNS 8 TO 16 . . . . JUMPED
COLUMNS 7 TO 63 . . . . ITEM RESPONSES
>NAMES ITEM01,ITEM02,ITEM03,ITEM04,ITEM05,ITEM06,ITEM07,ITEM08,
ITEM09,ITEM10,ITEM11,ITEM12,ITEM13,ITEM14,ITEM15,ITEM16,
ITEM17,ITEM18,ITEM19,ITEM20,ITEM21,ITEM22,ITEM23,ITEM24,
ITEM25,ITEM26,ITEM27,ITEM28,ITEM29,ITEM30,ITEM31,ITEM32,
ITEM33,ITEM34,ITEM35,ITEM36,ITEM37,ITEM38,ITEM39,ITEM40,
ITEM41,ITEM42,ITEM43,ITEM44,ITEM45,ITEM46,ITEM47,ITEM48,
ITEM49,ITEM50,ITEM51,ITEM52,ITEM53,ITEM54,ITEM55,ITEM56,
ITEM57,ITEM58,ITEM59,ITEM60,ITEM61,ITEM62,ITEM63;
>RESPONSE ”,’0’,’1’,’*’;
>KEY 111111111111111111111111111111111111111111111111111111111111111;
>SELECT 1(1)22,24(1)55,57(1)63;
>RELIABILITY KR20;
>TETRACHORIC RECODE,NDEC=3,LIST;
>FACTOR NFAC=5,NROOT=5,NIT=(6,0.01),ROTATE=PROMAX,RESIDUAL,SMOOTH;
>FULL ITER=(50,50,0.005),OMIT=RECODE,CPARMS=
(0.334, 0.12358, 0.07198, 0.16178, 0.15322, 0.15043, 0.24724,
0.06852, 0.12773, 0.1648, 0.23496, 0.23526, 0.13355, 0.22698,
0.32731, 0.21885, 0.1636, 0.15126, 0.08144, 0.15183, 0.1362,
0.11761, 0.18556, 0.42145, 0.30987, 0.15374, 0.2791, 0.22367,
0.19358, 0.18871, 0.15097, 0.10481, 0.34189, 0.17412, 0.22842,
0.34541, 0.13157, 0.25201, 0.11758, 0.16832, 0.20015, 0.12944,
0.13031, 0.1808, 0.15825, 0.1962, 0.34217, 0.30786, 0.15067,
0.10278, 0.1904, 0.26232, 0.1869, 0.10966, 0.1539, 0.18213,
0.21443, 0.09266, 0.20145, 0.25849, 0.28398 );
>SAVE SMOOTH,ROTATED,TRIAL;
>INPUT NIDW=7,SCORES,FILE=’C01.DAT’;
(7A1,9X,63A1)
>STOP;
```

B.2 BILOGMG

```
ENEM DE 98 A 08 MAIS PROVA LINK APLICADA NA BAHIA.
Rodada final.
>COMMENTS

ITENS RETIRADOS POR APRESENTAREM CORRELAÇÃO BISSERIAL MENOR QUE 0.15
itens 98; 12, 23, 29, 34, 49 E 56;
itens 99; 84 E 108;
itens 00; 137, 162 e 179;
itens 01; 226 e 237;
itens 02; 260, 274, 284, 292 e 294;
itens 03; 324;
itens 04; 422;
itens 05; 444, 487 e 504;
itens 06; 507, 528, 532, 538, 539, 560 e 563;
itens 07; 571, 581, 591, 618 e 622;
itens 08; 651 e 672.

itens de prova link.

O ENEM tem os seguintes numeros de itens por ano.
3ª SÉRIE DE 98 : 63
3ª SÉRIE DE 99 : 63
3ª SÉRIE DE 00 : 63
3ª SÉRIE DE 01 : 63
3ª SÉRIE DE 02 : 63
3ª SÉRIE DE 03 : 63
3ª SÉRIE DE 04 : 63
3ª SÉRIE DE 05 : 63
3ª SÉRIE DE 06 : 63
3ª SÉRIE DE 07 : 63
3ª SÉRIE DE 08 : 63
Prova link : 155 RETIRADOS DAS PROVAS ANTERIORES.
total : 693
ENEM tem os seguintes numeros de alunos por ano.

3ª SÉRIE DE 98 : 29978/115575 :29978 :34
3ª SÉRIE DE 99 : 80251/315960 :25000 :18
3ª SÉRIE DE 00 : 89592/ :25000 :11
3ª SÉRIE DE 01 : 311494/1200883 :22864 :27
3ª SÉRIE DE 02 : 341094/1318820 :30000 :27
3ª SÉRIE DE 03 : 343396/1322645 :30000 :17
3ª SÉRIE DE 04 : 269172/ :25000 :29
3ª SÉRIE DE 05 : 569761/2200618 :25000 :18
3ª SÉRIE DE 06 : 704623/2783968 :25000 :25
3ª SÉRIE DE 07 : :40000 :16
3ª SÉRIE DE 08 : :40000 :16
Prova link : 8269 :8269 :08
total : :326111 :281

>GLOBAL DFNAME='BAENEM.DAT',NPARM=3,NSUBJECT=325830,
NWGHT=0,NTEST=1,SAVE;
>SAVE PARM='ENEMGERAL.PAR',
```

```
SCORE='ENEMGERAL.SCO',
EXPECTED='EN1MGERAL.EXP';
>LENGTH NITEMS=693;
>INPUT NTOTAL=693,NFMT=1,TYPE=1,SAMPLING=325830,
NALT=5,NIDCHAR=7,
NFORM=21,NGROUP=12;
¿ITEMS INUMBERS=(1(1)693),
INAMES=(
'E9801', 'E9802', 'E9803', 'E9804', 'E9805',
'E9806', 'E9807', 'E9808', 'E9809', 'E9810',
'E9811', 'E9812', 'E9813', 'E9814', 'E9815',
'E9816', 'E9817', 'E9818', 'E9819', 'E9820',
'E9821', 'E9822', 'E9823', 'E9824', 'E9825',
'E9826', 'E9827', 'E9828', 'E9829', 'E9830',
'E9831', 'E9832', 'E9833', 'E9834', 'E9835',
'E9836', 'E9837', 'E9838', 'E9839', 'E9840',
'E9841', 'E9842', 'E9843', 'E9844', 'E9845',
'E9846', 'E9847', 'E9848', 'E9849', 'E9850',
'E9851', 'E9852', 'E9853', 'E9854', 'E9855',
'E9856', 'E9857', 'E9858', 'E9859', 'E9860',
'E9861', 'E9862', 'E9863',

'E9901', 'E9902', 'E9903', 'E9904', 'E9905',
'E9906', 'E9907', 'E9908', 'E9909', 'E9910',
'E9911', 'E9912', 'E9913', 'E9914', 'E9915',
'E9916', 'E9917', 'E9918', 'E9919', 'E9920',
'E9921', 'E9922', 'E9923', 'E9924', 'E9925',
'E9926', 'E9927', 'E9928', 'E9929', 'E9930',
'E9931', 'E9932', 'E9933', 'E9934', 'E9935',
'E9936', 'E9937', 'E9938', 'E9939', 'E9940',
'E9941', 'E9942', 'E9943', 'E9944', 'E9945',
'E9946', 'E9947', 'E9948', 'E9949', 'E9950',
'E9951', 'E9952', 'E9953', 'E9954', 'E9955',
'E9956', 'E9957', 'E9958', 'E9959', 'E9960',
'E9961', 'E9962', 'E9963',

'E0001', 'E0002', 'E0003', 'E0004', 'E0005',
'E0006', 'E0007', 'E0008', 'E0009', 'E0010',
'E0011', 'E0012', 'E0013', 'E0014', 'E0015',
'E0016', 'E0017', 'E0018', 'E0019', 'E0020',
'E0021', 'E0022', 'E0023', 'E0024', 'E0025',
'E0026', 'E0027', 'E0028', 'E0029', 'E0030',
'E0031', 'E0032', 'E0033', 'E0034', 'E0035',
'E0036', 'E0037', 'E0038', 'E0039', 'E0040',
'E0041', 'E0042', 'E0043', 'E0044', 'E0045',
'E0046', 'E0047', 'E0048', 'E0049', 'E0050',
'E0051', 'E0052', 'E0053', 'E0054', 'E0055',
'E0056', 'E0057', 'E0058', 'E0059', 'E0060',
'E0061', 'E0062', 'E0063',

'E0101', 'E0102', 'E0103', 'E0104', 'E0105',
'E0106', 'E0107', 'E0108', 'E0109', 'E0110',
'E0111', 'E0112', 'E0113', 'E0114', 'E0115',
'E0116', 'E0117', 'E0118', 'E0119', 'E0120',
```

'E0121', 'E0122', 'E0123', 'E0124', 'E0125',
'E0126', 'E0127', 'E0128', 'E0129', 'E0130',
'E0131', 'E0132', 'E0133', 'E0134', 'E0135',
'E0136', 'E0137', 'E0138', 'E0139', 'E0140',
'E0141', 'E0142', 'E0143', 'E0144', 'E0145',
'E0146', 'E0147', 'E0148', 'E0149', 'E0150',
'E0151', 'E0152', 'E0153', 'E0154', 'E0155',
'E0156', 'E0157', 'E0158', 'E0159', 'E0160',
'E0161', 'E0162', 'E0163',

'E0201', 'E0202', 'E0203', 'E0204', 'E0205',
'E0206', 'E0207', 'E0208', 'E0209', 'E0210',
'E0211', 'E0212', 'E0213', 'E0214', 'E0215',
'E0216', 'E0217', 'E0218', 'E0219', 'E0220',
'E0221', 'E0222', 'E0223', 'E0224', 'E0225',
'E0226', 'E0227', 'E0228', 'E0229', 'E0230',
'E0231', 'E0232', 'E0233', 'E0234', 'E0235',
'E0236', 'E0237', 'E0238', 'E0239', 'E0240',
'E0241', 'E0242', 'E0243', 'E0244', 'E0245',
'E0246', 'E0247', 'E0248', 'E0249', 'E0250',
'E0251', 'E0252', 'E0253', 'E0254', 'E0255',
'E0256', 'E0257', 'E0258', 'E0259', 'E0260',
'E0261', 'E0262', 'E0263',

'E0301', 'E0302', 'E0303', 'E0304', 'E0305',
'E0306', 'E0307', 'E0308', 'E0309', 'E0310',
'E0311', 'E0312', 'E0313', 'E0314', 'E0315',
'E0316', 'E0317', 'E0318', 'E0319', 'E0320',
'E0321', 'E0322', 'E0323', 'E0324', 'E0325',
'E0326', 'E0327', 'E0328', 'E0329', 'E0330',
'E0331', 'E0332', 'E0333', 'E0334', 'E0335',
'E0336', 'E0337', 'E0338', 'E0339', 'E0340',
'E0341', 'E0342', 'E0343', 'E0344', 'E0345',
'E0346', 'E0347', 'E0348', 'E0349', 'E0350',
'E0351', 'E0352', 'E0353', 'E0354', 'E0355',
'E0356', 'E0357', 'E0358', 'E0359', 'E0360',
'E0361', 'E0362', 'E0363',

'E0401', 'E0402', 'E0403', 'E0404', 'E0405',
'E0406', 'E0407', 'E0408', 'E0409', 'E0410',
'E0411', 'E0412', 'E0413', 'E0414', 'E0415',
'E0416', 'E0417', 'E0418', 'E0419', 'E0420',
'E0421', 'E0422', 'E0423', 'E0424', 'E0425',
'E0426', 'E0427', 'E0428', 'E0429', 'E0430',
'E0431', 'E0432', 'E0433', 'E0434', 'E0435',
'E0436', 'E0437', 'E0438', 'E0439', 'E0440',
'E0441', 'E0442', 'E0443', 'E0444', 'E0445',
'E0446', 'E0447', 'E0448', 'E0449', 'E0450',
'E0451', 'E0452', 'E0453', 'E0454', 'E0455',
'E0456', 'E0457', 'E0458', 'E0459', 'E0460',
'E0461', 'E0462', 'E0463',

'E0501', 'E0502', 'E0503', 'E0504', 'E0505',
'E0506', 'E0507', 'E0508', 'E0509', 'E0510',

'E0511', 'E0512', 'E0513', 'E0514', 'E0515',
'E0516', 'E0517', 'E0518', 'E0519', 'E0520',
'E0521', 'E0522', 'E0523', 'E0524', 'E0525',
'E0526', 'E0527', 'E0528', 'E0529', 'E0530',
'E0531', 'E0532', 'E0533', 'E0534', 'E0535',
'E0536', 'E0537', 'E0538', 'E0539', 'E0540',
'E0541', 'E0542', 'E0543', 'E0544', 'E0545',
'E0546', 'E0547', 'E0548', 'E0549', 'E0550',
'E0551', 'E0552', 'E0553', 'E0554', 'E0555',
'E0556', 'E0557', 'E0558', 'E0559', 'E0560',
'E0561', 'E0562', 'E0563',

'E0601', 'E0602', 'E0603', 'E0604', 'E0605',
'E0606', 'E0607', 'E0608', 'E0609', 'E0610',
'E0611', 'E0612', 'E0613', 'E0614', 'E0615',
'E0616', 'E0617', 'E0618', 'E0619', 'E0620',
'E0621', 'E0622', 'E0623', 'E0624', 'E0625',
'E0626', 'E0627', 'E0628', 'E0629', 'E0630',
'E0631', 'E0632', 'E0633', 'E0634', 'E0635',
'E0636', 'E0637', 'E0638', 'E0639', 'E0640',
'E0641', 'E0642', 'E0643', 'E0644', 'E0645',
'E0646', 'E0647', 'E0648', 'E0649', 'E0650',
'E0651', 'E0652', 'E0653', 'E0654', 'E0655',
'E0656', 'E0657', 'E0658', 'E0659', 'E0660',
'E0661', 'E0662', 'E0663',

'E0701', 'E0702', 'E0703', 'E0704', 'E0705',
'E0706', 'E0707', 'E0708', 'E0709', 'E0710',
'E0711', 'E0712', 'E0713', 'E0714', 'E0715',
'E0716', 'E0717', 'E0718', 'E0719', 'E0720',
'E0721', 'E0722', 'E0723', 'E0724', 'E0725',
'E0726', 'E0727', 'E0728', 'E0729', 'E0730',
'E0731', 'E0732', 'E0733', 'E0734', 'E0735',
'E0736', 'E0737', 'E0738', 'E0739', 'E0740',
'E0741', 'E0742', 'E0743', 'E0744', 'E0745',
'E0746', 'E0747', 'E0748', 'E0749', 'E0750',
'E0751', 'E0752', 'E0753', 'E0754', 'E0755',
'E0756', 'E0757', 'E0758', 'E0759', 'E0760',
'E0761', 'E0762', 'E0763',

'E0801', 'E0802', 'E0803', 'E0804', 'E0805',
'E0806', 'E0807', 'E0808', 'E0809', 'E0810',
'E0811', 'E0812', 'E0813', 'E0814', 'E0815',
'E0816', 'E0817', 'E0818', 'E0819', 'E0820',
'E0821', 'E0822', 'E0823', 'E0824', 'E0825',
'E0826', 'E0827', 'E0828', 'E0829', 'E0830',
'E0831', 'E0832', 'E0833', 'E0834', 'E0835',
'E0836', 'E0837', 'E0838', 'E0839', 'E0840',
'E0841', 'E0842', 'E0843', 'E0844', 'E0845',
'E0846', 'E0847', 'E0848', 'E0849', 'E0850',
'E0851', 'E0852', 'E0853', 'E0854', 'E0855',
'E0856', 'E0857', 'E0858', 'E0859', 'E0860',
'E0861', 'E0862', 'E0863');

```
>TEST TNAME='PBAHIA',INUMBERS=(1(1)693);
>FORM01 LENGTH= 63, INUMBERS=(
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14,
15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26,
27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38,
39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50,
51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62,
63);
>FORM02 LENGTH= 63, INUMBERS=(
64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75,
76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87,
88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99,
100, 101, 102, 103, 104, 105, 106, 107, 108, 109,
110, 111, 112, 113, 114, 115, 116, 117, 118, 119,
120, 121, 122, 123, 124, 125, 126);
>FORM03 LENGTH= 63, INUMBERS=(
127, 128, 129, 130, 131, 132, 133, 134, 135, 136,
137, 138, 139, 140, 141, 142, 143, 144, 145, 146,
147, 148, 149, 150, 151, 152, 153, 154, 155, 156,
157, 158, 159, 160, 161, 162, 163, 164, 165, 166,
167, 168, 169, 170, 171, 172, 173, 174, 175, 176,
177, 178, 179, 180, 181, 182, 183, 184, 185, 186,
187, 188, 189);
>FORM04 LENGTH= 63, INUMBERS=(
190, 191, 192, 193, 194, 195, 196, 197, 198, 199,
200, 201, 202, 203, 204, 205, 206, 207, 208, 209,
210, 211, 212, 213, 214, 215, 216, 217, 218, 219,
220, 221, 222, 223, 224, 225, 226, 227, 228, 229,
230, 231, 232, 233, 234, 235, 236, 237, 238, 239,
240, 241, 242, 243, 244, 245, 246, 247, 248, 249,
250, 251, 252);
>FORM05 LENGTH= 63, INUMBERS=(
253, 254, 255, 256, 257, 258, 259, 260, 261, 262,
263, 264, 265, 266, 267, 268, 269, 270, 271, 272,
273, 274, 275, 276, 277, 278, 279, 280, 281, 282,
283, 284, 285, 286, 287, 288, 289, 290, 291, 292,
293, 294, 295, 296, 297, 298, 299, 300, 301, 302,
303, 304, 305, 306, 307, 308, 309, 310, 311, 312,
313, 314, 315);
>FORM06 LENGTH= 63, INUMBERS=(
316, 317, 318, 319, 320, 321, 322, 323, 324, 325,
326, 327, 328, 329, 330, 331, 332, 333, 334, 335,
336, 337, 338, 339, 340, 341, 342, 343, 344, 345,
346, 347, 348, 349, 350, 351, 352, 353, 354, 355,
356, 357, 358, 359, 360, 361, 362, 363, 364, 365,
366, 367, 368, 369, 370, 371, 372, 373, 374, 375,
376, 377, 378);
>FORM07 LENGTH= 63, INUMBERS=(
379, 380, 381, 382, 383, 384, 385, 386, 387, 388,
389, 390, 391, 392, 393, 394, 395, 396, 397, 398,
399, 400, 401, 402, 403, 404, 405, 406, 407, 408,
409, 410, 411, 412, 413, 414, 415, 416, 417, 418,
419, 420, 421, 422, 423, 424, 425, 426, 427, 428,
429, 430, 431, 432, 433, 434, 435, 436, 437, 438,
```

439, 440, 441);
>FORM08 LENGTH= 63, INUMBERS=(
442, 443, 444, 445, 446, 447, 448, 449, 450, 451,
452, 453, 454, 455, 456, 457, 458, 459, 460, 461,
462, 463, 464, 465, 466, 467, 468, 469, 470, 471,
472, 473, 474, 475, 476, 477, 478, 479, 480, 481,
482, 483, 484, 485, 486, 487, 488, 489, 490, 491,
492, 493, 494, 495, 496, 497, 498, 499, 500, 501,
502, 503, 504);
>FORM09 LENGTH= 63, INUMBERS=(
505, 506, 507, 508, 509, 510, 511, 512, 513, 514,
515, 516, 517, 518, 519, 520, 521, 522, 523, 524,
525, 526, 527, 528, 529, 530, 531, 532, 533, 534,
535, 536, 537, 538, 539, 540, 541, 542, 543, 544,
545, 546, 547, 548, 549, 550, 551, 552, 553, 554,
555, 556, 557, 558, 559, 560, 561, 562, 563, 564,
565, 566, 567);
>FORM10 LENGTH= 63, INUMBERS=(
568, 569, 570, 571, 572, 573, 574, 575, 576, 577,
578, 579, 580, 581, 582, 583, 584, 585, 586, 587,
588, 589, 590, 591, 592, 593, 594, 595, 596, 597,
598, 599, 600, 601, 602, 603, 604, 605, 606, 607,
608, 609, 610, 611, 612, 613, 614, 615, 616, 617,
618, 619, 620, 621, 622, 623, 624, 625, 626, 627,
628, 629, 630);
>FORM11 LENGTH= 63, INUMBERS=(
631, 632, 633, 634, 635, 636, 637, 638, 639, 640,
641, 642, 643, 644, 645, 646, 647, 648, 649, 650,
651, 652, 653, 654, 655, 656, 657, 658, 659, 660,
661, 662, 663, 664, 665, 666, 667, 668, 669, 670,
671, 672, 673, 674, 675, 676, 677, 678, 679, 680,
681, 682, 683, 684, 685, 686, 687, 688, 689, 690,
691, 692, 693);
>FORM12 LENGTH= 62, INUMBERS=(
4, 15, 87, 100, 103, 119, 124, 155, 157, 171, 191,
180, 181, 252, 243, 295, 313, 382, 397, 451, 429,
454, 456, 509, 540, 548, 693, 614, 606, 642, 639,
13, 33, 59, 72, 78, 10, 82, 105, 94, 120, 197, 209,
309, 315, 342, 340, 123, 330, 346, 359, 392, 472,
525, 566, 576, 584, 343, 624, 664, 678, 679);
>FORM13 LENGTH= 62, INUMBERS=(
13, 33, 59, 72, 78, 10, 82, 105, 94, 120, 197, 209,
309, 315, 342, 340, 123, 330, 346, 359, 392, 472,
525, 566, 576, 584, 343, 624, 664, 678, 679, 7, 42,
44, 58, 79, 113, 319, 457, 110, 112, 116, 135, 250,
358, 261, 328, 312, 334, 331, 385, 386, 431, 439,
464, 470, 676, 628, 630, 593, 641, 668);
>FORM14 LENGTH= 62, INUMBERS=(
7, 42, 44, 58, 79, 113, 319, 457, 110, 112, 116, 135,
250, 358, 261, 328, 312, 334, 331, 385, 386, 431, 439,
464, 470, 676, 628, 630, 593, 641, 668, 5, 37, 89, 11,
91, 108, 212, 109, 147, 196, 219, 207, 347, 233, 290,
282, 333, 351, 450, 371, 515, 379, 634, 400, 627, 406,
519, 670, 542, 651, 625);

```
>FORM15 LENGTH= 62, INUMBERS=(
5, 37, 89, 11, 91, 108, 212, 109, 147, 196, 219, 207,
347, 233, 290, 282, 333, 351, 450, 371, 515, 379, 634,
400, 627, 406, 519, 670, 542, 651, 625, 28, 53, 60, 85,
163, 128, 134, 173, 316, 178, 475, 185, 348, 189, 193,
204, 272, 304, 352, 367, 402, 408, 478, 685, 427, 489,
494, 565, 613, 635, 655);
>FORM16 LENGTH= 62, INUMBERS=(
28, 53, 60, 85, 163, 128, 134, 173, 316, 178, 475, 185,
348, 189, 193, 204, 272, 304, 352, 367, 402, 408, 478,
685, 427, 489, 494, 565, 613, 635, 655, 4, 15, 87, 100,
103, 119, 124, 155, 157, 171, 191, 180, 181, 252, 243,
295, 313, 382, 397, 451, 429, 454, 456, 509, 540, 548,
693, 614, 606, 642, 639);
>FORM17 LENGTH= 62, INUMBERS=(
4, 15, 87, 100, 103, 119, 124, 155, 157, 171, 191, 180,
181, 252, 243, 295, 313, 382, 397, 451, 429, 454, 456,
509, 540, 548, 693, 614, 606, 642, 639, 7, 42, 44, 58,
79, 113, 319, 457, 110, 112, 116, 135, 250, 358, 261,
328, 312, 334, 331, 385, 386, 431, 439, 464, 470, 676,
628, 630, 593, 641, 668);
>FORM18 LENGTH= 62, INUMBERS=(
13, 33, 59, 72, 78, 10, 82, 105, 94, 120, 197, 209, 309,
315, 342, 340, 123, 330, 346, 359, 392, 472, 525, 566, 576,
584, 343, 624, 664, 678, 679, 5, 37, 89, 11, 91, 108, 212,
109, 147, 196, 219, 207, 347, 233, 290, 282, 333, 351, 450,
371, 515, 379, 634, 400, 627, 406, 519, 670, 542, 651, 625);
>FORM19 LENGTH= 62, INUMBERS=(
7, 42, 44, 58, 79, 113, 319, 457, 110, 112, 116, 135, 250,
358, 261, 328, 312, 334, 331, 385, 386, 431, 439, 464,
470, 676, 628, 630, 593, 641, 668, 28, 53, 60, 85, 163,
128, 134, 173, 316, 178, 475, 185, 348, 189, 193, 204,
272, 304, 352, 367, 402, 408, 478, 685, 427, 489, 494,
565, 613, 635, 655);
>FORM20 LENGTH= 62, INUMBERS=(
5, 37, 89, 11, 91, 108, 212, 109, 147, 196, 219, 207, 347,
233, 290, 282, 333, 351, 450, 371, 515, 379, 634, 400, 627,
406, 519, 670, 542, 651, 625, 4, 15, 87, 100, 103, 119,
124, 155, 157, 171, 191, 180, 181, 252, 243, 295, 313,
382, 397, 451, 429, 454, 456, 509, 540, 548, 693, 614,
606, 642, 639);
>FORM21 LENGTH= 62, INUMBERS=(
28, 53, 60, 85, 163, 128, 134, 173, 316, 178, 475, 185, 348,
189, 193, 204, 272, 304, 352, 367, 402, 408, 478, 685, 427,
489, 494, 565, 613, 635, 655, 13, 33, 59, 72, 78, 10, 82,
105, 94, 120, 197, 209, 309, 315, 342, 340, 123, 330, 346,
359, 392, 472, 525, 566, 576, 584, 343, 624, 664, 678, 679);
>GROUP1 GNAME='E98', LENGTH=56, INUMBERS=(
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 14,
15, 16, 17, 18, 19, 20, 21, 22, 24, 25, 26,
27, 28, 30, 31, 32, 33, 35, 36, 37, 38,
39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 50,
51, 52, 53, 54, 55, 57, 58, 59, 61, 62,
63);
```

```
>GROUP2 GNAME='E99', LENGTH=60, INUMBERS=(
64, 65, 66, 67, 68, 69, 70, 71, 73, 74, 75,
76, 77, 78, 79, 80, 81, 82, 83, 85, 86, 87,
88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99,
100, 101, 102, 103, 104, 105, 106, 107, 109,
110, 111, 112, 113, 114, 115, 116, 117, 118, 119,
120, 121, 122, 123, 124, 125, 126);
>GROUP3 GNAME='E00', LENGTH=60, INUMBERS=(
127, 128, 129, 130, 131, 132, 133, 134, 135, 136,
138, 139, 140, 141, 142, 143, 144, 145, 146,
147, 148, 149, 150, 151, 152, 153, 154, 155, 156,
157, 158, 159, 160, 161, 163, 164, 165, 166,
167, 168, 169, 170, 171, 172, 173, 174, 175, 176,
177, 178, 180, 181, 182, 183, 184, 185, 186,
187, 188, 189);
>GROUP4 GNAME='E01', LENGTH=60, INUMBERS=(
190, 191, 192, 193, 194, 195, 197, 198, 199,
200, 201, 202, 203, 204, 205, 206, 207, 208, 209,
210, 211, 212, 213, 214, 215, 216, 217, 218, 219,
220, 221, 222, 223, 224, 225, 227, 228, 229,
230, 231, 232, 233, 234, 235, 236, 238, 239,
240, 241, 242, 243, 244, 245, 246, 247, 248, 249,
250, 251, 252);
>GROUP5 GNAME='E02', LENGTH=55, INUMBERS=(
253, 254, 255, 256, 257, 258, 259, 261, 262,
263, 264, 265, 266, 267, 268, 269, 270, 271, 272,
273, 275, 276, 277, 278, 279, 280, 281, 282,
283, 285, 286, 287, 288, 289, 291,
293, 295, 296, 297, 298, 299, 300, 301, 302,
303, 305, 306, 307, 308, 309, 310, 311,
313, 314, 315);
>GROUP6 GNAME='E03', LENGTH=62, INUMBERS=(
316, 317, 318, 319, 320, 321, 322, 323, 325,
326, 327, 328, 329, 330, 331, 332, 333, 334, 335,
336, 337, 338, 339, 340, 341, 342, 343, 344, 345,
346, 347, 348, 349, 350, 351, 352, 353, 354, 355,
356, 357, 358, 359, 360, 361, 362, 363, 364, 365,
366, 367, 368, 369, 370, 371, 372, 373, 374, 375,
376, 377, 378);
>GROUP7 GNAME='E04', LENGTH=60, INUMBERS=(
379, 380, 381, 382, 383, 384, 385, 387, 388,
389, 390, 391, 393, 394, 395, 396, 397, 398,
399, 400, 401, 402, 403, 404, 405, 406, 407, 408,
409, 410, 411, 412, 413, 414, 415, 416, 417, 418,
419, 420, 421, 423, 424, 425, 426, 427, 428,
429, 430, 431, 432, 433, 434, 435, 436, 437, 438,
439, 440, 441);
>GROUP8 GNAME='E05', LENGTH=57, INUMBERS=(
442, 443, 445, 446, 447, 448, 449, 450, 451,
452, 453, 454, 455, 457, 458, 459, 460, 461,
462, 463, 464, 465, 466, 467, 468, 469, 471,
472, 473, 474, 476, 477, 478, 479, 480, 481,
482, 483, 484, 485, 486, 488, 489, 490, 491,
492, 493, 494, 495, 496, 497, 498, 499, 500, 501,
```

```
502, 503);
>GROUP9 GNAME='E06', LENGTH=54, INUMBERS=(
505, 506, 508, 509, 510, 511, 512, 513, 514,
515, 516, 517, 518, 519, 520, 521, 522, 523, 524,
525, 526, 527, 529, 530, 531, 533, 534,
535, 536, 537, 540, 541, 542, 543, 544,
545, 546, 547, 548, 549, 550, 551, 552, 553, 554,
555, 556, 557, 558, 559, 561, 562, 564,
567);
>GROUP10 GNAME='E07', LENGTH=58, INUMBERS=(
568, 569, 570, 572, 573, 574, 575, 576, 577,
578, 579, 580, 582, 583, 584, 585, 586, 587,
588, 589, 590, 592, 593, 594, 595, 596, 597,
598, 599, 600, 601, 602, 603, 604, 605, 606, 607,
608, 609, 610, 611, 612, 613, 614, 615, 616, 617,
619, 620, 621, 623, 624, 625, 626, 627,
628, 629, 630);
>GROUP11 GNAME='E08', LENGTH=58, INUMBERS=(
631, 632, 633, 634, 635, 636, 637, 638, 639, 640,
641, 642, 643, 644, 645, 646, 647, 648, 649, 650,
652, 653, 654, 655, 656, 657, 658, 659, 660,
661, 662, 663, 664, 665, 666, 667, 669, 670,
671, 673, 674, 675, 677, 678, 679, 680,
681, 682, 683, 684, 685, 686, 687, 688, 689, 690,
691, 692);
>GROUP12 GNAME='BAHIA08', LENGTH=136, INUMBERS=(
4, 15, 87, 100, 103, 119, 124, 155, 157, 171, 191,
180, 181, 252, 243, 295, 313, 382, 397, 451, 429,
454, 509, 540, 548, 614, 606, 642, 639,
13, 33, 59, 78, 10, 82, 105, 94, 120, 197, 209,
309, 315, 342, 340, 123, 330, 346, 359, 472,
525, 576, 584, 343, 624, 664, 678, 679, 7, 42,
44, 58, 79, 113, 319, 457, 110, 112, 116, 135, 250,
358, 261, 328, 334, 331, 385, 431, 439,
464, 628, 630, 593, 641, 5, 37, 89,
11, 91, 212, 109, 147, 219, 207, 347, 233,
282, 333, 351, 450, 371, 515, 379, 634, 400,
627, 406, 519, 670, 542, 625, 28, 53, 85,
163, 128, 134, 173, 316, 178, 185, 348, 189,
193, 204, 272, 352, 367, 402, 408, 478, 685,
427, 489, 494, 613, 655);
(7A1,5X,I2,I2,63A1)
>CALIB NQPT=20, CYCLES=50, NEWTON=0,CRIT=0.01,
IDIST=0, ACCEL=0.5, DIAGNOSIS=0,REFERENCE=1,
NORMAL;
>SCORE METHOD=2,NQPT=20,IDIST=0,Noprint;
```

Apêndice C

FIGURAS

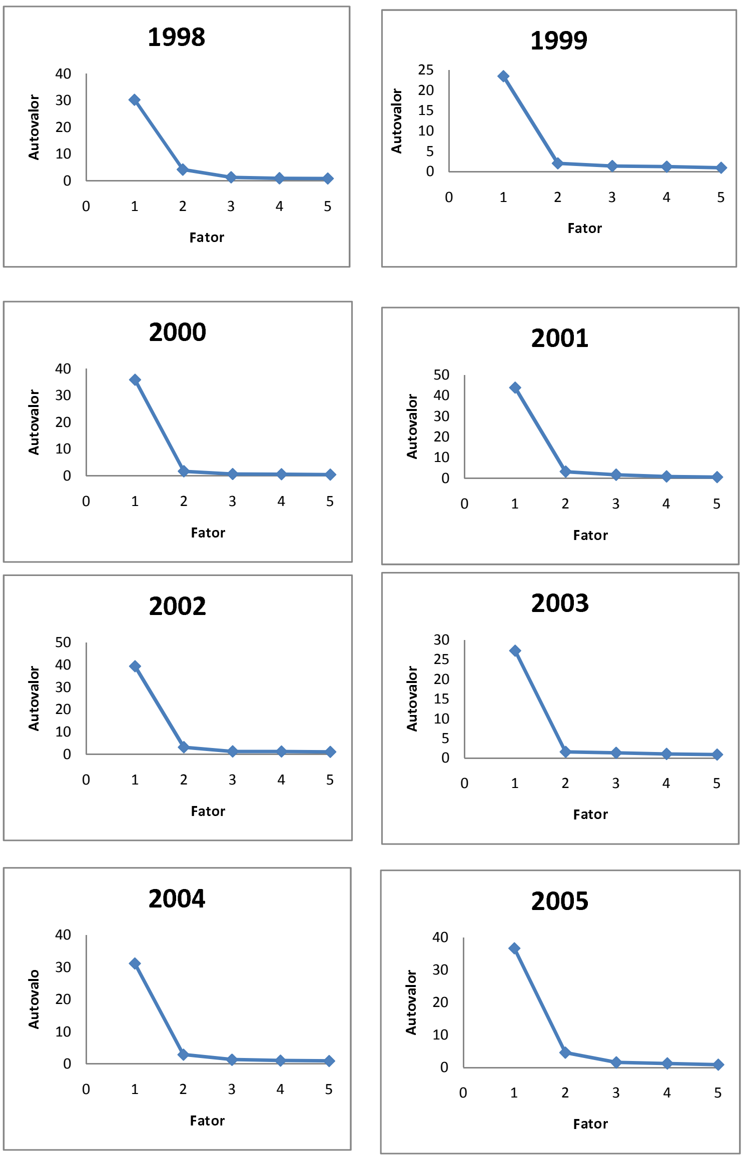

Figura C.1 *Gráfico Scree-Plot 1998 - 2005.*

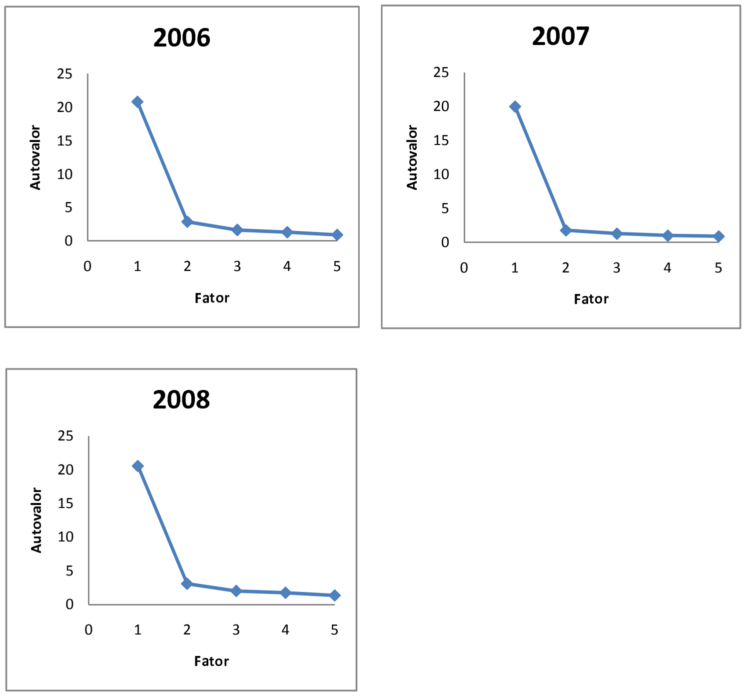

Figura C.2 *Gráfico Scree-Plot 2006 - 2008.*

BIBLIOGRAFIA

ANDRADE, D.F., TAVARES, H.R., VALLE, R.C. (2000). *Teoria da Resposta ao Item: Conceitos e Aplicações.* Associação Brasileira de Estatística: São Paulo.

PASQUALI, L. (2004) *Teoria dos Testes na Psicologia e na Educação* - Petrópolis, RJ: Vozes 2003.

ANDERSEN E. B. (1980). *Discrete Statistical Models with Social Science Applications.* New York: North-Holand Publishing Company.

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.

ANDRADE, D. F. e VALLE, R. C. (1998). *Introdução à teoria da resposta ao item: conceitos e aplicações. Estudos em Avaliação Educacional*, *18*, 13-32.

AZEVEDO, C. L. N. (2003) *Métodos de Estimação na Teoria da Resposta ao Item.* UFC.

BAKER, F. B. (1992). *Item Response Theory - Parameter Estimation Techniques.* New York: Marcel Dekker, Inc.

BIRNBAUM, A. (1957). *Efficient design and use of tests of a mental ability for various decision-making problems*, (Series Report No. 58-16. Project No. 7755-23). USAF School of Aviation Medicine, Texas: Randolph Air Force Base.

BIRNBAUM, A. (1968). Some Latent Trait Models and Their Use in Infering an Examinee's Ability. In F.M. Lord & M.R. Novick. *Statistical Theories of Mental Test Scores.* Reading, MA:Addison-Wesley.

BOCK, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika, 37*, 29-51.

BOCK, R. D. and AITKIN, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of a EM algorithm. *Psychometrika*, *46*, 433-459.

BOCK, R. D. and LIEBERMAN, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika*, *35*, 179-197.

BOCK, R. D. and 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: Spring-Verlag.

CATTEL, R. B. (1966) *The Screen Test For Number de Factors.* Multivariate Behavioral Research, 1, p. 140 - 161.

CONDÉ, F. N. (2001) *Análise Empirica de Itens, Technical Report. Instituto Nacional de Estudos e Pesquisas Educacionais* - DAEB/INEP/MEC.

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.

DIVGI, D. R. (1979). Calculation of the tetrachoric correlation coefficient. *Psychometrika, 44*, 169-172.

GULLIKSEN, H. (1950). *Theory of Mental Tests.* New York : John Wiley and Sons.

HAIR, Joseph F.; TATHAN, Roanald L.; ANDERSON, Rolph E.; BLACK, Williiam. *Análse Maultivariada de Dados*; trad. Adonai Schlup Sant'Anna e Anselmo Chaves Neto. - 5. ed. - Porto Alegre: Bookman, 2005.

HAMBLETON, R.K. and COOK, L.L. (1997). Latent trait models and their use in the analysis of educational test data. *Journal of Educational Measurement, 14*, 75-96.

HAMBLETON, R. K. and SWAMINETHAN, H. (1985). *Item Response Theory: Principles and Applications.* Boston: Kluwer Academic Publishers.

HAMBLETON, R. K., SWAMINETHAN, H. and Rogers, H. J. (1991). *Fundamentals of Item Response Theory.* Newbury Park : Sage Publications.

HEDGES, L. V. and VEVEA, J. L. (1997). A study of equating in NAEP. *Paper presented at The NAEP Validity Studies Panel.* Palo Alto : American Institutes for Research.

ISSAC, E. and KELLER, H. B. (1966). *Analysis of Numerical Methods.* New York: Wiley & Sons.

LINDEN, W. J. VAN DER and HAMBLETON, R. K. (1997). *Handbook of Modern Item Response Theory.* New York : Springer-Verlag.

LORD, F. M. (1952). A theory of test scores (No. 7). *Psychometric Monograph.*

LORD, F. M. (1968). An analysis of the verbal scholastic aptitude test using Birnbaum's three-parameter logistic model. *Educational and Psychological Measurement*, *28*, 989-1020.

LORD, F. M. (1974). Estimation of latent ability and item parameters when there are omitted responses. *Psychometrika*, *39*, 247-264.

LORD, F. M. (1975). *Evaluation with artificial data of a procedure for estimating ability and item characteristic curve parameters*, (Research Bulletin RB-75-33). Princeton, NJ: Educational Testing Service.

LORD, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems.*

Hillsdale: Lawrence Erlbaum Associates, Inc.

LORD, F. M. and NOVICK, M. R. (1968). *Statistical Theories of Mental Test Scores.* Reading, MA: Addison-Wesley.

MINGOTTI, S. A. (2007) *Análise de Dados Através de Métodos de Estatística Multivariada.* Belo Horizonte: UFMG, 2007.

MISLEVY, R. J. (1986a). Bayes modal estimation in item response models. *Psychometrika, 51*, 177-195.

MISLEVY, R. J. (1986b). Recent developments in the factor analysis of categorical variables. *Journal of Educational Statistics, 11*, 3-31.

MISLEVY, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56*, 177-196.

MISLEVY, R. J. (1992). *Linking Educational Assessments : concepts, issues, methods and prospects.* Princeton : Educational Testing Service.

MISLEVY, R. J. and STOCKING, M. L. (1989). A Consumer's Guide to LOGISTIC and BILOG. *Applied Psychological Measurement, 13* 57-75.

MISLEVY, R. J. and BOCK, R. D. (1990). *BILOG 3 : Item Analysis and Test Scoring with Binary Logistic Models.* Chicago : Scientific Software, Inc.

MURAKI, E. (1992). A generalized partial credit model : Application of an EM algorithm. *Applied Psychological Measurement, 16*, 159-176.

MURAKI, E. and BOCK, R. D. (1997). *PARSCALE : IRT Based Test Scoring and Item Analysis for Graded Open-Ended Exercices and Performance Tasks.* Chicago : Scientific Software, Inc.

NOJOSA, R. T. (2001) *Modelos Multimensionais para a Teoria da Resposta ao Item, Master's thesis.* Universidade Federal de Pernambuco.

PEARSON, K. (1900). *On th Correlation of Characters Not Quantitativaly Measurable.* Oyal Society Philosophical Transactions.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications.* New York: Wiley & Sons.

RASCH, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen : Danish Institute for Educational Research.

RECKASE, M. D. (1989). *The Diffculty of Test Itens That Measure More Than One Ability.* Aplied Psychological Measurement.

RICHARDSON, M. W. (1936). The relationship between difficulty and the differential validity of a test. *Psychometrika, 1*, 33-49.

SAMEJIMA, F. A. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph, 17.*

SOARES, J. F. , MARTINS, M. I. e Assunção, C. N. B. (1998). Heterogeneidade acadêmica dos alunos admitidos na UFMG e PUC-MG. *Estudos em Avaliação Educacional, 17*, 61-72. São Paulo : Fundação Carlos Chagas.

SWAMINATHAN, H. and Gifford J. A. (1983). Estimation of Parameters in the Three-Parameter Latent Trait Model. In D. Weiss (Ed.),*New Horizonts in Testing.* New York: Academic Press.

ZIMOWSKI, M. F. , MURAKI, E. , MISLEVY, R. J. and BOCK, R.D. (1996). *BILOG-MG: Multiple-Group IRT Analysis and Test Maintenance for Binary Items.* Chicago : Scientific Software, Inc.