



Modelo Linear do Efeito-Posição do Item via Estimação Bayesiana na TRI

Shirlaine Moraes e Souza

DISSERTAÇÃO APRESENTADA
AO
PROGRAMA DE PÓS-GRADUAÇÃO
EM
MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE FEDERAL DO PARÁ
PARA
OBTENÇÃO DO GRAU DE MESTRE
EM
ESTATÍSTICA

Orientadora: **Profa. Dra. Maria Regina Madruga Tavares.**
Co-orientador: **Prof. Dr. Héilton Ribeiro Tavares.**

Belém, 26 de Maio de 2017.

Modelo Linear do Efeito-Posição do Item via Estimação Bayesiana na TRI

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Shirlaine Moraes e Souza e aprovada pela comissão julgadora.

Orientadora: **Profa. Dra. Maria Regina Madruga Tavares.**

Co-orientador: **Prof. Dr. Héilton Ribeiro Tavares.**

Belém, 26 de Maio de 2017.

Dados Internacionais de Catalogação - na - Publicação (CIP)
Biblioteca de Pós-Graduação do ICEN/UFPA

Souza, Shirlaine Moraes e

Modelo linear do efeito-posição do item via estimação bayesiana na TRI/Shirlaine Moraes e Souza; orientadora, Maria Regina Madruga Tavares.-2017.

92 f.: il. 29 cm

Inclui bibliografias

Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Ciências Exatas e Naturais, Programa de Pós-Graduação em Matemática e Estatística, Belém, 2017.

1. Estatística – Estudo e ensino-Avaliação. 2. Teoria bayesiana de decisão estatística. 3. Teoria da resposta ao item (TRI) – Avaliação. 4. Teoria estimativa. I. Tavares, Maria Regina Madruga, orientadora. II. Título.

CDD – 22 ed. 519.542

Shirlaine Moraes e Souza

Modelo Linear do Efeito-Posição do Item via Estimação Bayesiana na TRI

Esta Dissertação, foi julgada e aprovada, para a obtenção do grau de Mestre em Estatística, no Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará.

Belém, 26 de Maio de 2017

João Marcelo Brazão Protázio

Prof. Dr. João Marcelo Brazão Protázio
Coordenador do Programa de Pós-Graduação em Matemática e Estatística - UFPA).

Banca Examinadora

Maria Regina Madrugada Tavares

Profa. Dra. Maria Regina Madruga Tavares
PPGME/UFPA
Orientadora

Héilton Ribeiro Tavares

Prof. Dr. Héilton Ribeiro Tavares
PPGME/UFPA
Co-orientador

Valcir João da Cunha Farias

Prof. Dr. Valcir João da Cunha Farias
PPGME/UFPA
Examinador

Vinícius Duarte Lima

Prof. Dr. Vinícius Duarte Lima
Faculdade de Estatística/UFPA
Examinador

A minha mãe Cipriana Neta.

Agradecimentos

A Deus, por iluminar todos os meus caminhos e por realizar os meus sonhos no tempo certo;

Aos meus pais, Cipriana Neta Moraes e Souza e José do Amaral e Souza, que me mostraram e continuam mostrando o melhor caminho para eu seguir e pelo amor incondicional que até hoje recebo dos dois;

A minha orientadora, Profa. Dra. Maria Regina Madruga Tavares, pelo estímulo e parceria na realização deste trabalho e pelo aprendizado durante a graduação e o Mestrado em estatística;

À Universidade Federal do Pará;

Ao PPGME;

À Profa. Dra. Maria Regina Madruga Tavares, ao Prof. Dr. Héilton Ribeiro Tavares e à Profa. Dra. Marina Yassuko Toma, pela atenção, carinho e apoio proporcionado durante todos esses anos;

À Profa. Dra. Terezinha Ferreira de Oliveira, ao Prof. Dr. Valcir João da Cunha Farias e ao Prof. Dr. Anderson David de Souza Campelo, pelo apoio durante o mestrado;

Ao meu namorado Edelson Sena Souza, por sua dedicação, compreensão e preocupação durante todos esses anos. Por me apoiar em todas as minhas decisões e pelo companheirismo #Teamo.

Aos meus colegas da turma do mestrado de 2015: Fádua, Roger, Miguel, Charles e Francival, pelo apoio, carinho e atenção durante o mestrado;

Às minhas irmãs, sobrinhos(as), primas(os), tios(as) e minhas avós, que sempre me deram atenção nos momentos difíceis e que me proporcionam momentos de felicidades;

À minha querida "Fadinha" e toda sua família, em especial sua mãe Leonetti, que me acolheram durante esses dois anos. Obrigada pelas xícaras de café e etc..

Ao IASEP, por ter me dado todo apoio necessário para a realização dessa dissertação;

Agradeço a equipe da Biblioteca Central Sesc Doca pelo incentivo e pelo apoio constante, sendo este um dos locais onde encontrei serenidade e tranquilidade para escrever boa parte do meu trabalho;

Finalmente, a todos que de alguma forma contribuíram para a realização desta Dissertação.

“ Tudo posso naquele que me fortalece.”

Filipenses 4.13

Resumo

Este trabalho aborda um tema muito relevante no ajuste de modelos da teoria da resposta ao item, o efeito da posição dos itens nos cadernos de prova e/ou questionários. É comum nas avaliações em grande escala usar diferentes cadernos de prova, contendo os mesmos itens em posições distintas, a fim de evitar fraude nessas avaliações. Com isso, há uma grande preocupação com possíveis efeitos associados a mudança de posição do item nos cadernos de prova, que possam alterar os parâmetros do item e, conseqüentemente, o escore (habilidade) de interesse mensurado na avaliação. São apresentados alguns estudos que propõem modelos para avaliar esse efeito, com modelagem em uma e em duas etapas. Nesse trabalho é proposto um modelo, baseado em uma extensão do modelo logístico de 3 parâmetros da TRI, com uma relação linear para avaliar o efeito da posição do item nos parâmetros de discriminação e de dificuldade do item. O processo de estimação dos parâmetros desse modelo considera a abordagem clássica, via método da máxima verossimilhança, e a abordagem bayesiana. O ajuste do modelo proposto em dados simulados produziu resultados satisfatórios, com estimativas muito próximas dos verdadeiros valores adotados. Uma aplicação do modelo proposto aos dados da prova de Ciências da Natureza e suas tecnologias do Exame nacional do Ensino Médio do ano de 2014 (ENEM-2014) foi realizada, e os resultados indicaram efeito significativo da posição em alguns itens do ENEM-2014.

Palavras-chave: modelos de resposta ao item, efeito da posição do item, estimação bayesiana.

Abstract

This work deals with a very relevant subject in the adjustment of item response theory models (IRT), the position effect of items in test booklets and / or questionnaires. It is common in large-scale assessments to use different booklets containing the set of same items in different positions, in order to avoid fraud in that assessments. Thus, there is a great concern with possible effects associated with changing in the position of the item in the test booklets, which may alter the item parameters and, consequently, the estimated ability score for a set of examinees measured in the evaluation. We present some studies that propose models to evaluate that effect modeling it in one and two stages. In this work a model is proposed based on an extension of the 3 parameters logistic model of the IRT, with additive components to evaluate the effect of the position of each item in the discrimination and difficulty parameters. The estimation process of the parameters of this model considers the classical approach, through the maximum likelihood method, and the Bayesian approach as well. The fit of the proposed model in simulated data produced satisfactory results, with estimates very close to the true values adopted. An application of the proposed model to the data of the Nature Science Test and Its Technologies of the National High School Examination of 2014 (NC-ENEM/2014) was carried out, and the results indicated a significant effect of the position in some items.

Keywords: item response models; item position effect; bayesian inference.

Sumário

Resumo	viii
Abstract	ix
Lista de Tabelas	xii
Lista de Figuras	xiv
1 Introdução	1
1.1 Aspectos Gerais	1
1.2 Exame Nacional do Ensino Médio	2
1.3 Justificativa e Importância da Dissertação	3
1.4 Objetivos	4
1.4.1 Objetivo Geral	4
1.4.2 Objetivos Específicos	4
1.5 Estrutura da Dissertação	5
2 Efeito da Posição do Item nos Modelos da TRI	6
2.1 Modelo Logístico Unidimensional de 3 Parâmetros	6
2.1.1 Estimaco dos parâmetros dos itens e das habilidades	8
2.2 Efeito da Posico do Item (EPI)	10
2.2.1 Abordagem clssica em duas etapas na modelagem do EPI	12
2.2.2 Modelos para o EPI em uma etapa	13
2.2.2.1 Modelos para o EPI via parametrizaco DIF na TRI	13
2.2.2.2 Modelo Logístico Linear para o EPI, via Modelos Lineares Generalizados Mistos	15
2.2.2.3 Modelagem Multinível para o EPI	17
2.2.2.4 Modelagem Linear e no Linear do EPI	19
3 Modelo Linear para o EPI: Estimaco por Mxima Verossimilhança	22
3.1 Introduco	22
3.2 Estimaco do EPI em duas etapas	22
3.3 Estimaco do EPI em uma etapa	24
4 Modelo Linear para o EPI: Estimaco Bayesiana	28
4.1 Estimaco Bayesiana	28

4.1.1	Distribuições <i>a priori</i>	29
4.1.2	Distribuição <i>a posteriori</i> conjunta para os parâmetros dos itens	30
4.1.3	Equações de estimação dos parâmetros dos itens	30
5	Resultados e Discussões	33
5.1	Aplicação a dados simulados	33
5.1.1	Estimativas via máxima verossimilhança	34
5.1.2	Estimativas bayesianas	39
5.2	Aplicação aos dados do ENEM-2014	43
5.2.1	Estrutura do exame e da base de dados	43
5.2.2	Distribuição dos Escores	48
5.2.3	Proporções de acertos	50
5.2.4	A Modelagem do EPI em 2 Etapas	54
5.2.4.1	Etapa 1: calibração dos itens	54
5.2.4.2	Etapa 2: teste de hipóteses para a significância do EPI	59
5.2.5	A Modelagem do EPI em uma Etapa	62
5.2.5.1	Estimativas de máxima verossimilhança e bayesianas	62
5.2.5.2	Testes de hipóteses para a significância do EPI em uma etapa	64
6	Considerações finais e recomendações para trabalhos futuros	67
6.1	Considerações finais	67
6.2	Recomendações para trabalhos futuros	68
	Bibliografia	69
A	Intervalos de Confiança e Credibilidade	72

Lista de Tabelas

5.1	Valores adotados para os parâmetros dos itens na simulação dos dados. . .	35
5.2	Estimativas médias dos parâmetros dos itens para os dados simulados, via método da máxima verossimilhança.	36
5.3	Desvios-padrão das estimativas de máxima verossimilhança dos parâmetros dos itens para os dados simulados.	37
5.4	Erros na estimação de máxima verossimilhança dos parâmetros dos itens para os dados simulados.	38
5.5	Estimativas médias dos parâmetros dos itens para os dados simulados, via método bayesiano.	40
5.6	Desvios-padrão das estimativas bayesianas dos parâmetros dos itens para os dados simulados.	41
5.7	Erros na estimação bayesiana dos parâmetros dos itens para os dados simulados.	42
5.8	População de candidatos e tamanho amostral por caderno de prova de Ciências da Natureza, ENEM-2014.	43
5.9	Distribuição dos 45 itens da prova de Ciências da Natureza por caderno, ENEM-2014.	44
5.10	Proporção de acerto, seus valores mínimo e máximo e a diferença, por item e por caderno de prova, ENEM-2014.	51
5.11	Estimativas do parâmetro de discriminação a_i por caderno, seus valores máximos e mínimos e a diferença entre eles.	55
5.12	Estimativas do parâmetro de dificuldade b_i por caderno, seus valores máximos e mínimos e a diferença entre eles.	56
5.13	Estimativas do parâmetro de acerto casual c_i por caderno, seus valores máximos e mínimos e a diferença entre eles.	57
5.14	Estimativas dos Parâmetros a_i , b_i e c_i , considerando os dados conjuntamente.	58
5.15	Estimativas dos parâmetros de EPI no processo de estimação em 2 Etapas.	60
5.16	Itens com EPI nos parâmetros de discriminação e de dificuldade por caderno na estimação em duas etapas.	61
5.17	Estimativas de máxima verossimilhança dos parâmetros dos itens na modelagem do EPI em uma etapa.	63
5.18	Estimativas bayesianas dos parâmetros dos itens na modelagem do EPI em uma etapa.	64
5.19	Itens com EPI nos parâmetros a_i , b_i e em ambos, por caderno, segundo os intervalos de confiança assintóticos.	65

5.20	Itens com EPI nos parâmetros a_i, b_i e em ambos, por caderno, segundo os intervalos de credibilidade.	66
A.1	Limites inferior (l_{inf}^a) e superior (l_{sup}^a) dos intervalos de confiança para os parâmetros $\delta_1^a, \delta_2^a, \delta_3^a$ e δ_4^a na modelagem do EPI em 2 etapas.	73
A.2	Limites inferior (l_{inf}^b) e superior (l_{sup}^b) dos intervalos de confiança para os parâmetros $\delta_1^b, \delta_2^b, \delta_3^b$ e δ_4^b na modelagem do EPI em 2 etapas.	74
A.3	Limites inferior (l_{inf}^a) e superior (l_{sup}^a) dos intervalos de confiança para os parâmetros $\delta_1^a, \delta_2^a, \delta_3^a$ e δ_4^a na modelagem do EPI em 1 etapa.	75
A.4	Limites inferior (l_{inf}^b) e superior (l_{sup}^b) dos intervalos de confiança para os parâmetros $\delta_1^b, \delta_2^b, \delta_3^b$ e δ_4^b na modelagem do EPI em 1 etapa.	76
A.5	Limites inferior (l_{inf}^a) e superior (l_{sup}^a) dos intervalos de credibilidade para os parâmetros $\delta_1^a, \delta_2^a, \delta_3^a$ e δ_4^a na modelagem do EPI em 1 etapa.	77
A.6	Limites inferior (l_{inf}^b) e superior (l_{sup}^b) dos intervalos de credibilidade para os parâmetros $\delta_1^b, \delta_2^b, \delta_3^b$ e δ_4^b na modelagem do EPI em 1 etapa.	78

Lista de Figuras

5.1	Distribuição dos itens da prova de Ciências da Natureza nos cadernos azul, amarelo, branco e rosa, segundo a ordem destes itens, ENEM-2014.	45
5.2	Curva Característica Empírica do Item 6 por caderno de prova.	46
5.3	Curva Característica Empírica do Item 8 por caderno de prova.	46
5.4	Curva Característica Empírica do Item 17 por caderno de prova.	47
5.5	Curva Característica Empírica do Item 20 por caderno de prova.	47
5.6	Curva Característica Empírica do Item 25 por caderno de prova.	48
5.7	Histograma do escore dos respondentes do caderno azul.	48
5.8	Histograma do escore dos respondentes do caderno amarelo.	49
5.9	Histograma do escore dos respondentes do caderno branco.	49
5.10	Histograma do escore dos respondentes do caderno rosa.	50
5.11	Proporção de acertos segundo a dificuldade do item no caderno azul.	52
5.12	Proporção de acertos segundo a dificuldade do item no caderno amarelo.	52
5.13	Proporção de acertos segundo a dificuldade do item no caderno branco.	53
5.14	Proporção de acertos segundo a dificuldade do item no caderno rosa.	53

Capítulo 1

Introdução

1.1 Aspectos Gerais

Atualmente, muitos países vêm investindo em suas políticas educacionais, com o objetivo de promover uma educação eficiente e de qualidade para a formação dos seus cidadãos. Essa preocupação requer o conhecimento e avaliação dos sistemas educacionais vigentes, com a utilização de uma metodologia mais sofisticada e precisa na área de avaliação educacional. A Teoria da Resposta ao Item (TRI), que vem sendo progressivamente introduzida na análise estatística de dados, é um instrumento poderoso nos processos quantitativos de avaliação educacional, pelo fato de permitir a construção de escalas de conhecimento calibradas (ANDRADE et al., 2000).

A TRI surgiu, formalmente, a partir dos trabalhos de Lord (1952) e Rasch (1960). Atualmente, a TRI é uma das técnicas mais utilizada nas áreas de avaliação educacional e psicologia, sendo munida de ferramentas estatísticas para a produção de escalas quantitativas em ambas as áreas. A teoria também pode ter outras aplicações, como nas áreas das Ciências Sociais e Humanas, nas Engenharias, Saúde, etc., basta que esteja em foco a estimação de alguma variável que não seja diretamente observável, as chamadas *variáveis latentes*.

Na área de avaliação educacional deseja-se mensurar o conhecimento dos alunos em uma ou mais áreas, a partir da aplicação de uma prova contendo itens relacionados àquela área de estudo. A Teoria Clássica das Medidas (TCM), que baseia-se no escore total (número de acertos na prova), tem como característica principal analisar e interpretar a prova como um todo. A TRI apresenta grandes vantagens sobre a TCM, uma delas é que ela permite a comparação entre populações, desde que submetidas a provas que tenham alguns itens comuns, ou ainda, a comparação entre indivíduos da mesma população que tenham sido submetidos a provas totalmente diferentes. Segundo Andrade, Tavares e Valle

(2000), uma das principais características da TRI é que ela tem como elementos centrais os itens, e não a prova como um todo.

Alguns países já adotaram os modelos da TRI em suas avaliações. Por exemplo, em 1980 a TRI foi introduzida nos exames do *Scholastic Aptitude Test* (SAT), que mede a competência de estudantes secundaristas nos EUA e tem servido como um dos critérios de ingresso no ensino superior. A metodologia também é adotada no *Graduate Record Examinaton* (GRE), para estudantes candidatos à pós-graduação. O *Programa Internacional de Avaliação de Estudantes* (PISA), desenvolvido e coordenado pela Organização para Cooperação e Desenvolvimento Econômico (OCDE), passou a utilizar a metodologia da TRI a partir de 2000.

No Brasil, a TRI foi usada pela primeira vez em 1995 na análise dos dados do Sistema Nacional de Ensino Básico - SAEB. A utilização da TRI permitiu que os desempenhos de alunos de séries diferentes fossem comparados e colocados em uma escala única de conhecimento (NETO et al., 2016). Alguns estados, como São Paulo e Pará, já utilizam a TRI na análise dos dados de suas avaliações internas. Mas, a avaliação mais conhecida no Brasil pelo uso da TRI é o Exame Nacional do Ensino Médio (ENEM), elaborado pelo Instituto Nacional de estudos e Pesquisas Educacionais Anísio Teixeira (INEP). As principais características do ENEM serão apresentadas na próxima seção.

1.2 Exame Nacional do Ensino Médio

No Brasil, o ENEM teve início na década de 90, mais precisamente em 1998, e foi criado com o objetivo de avaliar o desempenho dos estudantes ao fim da educação básica. O ENEM acontece uma vez no ano e mobiliza milhões de estudantes, dentre os quais alguns estão terminando o ensino médio e os outros já concluíram.

Em 2009 aconteceu uma reformulação no ENEM, e ele passou a ser utilizado também como mecanismo de seleção para o ingresso no ensino superior. Com as mudanças introduzidas no ENEM, a nota obtida passou a dar acesso às vagas oferecidas por Instituições Federais de Ensino Superior (IFES), à mobilidade acadêmica e para induzir a reestruturação dos currículos do ensino médio, assim como obter o certificado de conclusão do ensino médio (mais detalhes em www.inep.gov.br).

Para o acesso às universidades, o ENEM pode ser utilizado como fase única de seleção ou combinada com seus processos seletivos próprios. Atualmente o ENEM se transformou

no maior vestibular do Brasil. A nota do ENEM também é utilizada na seleção de bolsistas do Programa Universidade para todos (ProUni), programa do Governo Federal que oferece bolsas de estudos em instituições privadas de ensino superior. É utilizada, também, no Programa Federal de Financiamento Estudantil (FIES).

A reformulação do ENEM foi muito importante para a avaliação educacional no Brasil. Atualmente, é uma avaliação em larga escala e o número de inscritos vem aumentando significativamente. Até hoje, as provas do ENEM são organizadas pelo INEP, instituto vinculado ao Ministério da Educação (MEC). Antes da reformulação do exame, a metodologia utilizada era a Teoria Clássica das Métricas, hoje faz uso da TRI, a principal técnica estatística de modelagem para variáveis latentes.

As provas do ENEM são aplicadas em dois dias. No primeiro dia os inscritos deverão responder a 90 questões (itens) de Ciências da Natureza e Ciências Humanas em até quatro horas e meia.

Já no segundo dia, os cadernos de prova conterão questões de Linguagens e Matemática. Neste dia também serão aplicadas 90 questões, porém a duração do exame é de cinco horas e meia, considerando que os inscritos precisam reservar um determinado tempo (estimado em uma hora) para escrever uma redação.

Para cada dia de prova são utilizados quatro cadernos de provas diferentes (diferenciados pela cor); no entanto, esses cadernos apresentam as mesmas questões (itens) organizadas em posições distintas. Geralmente, estas estratégias de reorganização do item são utilizadas por razões de segurança do teste, ou seja, para limitar a eficácia de memorização dos itens e impedir a cópia de respostas entre os estudantes durante o teste.

1.3 Justificativa e Importância da Dissertação

Em avaliações de grande escala há uma grande preocupação com a segurança do certame, a fim de evitar violações dos cadernos de prova e a ocorrência de cópia entre os candidatos. Nesse sentido, uma das estratégias adotadas é organizar os itens da prova em diferentes cadernos, com os itens comuns ocorrendo em diferentes posições em cada caderno de prova (como ocorre no ENEM, por exemplo).

Outra alternativa é o uso de testes computadorizados, que facilitam na administração dos itens, na logística da aplicação e na redução do custo, além de contribuir com uma maior segurança para a avaliação. Nesses testes são elaborados diferentes cadernos de

prova, contendo ou não itens comuns, e cada indivíduo recebe um desses cadernos para ser resolvido no computador. Um exemplo de teste computadorizado são os Testes Adaptativos Computadorizados (CAT, em inglês), em que os cadernos de prova são criados durante a administração do teste, e o conjunto de itens e a sequência em que eles aparecem na prova podem diferir significativamente entre os estudantes (LEARY e DORANS, 1985), uma vez que a escolha do próximo item a ser administrado dependerá do desempenho do estudante no item anterior, ou seja, o conjunto de itens a serem respondidos se adaptam ao desempenho individual de cada um.

Em algumas situações, essa mudança na posição do item pode afetar as características do item e das provas. Alguns trabalhos recentes (HOHENSINN et al., 2008; MEYERS et al., 2012; MOSES et al., 2007; POMMERICH e HARRIS, 2003; SCHWEIZER et al., 2009) apontam a possibilidade de ocorrer efeitos relacionados à posição do item, que podem tornar o item mais difícil quando aplicado no final da prova (efeito de fadiga do respondente), ou mais fácil, quando aplicado após uma sequência de itens relacionados (efeito de aprendizagem ou de contexto). Assim, é imprescindível que esses efeitos sejam detectados e mensurados, a fim de evitar viés nos resultados obtidos, uma vez que muitas dessas avaliações (como é o caso do ENEM no Brasil) são usadas para fins de classificação em processos seletivos.

Neste trabalho serão apresentados e discutidos métodos que contribuem na modelagem do efeito da posição dos itens, em avaliações baseadas na TRI.

1.4 Objetivos

1.4.1 Objetivo Geral

Utilizar um modelo linear que já incorpore explicitamente alguns parâmetros para detectar o efeito-posição do item, utilizando uma abordagem bayesiana no processo de estimação dos parâmetros do modelo.

1.4.2 Objetivos Específicos

i) Apresentar as principais abordagens discutidas na literatura para a construção de modelos que avaliem o efeito da posição do item na TRI;

ii) Avaliar o efeito-posição do item em dados reais, utilizando o processo de estimação em duas etapas via estimação por máxima verossimilhança;

iii) Analisar o efeito-posição do item em dados simulados, em uma única etapa, via estimação por máxima verossimilhança e estimação bayesiana;

iv) Modelar o efeito-posição do item em dados reais via estimação por máxima verossimilhança e estimação bayesiana.

1.5 Estrutura da Dissertação

Esta dissertação encontra-se dividida em 6 capítulos, a saber:

O Capítulo 1 refere-se à introdução do trabalho, onde são abordados os aspectos gerais, características do ENEM, justificativa e importância da dissertação, os objetivos geral e específicos, bem como o sumário da dissertação.

No Capítulo 2 é apresentada uma introdução aos modelos da TRI, à estimação dos parâmetros dos itens e das habilidades, e as principais abordagens na avaliação do efeito da posição do item nos modelos da TRI.

O Capítulo 3 mostrará a modelagem linear do efeito-posição do Item, com o processo de estimação em duas etapas e em uma etapa, via máxima verossimilhança;

O Capítulo 4 mostrará a modelagem do efeito-posição do item via estimação bayesiana, com a apresentação das distribuições *a priori* e da densidade da distribuição *a posteriori* conjunta.

O Capítulo 5 abordará as aplicações, resultados e discussões das técnicas desenvolvidas no estudo;

O Capítulo 6 apresentará as considerações finais e recomendações para trabalhos futuros.

Capítulo 2

Efeito da Posição do Item nos Modelos da TRI

2.1 Modelo Logístico Unidimensional de 3 Parâmetros

A ideia básica da TRI consiste no emprego de modelos, geralmente paramétricos, nos quais os parâmetros representam características importantes dos itens, sendo, assim, interpretáveis. Esses modelos têm grande utilização em áreas nas quais haja o interesse em estimar características latentes (quantidades não-observáveis). Com base nas respostas dadas a um conjunto de itens relacionados à característica latente de interesse, ajusta-se o modelo escolhido e obtém-se as estimativas dos parâmetros dos itens e dos escores associados à característica latente dos respondentes.

Neste trabalho, será considerada a utilização da TRI na área de avaliação educacional, em que os itens são as questões de uma prova, e o traço latente de interesse é a habilidade (ou proficiência) do aluno em alguma área do conhecimento.

Os modelos da TRI, ao levar em consideração as características dos itens na estimação das habilidades, permitem que a proficiência seja medida com maior precisão e que a escala produzida seja interpretável. Outra vantagem da TRI é possibilitar o cálculo da proficiência de um aluno mesmo que este não tenha respondido a todos os itens, o que é feito sem a necessidade de tratamento adicional dos dados.

Existem modelos da TRI para diferentes tipos de itens. Esses modelos dependem de três fatores, o primeiro é a natureza dos itens, se são dicotômicos (acerta-se ou não o item) ou politômicos, em que se considera níveis de acerto nas possíveis respostas. O segundo fator importante é o número de populações envolvidas, se é apenas uma ou mais de uma população. O terceiro e último fator é a quantidade de traços latentes que está sendo medida, se é apenas um (modelos unidimensionais) ou mais de um (modelos

multidimensionais). Para mais detalhes sobre os diferentes modelos da TRI ver Andrade, Tavares e Valle, 2000.

Neste trabalho serão considerados os modelos logísticos para itens dicotômicos, esses modelos diferem pelo número de parâmetros que utilizam para descrever o item. Existem três tipos, o mais utilizado e completo dentre os modelos propostos pela TRI é o modelo logístico unidimensional de 3 parâmetros (ML3), que é expresso por

$$P_{ji} = P(Y_{ji} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}} \quad (2.1)$$

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

Y_{ji} é uma variável dicotômica que assume o valor 1, quando o indivíduo j responde corretamente o item i , ou 0, quando o indivíduo j não responde corretamente o item i ;

θ_j é a habilidade ou proficiência (traço latente) do j -ésimo indivíduo;

$P(Y_{ji} = 1|\theta_j)$ é a probabilidade de um indivíduo j com habilidade θ_j responder corretamente o item i , e é chamada de Função de Resposta do Item (FRI);

a_i é o parâmetro que representa a discriminação (ou inclinação) do item i ;

b_i é o parâmetro que representa a dificuldade do item i , e é medido na mesma escala da habilidade;

c_i é o parâmetro que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item i (“chance” de acerto ao item para indivíduos com baixa habilidade), também chamado de parâmetro de acerto casual ou “chute”;

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.

A partir do ML3 apresentado na equação (2.1), dois outros modelos podem ser obtidos. Por exemplo, quando não existe possibilidade de acerto ao acaso, pode-se considerar $c_i = 0$ no ML3 e tem-se o chamado modelo logístico unidimensional de 2 parâmetros (ML2), que foi estabelecido por Lord (1952). Birnbaum (1968) modificou o modelo de Lord, utilizando a função de distribuição logística em vez da função de distribuição nor-

mal. Este modelo envolve apenas os seguintes parâmetros θ , b e a , os quais representam o traço latente, a dificuldade e a discriminação do item, respectivamente.

Se, além de não existir resposta ao acaso, ainda tivermos todos os itens com o mesmo poder de discriminação, tem-se o chamado modelo logístico unidimensional de 1 parâmetro (ML1), que é o modelo mais simples (RASCH, 1960) e envolve os parâmetros θ e b , os quais representam o traço latente e a dificuldade do item, respectivamente.

Uma outra maneira de representar os diversos modelos da TRI é utilizando a função *logito*. Por exemplo, considerando o modelo logístico de 2 parâmetros (modelo ML3 com $c_i = 0$), seja a $P(Y_{ji} = 1|\theta_j) = \pi_{ji}$; define-se a função $\text{logito}(\pi_{ji}) = \ln\left(\frac{\pi_{ji}}{1-\pi_{ji}}\right)$. Neste sentido, se $\text{logito}(\pi_{ji}) = \Delta_{ji}$, então $\pi_{ji} = \text{logito}^{-1}(\Delta_{ji}) = \frac{1}{1+e^{-\Delta_{ji}}}$, ou seja, $P(Y_{ji} = 1|\theta_j) = \text{logito}^{-1}(\Delta_{ji})$, onde $\Delta_{ji} = Da_i(\theta_j - b_i)$. Essa notação será usada em alguns modelos no decorrer deste trabalho.

Na próxima seção serão discutidos e apresentados os aspectos relacionados ao processo de estimação dos parâmetros dos itens e do traço latente nos modelos da TRI.

2.1.1 Estimação dos parâmetros dos itens e das habilidades

O ajuste de um modelo da TRI ocorre a partir da estimação dos parâmetros dos itens e das habilidades dos respondentes. Por (2.1) temos que a probabilidade de um indivíduo responder corretamente a um determinado item depende somente da habilidade do indivíduo e dos parâmetros que caracterizam o item. Mas, em geral ambos são desconhecidos. Apenas as respostas dos indivíduos aos itens do teste são observados.

No processo de estimação, tem-se 3 situações possíveis: *i*) os parâmetros dos itens são conhecidos e é necessário estimar apenas as habilidades, ou *ii*) as habilidades dos respondentes são conhecidas e o interesse é apenas a estimação dos parâmetros dos itens, ou *iii*) a situação mais comum, em que deseja-se estimar os parâmetros dos itens e as habilidades dos indivíduos simultaneamente. Na TRI, o processo de estimação dos parâmetros dos itens é conhecido como *calibração*.

Nas três situações citadas anteriormente, o processo de estimação é, geralmente, feito pelo método da máxima verossimilhança, através da aplicação de algum processo iterativo, e/ou usando procedimentos bayesianos (ANDRADE et al., 2000).

A seguir será apresentado o processo de estimação dos parâmetros dos itens pelo método de máxima verossimilhança quando as habilidades são conhecidas. Como já foi

dito anteriormente o modelo mais completo e mais utilizado é o ML3, logo, apenas esse modelo será considerado na apresentação das equações de estimação dos parâmetros.

Antes de iniciar o processo de estimação, algumas notações e suposições serão consideradas. Suponha que temos n indivíduos e I itens no estudo. Como já foi dito anteriormente θ_j é a habilidade do indivíduo j , Y_{ji} é a variável aleatória que representa a resposta (binária) do indivíduo j ao item i , sendo $j = 1, \dots, n$ e $i = 1, \dots, I$. Sejam $\mathbf{Y}_j = (Y_{j1}, Y_{j2}, \dots, Y_{jI})$ o vetor aleatório de respostas do indivíduo j e $\mathbf{Y}_{..} = (Y_{1.}, Y_{2.}, \dots, Y_{n.})$ o conjunto de respostas de todos os n indivíduos, sendo que normalmente $\mathbf{Y}_{..}$ é representado por uma matriz $n \times I$. De forma similar, as observações são representadas por y_{ji} , y_j e $y_{..}$. O vetor $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ representará as habilidades dos n indivíduos e $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_I)$, com $\zeta_i = (a_i, b_i, c_i)$, $i = 1, 2, \dots, I$, o vetor de parâmetros dos itens.

Ao considerar o modelo unidimensional, duas suposições são necessárias na implementação do processo de estimação: a independência entre as respostas de diferentes indivíduos (**S1**) e a independência local (**S2**), ou seja, dada sua habilidade, o indivíduo responde aos itens de forma independente. Assim, sob (**S1**) e (**S2**), pode-se escrever a função de verossimilhança, $L(\mathbf{y}|\boldsymbol{\zeta}) = P(\mathbf{Y}_{..} = \mathbf{y}_{..}|\boldsymbol{\theta}, \boldsymbol{\zeta})$, como

$$L(\mathbf{y}|\boldsymbol{\zeta}) = \prod_{j=1}^n P(\mathbf{Y}_j = y_j|\boldsymbol{\theta}, \boldsymbol{\zeta}) \tag{2.2}$$

$$= \prod_{j=1}^n \prod_{i=1}^I P(Y_{ji} = y_{ji}|\theta_j, \zeta_i). \tag{2.3}$$

Como a variável Y_{ji} segue uma distribuição Bernoulli, com parâmetro P_{ji} dado por (2.1), tem-se

$$L(\mathbf{y}|\boldsymbol{\zeta}) = \prod_{j=1}^n \prod_{i=1}^I P_{ji}^{y_{ji}} Q_{ji}^{1-y_{ji}},$$

com $Q_{ji} = 1 - P_{ji}$, de onde obtemos a log-verossimilhança, dada por

$$\log L(\mathbf{y}|\boldsymbol{\zeta}) = \sum_{j=1}^n \sum_{i=1}^I [y_{ji} \log P_{ji} + (1 - y_{ji}) \log Q_{ji}],$$

Os Estimadores de Máxima Verossimilhança (EMV) de ζ_i , $i = 1, \dots, I$, são os valores que maximizam a log-verossimilhança, ou equivalentemente, são as soluções do sistema de equações a seguir,

$$\frac{\partial \log L(\mathbf{y}|\boldsymbol{\zeta})}{\partial \zeta_i} = 0, \quad i = 1, \dots, I. \tag{2.4}$$

Substituindo a FRI dada em (2.1) na verossimilhança, o sistema de equações para um item i fixado é obtido pela derivada do $\log L(\boldsymbol{\zeta})$ em relação à cada componente de $\boldsymbol{\zeta}_i$. Assim, as equações de estimação para os parâmetros a_i , b_i e c_i são, respectivamente,

$$\frac{\partial \log L(\mathbf{y}|\boldsymbol{\zeta})}{\partial a_i} = 0 \Rightarrow D(1 - c_i) \sum_{j=1}^n (y_{ji} - P_{ji})(\theta_j - b_i)W_{ji} = 0 \quad (2.5)$$

$$\frac{\partial \log L(\mathbf{y}|\boldsymbol{\zeta})}{\partial b_i} = 0 \Rightarrow -Da_i(1 - c_i) \sum_{j=1}^n (y_{ji} - P_{ji})W_{ji} = 0 \quad (2.6)$$

$$\frac{\partial \log L(\mathbf{y}|\boldsymbol{\zeta})}{\partial c_i} = 0 \Rightarrow \sum_{j=1}^n (y_{ji} - P_{ji}) \frac{W_{ji}}{P_{ji}^*} = 0 \quad (2.7)$$

com $W_{ji} = \frac{P_{ji}^* Q_{ji}^*}{P_{ji} Q_{ji}}$, $P_{ji}^* = \{1 + e^{-Da_i(\theta_j - b_i)}\}^{-1}$ e $Q_{ji}^* = 1 - P_{ji}^*$.

Para a estimação das habilidades, considera-se que os parâmetros dos itens já foram calibrados (possivelmente, em outro teste), e toma-se a log-verossimilhança como função do vetor $\boldsymbol{\theta}$ e não de $\boldsymbol{\zeta}$. Assim, o EMV de θ_j é o valor que maximiza a log-verossimilhança, ou equivalentemente, é a solução da equação,

$$\frac{\partial \log L(\mathbf{y}|\boldsymbol{\theta})}{\partial \theta_j} = 0, \quad j = 1, \dots, n. \quad (2.8)$$

Logo, a equação de estimação para θ_j , $j = 1, \dots, n$ é dada por

$$D \sum_{i=1}^I a_i(1 - c_i)(y_{ji} - P_{ji})W_{ji} = 0. \quad (2.9)$$

As soluções das equações (2.5) a (2.7) e (2.9) não podem ser obtidas analiticamente, e algum procedimento numérico (por exemplo, o Newton-Raphson) deve ser utilizado para a solução.

Na próxima seção, serão discutidos os possíveis efeitos associados à mudança na posição dos itens de uma prova, e apresentadas algumas propostas existentes na literatura para a detecção e mensuração desses efeitos.

2.2 Efeito da Posição do Item (EPI)

Em muitas situações, alterar a ordem dos itens nos cadernos de provas para impedir fraudes e aumentar a segurança do teste é uma prática comum na realização de avaliações

de grande escala. Isto ocorre, por exemplo, nos 4 cadernos do ENEM, que são compostos pelos mesmos itens, mas apresentados em posições distintas nos diferentes cadernos. Esta prática nos leva a questionar se os indivíduos que responderam cadernos distintos de prova realizaram o mesmo teste, ou seja, será que esta mudança na ordem do item pode afetar as características do item e do teste, causando grandes impactos nas estimativas dos parâmetros e das notas?

Esse tema já motiva o interesse dos pesquisadores há bastante tempo, com registro dos primeiros trabalhos na literatura desde 1950 (MOLLENKOPF, 1950; MACNICOL, 1956). Uma revisão bibliográfica feita por Leary e Dorans (1985) indica que os estudos sobre o tema estão divididos em três períodos de tempo: os trabalhos iniciais investigaram o efeito da posição do item no desempenho do teste, já nos anos 1960 os trabalhos focaram em efeitos derivados de uma interação entre a posição do item e características (psicológicas e biológicas) dos respondentes e, posteriormente, os trabalhos estudaram os impactos da divulgação dos itens já aplicados no processo de equalização dos escores e o impacto da tecnologia, com a criação de testes adaptativos individuais, discutindo possíveis efeitos na estabilidade das estimativas dos parâmetros dos itens.

Dois tipos de efeitos associados a posição do item têm sido discutidos por Kington e Dorans (1984), o primeiro tipo é um efeito de aprendizagem, que ocorre quando o item torna-se mais fácil em posição posterior, ou seja, quando um determinado item ajuda a responder o item posterior, e o segundo é o efeito fadiga, que ocorre quando o item torna-se mais difícil se colocado no final da prova, por conta do cansaço/fadiga do respondente.

Debeer e Janssen (2013) apontam diferenças ao considerar efeitos associados à ordem do item e efeitos associados à posição do item. Eles afirmam que efeitos de ordem do item estão associados às características do caderno de prova, e terão impacto no escore do teste. Por outro lado, os efeitos de posição do item estão relacionados às características dos itens, produzindo impacto nas estimativas dos parâmetros do item. Nos estudos que avaliam a existência de efeitos associados à ordem do item considera-se que os cadernos de prova contém os mesmos itens, diferindo apenas na ordem em que os itens aparecem, e são aplicados a amostras aleatórias de indivíduos de uma mesma população (HANSON, 1996; DORANS e LAWRENCE, 1990; MOSES et al., 2007). O foco destes estudos é comparar o escore (total de acertos) entre os diferentes cadernos de prova. Em contrapartida, estudos que avaliam efeitos associados à posição do item têm por objetivo verificar se as estimativas

dos parâmetros do item diferem quando este item é apresentado em outra posição no caderno de prova.

Uma das grandes vantagens dos modelos da TRI na área de avaliação educacional é obter estimativas dos parâmetros dos itens e das proficiências equiparáveis, ou seja, são obtidas em uma mesma métrica (escala), mesmo quando os indivíduos que participam da avaliação respondem provas distintas, basta que hajam itens comuns fazendo uma ligação nas diferentes provas. Isso é feito através de uma metodologia incorporada no processo de estimação, chamada de Equalização, que exige invariância nas estimativas dos parâmetros dos itens nas diferentes provas. Se algum dos itens de ligação apresentar efeito de posição isto inviabiliza a equalização, pois este item poderá aparecer em posições distintas nas provas.

Nesta dissertação, o interesse é estudar os possíveis efeitos associados à posição do item. Pode-se classificar as técnicas para a detecção e mensuração desses efeitos em dois tipos: a abordagem em duas etapas e a abordagem em uma única etapa. Inicialmente, será apresentada uma breve visão da abordagem clássica em duas etapas na modelagem do EPI. Em seguida, serão apresentadas algumas propostas em uma única etapa, quando parâmetros associados ao EPI são introduzidos na função de resposta do item.

2.2.1 Abordagem clássica em duas etapas na modelagem do EPI

Alguns autores (MEYERS et al., 2012; WHITELY e DAWIS, 1976; YEN, 1980) mostraram a existência de efeito de posição de itens no modelo ML1, indicando alterações significativas no parâmetro de dificuldade dos itens. Esses estudos consideraram a abordagem em duas etapas: na primeira etapa o parâmetro de dificuldade do item é estimado separadamente em cada caderno de prova, e na segunda etapa ajusta-se um modelo que considera as diferenças nas estimativas obtidas nos diferentes cadernos como função da posição do item. Resultados obtidos por Meyers et al. (2012), usando o ML1, indicaram bons ajustes com modelos lineares, quadráticos e cúbicos para explicar as diferenças no parâmetro de dificuldade de itens de matemática e leitura em dois cadernos de prova.

Nessa abordagem, para o estudo de efeito associado à posição do item, são obtidas T estimativas \hat{b}_{ik} do parâmetro de dificuldade do item i na posição k , $i = 1, \dots, I$, $k \in \{1, 2, \dots, K\}$, onde T é o número de diferentes cadernos de prova, cada um contendo K itens. Seja \bar{b}_i a média das T estimativas do item i e \hat{b}_i a estimativa do parâmetro de dificuldade do item i obtida com base nas respostas de todos os cadernos de prova.

No procedimento em duas etapas, ajusta-se um modelo para detectar se há diferenças significativas entre essas T estimativas como função da posição em que o item aparece no caderno de prova. Geralmente, o modelo ajustado pode ser linear, quadrático ou cúbico.

Possíveis modelos lineares a serem ajustados são:

$$\widehat{b}_{is} - \widehat{b}_{it} = \beta_0 + \beta_1(s - t) \quad (2.10)$$

$$\widehat{b}_{ik} - \bar{b}_i = \beta_0 + \beta_1 k \quad (2.11)$$

$$\widehat{b}_{ik} - \widehat{b}_i = \beta_0 + \beta_1 k. \quad (2.12)$$

Assim, se houver efeito de posição do item espera-se que o parâmetro β_1 seja significativo nos modelos (2.10) a (2.12) apresentados.

Se o efeito detectado for associado à fadiga do respondente, espera-se nos modelos (2.11) e (2.12) que as diferenças $\widehat{b}_{ik} - \bar{b}_i$ e $\widehat{b}_{ik} - \widehat{b}_i$, respectivamente, sejam positivas nos valores maiores de k e negativas no valores menores de k , de forma que β_1 será positivo. Já as diferenças $\widehat{b}_{is} - \widehat{b}_{it}$ no modelo (2.10) deve ser positiva quanto maior for a distância entre as posições s e t .

Se o efeito detectado for de aprendizagem, espera-se que o item tenha uma dificuldade menor nas posições finais do caderno de prova, ou seja, as diferenças $\widehat{b}_{ik} - \bar{b}_i$ e $\widehat{b}_{ik} - \widehat{b}_i$ serão negativas nos valores maiores de k , de forma que o parâmetro β_1 será negativo.

Há algumas críticas a essa abordagem na literatura (BRENNAN, 1992; KINGSTON e DORANS, 1984; DAVEY e LEE, 2011), sugerindo que sejam incorporados parâmetros na função de resposta ao item (FRI) para avaliar possíveis efeitos significativos associados à posição do item. Algumas dessas propostas serão apresentadas na próxima seção.

2.2.2 Modelos para o EPI em uma etapa

Em vez de analisar primeiro as respostas do teste em cada caderno e depois comparar as estimativas dos parâmetros do item entre os cadernos, a utilização da abordagem da TRI em um procedimento de uma só etapa parece ser mais viável, em que o efeito-posição do item pode ser distinguido de efeitos gerais, associados às características dos respondentes e dos itens.

2.2.2.1 Modelos para o EPI via parametrização DIF na TRI

Debeer e Janssen (2013) argumentam que as diferenças detectadas na dificuldade de um item em diferentes cadernos de prova, associadas unicamente à diferença de posição

do item nos cadernos, podem ser abordadas como um caso de Funcionamento Diferencial do Item (DIF, em inglês). Modelos para itens com DIF consideram que estes itens têm características (parâmetros de dificuldade e/ou de discriminação) diferentes entre grupos de respondentes. No caso do efeito-posição do item, os grupos são formados pelos respondentes de cada caderno de prova.

Considerando o modelo ML1, a probabilidade π_{jik} de uma resposta correta do respondente \mathbf{j} ($j = 1, 2, \dots, n$) ao item \mathbf{i} ($i = 1, 2, \dots, I$) na posição \mathbf{k} ($k \in \{1, 2, \dots, K\}$) é uma função do traço latente θ_j e da dificuldade b_{ik} para o item \mathbf{i} na posição \mathbf{k} . Em forma *logito*, tem-se

$$\text{logito}[\pi_{jik}] = \theta_j - b_{ik}. \quad (2.13)$$

Com a parametrização DIF de Meulders e Xie (2004), o parâmetro de dificuldade b_{ik} em (2.13) pode ser decomposto em dois componentes,

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \delta_{ik}^b], \quad (2.14)$$

onde b_i é a dificuldade do item i na posição de referência (por exemplo, a posição do item no primeiro caderno) e δ_{ik}^b é o parâmetro DIF, ou parâmetro de posição, que modela a diferença na dificuldade do item entre a posição de referência e a posição k em outro caderno de prova.

A parametrização DIF pode ser estendida para o modelo ML2, considerando efeitos de posição do item nos parâmetros de dificuldade e de discriminação. Nesse caso, tem-se

$$\text{logito}[\pi_{jik}] = (a_i + \delta_{ik}^a)[\theta_j - (b_i + \delta_{ik}^b)], \quad (2.15)$$

onde o parâmetro δ_{ik}^a mede a mudança na discriminação do item, dependendo da sua posição k . Apesar dos estudos voltados ao efeito da posição de itens em avaliações educacionais não indicarem efeito no parâmetro de discriminação, em testes psicométricos esse efeito tem sido significativo no parâmetro de discriminação do item.

Em (2.14) e (2.15) observa-se que os efeitos associados à posição do item consideram uma interação entre o conteúdo do item e sua posição. Porém, pode-se considerar que o efeito não depende do item, apenas da sua posição. Por exemplo, em (2.14), teríamos

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \delta_k^b]. \quad (2.16)$$

Em (2.16) não foi admitida nenhum tipo de estrutura para o efeito das diferentes posições

δ_k^b na dificuldade do item, porém pode-se assumir um efeito linear na posição, ou seja,

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \gamma(k - 1)]. \quad (2.17)$$

Em (2.17), γ representa um peso linear da posição e b_i é a dificuldade do item quando ele é apresentado na primeira posição (se $k = 1$, o efeito é nulo). Se $\gamma < 0$ tem-se um efeito de aprendizagem e se $\gamma > 0$ um efeito de fadiga. Modelos não lineares, com funções quadráticas, cúbicas, exponenciais etc., também podem ser utilizados.

Outra extensão dos modelos apresentados é considerar que o efeito associado à posição do item é individual para cada respondente, ou seja, o peso γ passaria a ser γ_j , um peso específico do respondente j . Por exemplo, no ML2 teríamos

$$\text{logito}[\pi_{jik}] = a_i[\theta_j - (b_i + \gamma_j(k - 1))]. \quad (2.18)$$

Em (2.18) temos que γ_j é um efeito aleatório com distribuição normal e representa a mudança na proficiência do respondente durante o teste. Assim, esse modelo passa a ter duas dimensões, sendo possível estimar a correlação entre γ_j e θ_j .

No processo de estimação dos modelos (2.14) a (2.18) é necessário a identificabilidade dos modelos, ou seja, uma posição de referência deve ser estabelecida tal que seu efeito de posição seja fixado em zero. Em (2.14) e (2.15), deve ser definida uma posição de referência para cada item, que pode ser a posição do item no caderno de prova 1. Em (2.16) a (2.18), a posição de referência será a mesma para todos os itens em todos os cadernos. Em (2.16) pode-se tomar a primeira posição como sendo a de referência, da mesma forma que adotou-se em (2.17) e em (2.18).

2.2.2.2 Modelo Logístico Linear para o EPI, via Modelos Lineares Generalizados Mistos

Weirich, Hecht e Böhme (2014) propõe um Modelo Logístico Linear com um termo adicional de erro (MLL + ε) para explicar o efeito associado à posição do item em uma abordagem de Modelos Lineares Generalizados Mistos (MLGM). O uso de variáveis preditoras para explicar as respostas aos itens, via abordagem MLGM (MOLENBERGHS e VERBEKE, 2004), propõe modelar as respostas como função de características dos itens, características pessoais, ou uma combinação de ambas. Os efeitos das diferentes preditoras podem ser fixos ou aleatórios. O efeito associado à posição do item também pode ser modelado por algum modelo explicativo deste tipo (DE BOECK e WILSON, 2004), com

a posição do item sendo uma preditora na parte do item responsável por alguma variação na dificuldade do item. Um exemplo deste tipo de modelo é o MLL (FISCHER, 1973).

Seja o modelo de Rasch (ML1), dado por

$$P(Y_{ji} = 1) = \frac{e^{\eta_{ji}}}{1 + e^{\eta_{ji}}}, \quad (2.19)$$

onde η_{ji} é uma função da habilidade θ_j do respondente j e de β_i , a dificuldade do item i , dada por

$$\eta_{ji} = \theta_j - \beta_i. \quad (2.20)$$

Como em Debber e Janssen (2013) pode-se adicionar um parâmetro δ para o EPI,

$$\eta_{jik} = \theta_j - \beta_i - \delta_k, \quad (2.21)$$

e η_{jik} passa a ser a resposta *logito* do respondente j ao item i na posição k . O parâmetro δ_k pode ser explicado em termos da posição do item, através da combinação linear (ver De BOECK et al, 2011)

$$\eta_{jik} = \theta_j - \beta_i - \sum_{k=1}^K d_k X_k, \quad (2.22)$$

onde X_k representa a posição do item e d_k é o efeito linear da preditora X_k . A falta de um termo de erro em (2.22) implica que os coeficientes d_k são preditos sem erro pela posição do item. Esta má especificação do MLL é discutida por alguns autores (HOHENSINN et al., 2008; De BOECK, 2008), que apontaram baixo poder do teste da razão de verossimilhanças para detectar efeito de posição e ignorar a incerteza associada com a variação não-explicada do item.

Assim, a adição de um termo de erro no modelo é sugerida. Inicialmente, o modelo em (2.21) é modificado para permitir que o parâmetro associado ao EPI, além de ser dependente da posição, também seja dependente do item, ou seja,

$$\eta_{jik} = \theta_j - \beta_i - \delta_{ik} \quad (2.23)$$

e, posteriormente, toma-se

$$\eta_{jik} = \theta_j - \beta_i - \sum_{k=1}^K d_k X_k + \varepsilon_{ik}. \quad (2.24)$$

Em (2.24), com a introdução do termo ε_{ik} , o efeito δ_{ik} foi decomposto em uma parte fixa e uma parte aleatória. A parte fixa considera a posição do item na predição de δ_{ik} , e a parte aleatória reflete a incerteza nessa predição.

2.2.2.3 Modelagem Multinível para o EPI

Alexandrowicz e Matschinger (2008) também aplicaram regressão logística para estimar os efeitos associados à posição do item. Eles ajustaram um modelo de regressão logística misto para todos os I itens e todos os n indivíduos. Kamata (2001), em termos de Modelos Lineares Generalizados Hieárquicos (MLGH), afirmou que a combinação dos itens e das pessoas em um único vetor de respostas requer variáveis categóricas adicionais para indicar o item específico que corresponde a uma dada resposta ao item,

$$\eta_{ji} = \gamma_0 + \sum_{q=1}^{N-1} \gamma_q X_{qji} + u_j. \quad (2.25)$$

Neste modelo, $X_{qji} = 1$ se $q = i$ e $X_{qji} = 0$, caso contrário (q é incluído para permitir dados *missing*). Efeitos da Posição ainda não foram adicionados no modelo. O intercepto γ_0 representa o localização do item de referência, aqui será o último item no caderno. Para o item $q = i$, o modelo reduz a

$$\eta_{ji} = \gamma_0 + \gamma_q + u_j, \quad (2.26)$$

e para o item de referência, o modelo reduz a

$$\eta_{ji} = \gamma_0 + u_j. \quad (2.27)$$

Alexandrowicz e Matschinger (2008) analisaram uma versão condicional deste modelo, com efeitos principais adicionais para a posição. Assim, tem-se

$$\eta_{ij} = \gamma_0 + \sum_{q=1}^{N-1} \gamma_q X_{qji} + \sum_{r=1}^{N-1} \gamma_{r+N} X_{(r+N)ji} + u_j. \quad (2.28)$$

Neste caso, os efeitos de posição são tais que $X_{(r+N)ji} = 1$ se o item i aparece na posição r , e $X_{(r+N)ji} = 0$, caso contrário. A posição $r = N$ é omitida por ser uma posição de referência. O parâmetro γ_r representa a mudança aditiva em η_{ij} quando todos os itens aparecem na posição r .

Em (2.28), cada efeito de posição é estimado através de todos os itens. O número total de possíveis EPI será o número de itens no teste menos a posição de referência, e cada efeito de posição acontecerá para todos os itens que aparecerem na posição correspondente.

Com o objetivo de estimar simultaneamente os efeitos de item, de posição e de indivíduos, Albano (2013) adotou uma abordagem de MLGH por sua flexibilidade e por oferecer uma combinação de informação parcial na estimação desses efeitos e outros, através

de estruturas de dados complexas. O segundo objetivo em sua proposta era estimar efeitos de posição específicos do item.

Albano (2013) considerou a posição como uma covariável contínua, e os efeitos da interação posição-item tornaram-se inclinações, estimativas da mudança linear média no *logito* da resposta correta através de todas as posições na qual um item aparece.

O modelo incondicional na Equação (2.25) é escrito na forma hierárquica como

$$\begin{aligned}\eta_{ji} &= \beta_{0j} + \sum_{q=1}^{N-1} \beta_{qj} X_{qji}, \\ \beta_{0j} &= \gamma_{00} + u_{0j}, \\ \beta_{qj} &= \gamma_{q0}.\end{aligned}\tag{2.29}$$

O termo aleatório u_{0j} representa a habilidade ou traço latente do indivíduo j . A interpretação dos parâmetros γ_{00} e γ_{q0} depende dos valores das variáveis indicadoras X_{qji} . Usando um código dummy adequado (como mostrado na Tabela 1 de Albano (2013)), o intercepto γ_{00} é o efeito (dificuldade) do item de referência selecionado e γ_{q0} é o efeito do item $q = i$, expresso como uma diferença do item de referência.

Baseando-se na Equação (2.29), que serve como um modelo base ($M0$), o modelo $M1$ inclui um efeito principal para a posição β_{Nj} no nível 1, e p_{ji} é a posição ($1, 2, \dots, N$) do item $q = i$ para o examinando j :

$$\begin{aligned}\eta_{ji} &= \beta_{0j} + \sum_{q=1}^{N-1} \beta_{qj} X_{qji} + \beta_{Nj} p_{ji}, \\ \beta_{0j} &= \gamma_{00} + u_{0j}, \\ \beta_{qj} &= \gamma_{q0}, \\ \beta_{Nj} &= \gamma_{N0}.\end{aligned}\tag{2.30}$$

O efeito fixo para a posição γ_{N0} é a mudança média no desempenho através de todas as posições, e pode ser interpretada como um efeito de impacto para a posição ou um efeito de posição do teste.

Baseando-se em $M1$, o modelo $M2$ contém $N - 1$ termos adicionais de interação posição-item $\beta_{(N+q)j}$ no nível 1 para os efeitos posição-item:

$$\eta_{ji} = \beta_{0j} + \sum_{q=1}^{N-1} \beta_{qj} X_{qji} + \beta_{Nj} p_{ji} + \sum_{q=1}^{N-1} \beta_{(N+q)j} X_{qji} p_{ji},$$

$$\beta_{0j} = \gamma_{00} + u_{0j}, \quad (2.31)$$

$$\beta_{qj} = \gamma_{q0},$$

$$\beta_{Nj} = \gamma_{N0},$$

$$\beta_{(N+q)j} = \gamma_{(N+q)0}.$$

No Modelo $M2$ o intercepto novamente varia sobre os indivíduos no nível 2, com o efeito principal da posição, $N - 1$ efeitos dos itens, e $N - 1$ efeitos de interações fixas no nível 2. Combinando os modelos $M0$, $M1$ e $M2$, restrito ao item $q = i$, tem-se

$$\eta_{ji} = \gamma_{00} + \gamma_{q0} + u_{0j}, \quad (2.32)$$

$$\eta_{ji} = \gamma_{00} + \gamma_{q0} + u_{0j} + \gamma_{N0}p_{ji}, \quad (2.33)$$

$$\eta_{ji} = \gamma_{00} + \gamma_{q0} + u_{0j} + \gamma_{N0}p_{ji} + \gamma_{(N+q)0}p_{ji}. \quad (2.34)$$

2.2.2.4 Modelagem Linear e não Linear do EPI

Kang (2014) analisou a forma como diferentes representações da posição do item, ou seja, usando termos categóricos, lineares e quadráticos, podem afetar o modo como a relação entre a posição do item e a dificuldade do item é expressa, bem como o impacto dessas representações no desempenho do teste.

Segundo Kang (2014), os estudos recentes indicam 3 formas de modelar o EPI:

- Se o EPI não se altera linearmente, variáveis categóricas (indicadoras) para a posição serão mais apropriadas (ALEXANDROWICZ e MATSCHINGER, 2008; DEBEER e JANSSEN, 2013; POMPLUN e RITCHIE, 2004).
- Se o EPI tem uma alteração linear, variáveis contínuas poderão estimar melhor o EPI (ALBANO, 2013; DAVEY e LEE, 2011).
- Se o EPI tem uma alteração curvilínea, termos polinomiais para variáveis contínuas, tais como uma forma quadrática, poderão melhorar a estimação do EPI (MEYERS et al., 2012).

Sob a ótica das 3 formas de modelagem apresentadas, Kang (2014) propôs sete modelos a serem estudados, denotados por: $(M0)$, $(M1A)$, $(M1B)$, $(M2A)$, $(M2B)$, $(M3)$ e $(M4)$. Ele utilizou três tipos diferentes de indicadores para o efeito de posição simultaneamente: indicador categórico ($M1A$ e $M2A$), indicador linear ($M1B$ e $M2B$) e indicador quadrático ($M3$ e $M4$). Estudos anteriores não incluíram os três tipos diferentes de indicadores em um mesmo artigo. Também foram analisados efeitos principais ($M1A$, $M1B$ e $M3$) e efeitos de interação ($M2A$, $M2B$ e $M4$) para cada tipo de indicador. Desta forma, a variabilidade dos efeitos da posição foi analisada.

Ele considerou o modelo de Rasch como o modelo base ($M0$), expresso como

$$\text{logito}[\pi_{jik}] = \theta_j - b_{ik}. \quad (2.35)$$

O próximo modelo inclui um efeito principal adicional, com a decomposição de b_{ik} em dois componentes:

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \delta_k^b], \quad (2.36)$$

O modelo em (2.36) é um modelo de efeito-posição categórica (fixado), e será denotado por $(M1A)$, com b_i representando a dificuldade do item i na posição de referência, e δ_k^b representando o parâmetro da posição $(1, 2, \dots, N)$. Note que haverá um efeito da posição para cada posição k .

Em $(M1A)$ o modelo de efeito-posição pode também ser assumido como um efeito linear na dificuldade,

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \gamma_1(k - 1)]. \quad (2.37)$$

No modelo em (2.37), referido como $(M1B)$, γ_1 representa o peso linear da posição e b_i representa a dificuldade do item i na posição 1.

O próximo modelo adiciona efeitos de interação no efeito-posição do item,

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \delta_k^b + \delta_{ik}^b], \quad (2.38)$$

onde δ_{ik}^b representa um parâmetro associado ao efeito de interação na posição e é dependente do item. Ele modela a diferença na dificuldade do item na posição k da dificuldade na posição de referência. O δ_k^b representa um parâmetro associado ao efeito principal da posição do item. O modelo em (2.38) será chamado de $(M2A)$.

O modelo com efeitos de interação na posição pode também assumir um efeito-posição linear na dificuldade,

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \gamma_1(k - 1) + \gamma_{1i}(k - 1)], \quad (2.39)$$

onde γ_{1i} representa o peso linear do parâmetro de posição, que é dependente do item (efeito de interação). Portanto, neste modelo, chamado de (*M2B*), cada item tem a sua própria inclinação e inclui, também, o parâmetro do efeito-posição linear principal $\gamma_1(k - 1)$.

Se o modelo *M1B* não representa com precisão uma relação não-linear entre a dificuldade e a posição do item, podemos considerar um modelo, denotado por *M3*, com efeito quadrático principal associado á posição do item,

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \gamma_1(k - 1) + \gamma_2(k - 1)^2], \quad (2.40)$$

com γ_2 representando o peso quadrático do efeito-posição principal, que é aplicado a todos os itens. O modelo *M3* é como uma extensão do modelo anterior *M1B*.

O próximo modelo, denotado por *M4*, também considera uma relação quadrática entre os efeitos da posição e a dificuldade do item, e adiciona um efeito de interação associado à posição do item,

$$\text{logito}[\pi_{jik}] = \theta_j - [b_i + \gamma_1(k - 1) + \gamma_{1i}(k - 1) + \gamma_2(k - 1)^2 + \gamma_{2i}(k - 1)^2], \quad (2.41)$$

com γ_{2i} representando o peso quadrático do parâmetro de posição, que é dependente do item (efeito de interação da posição quadrática). Portanto, neste modelo cada item tem a sua própria inclinação na forma da curva.

Capítulo 3

Modelo Linear para o EPI: Estimação por Máxima Verossimilhança

3.1 Introdução

Neste capítulo apresentaremos a modelagem linear do Efeito-Posição do Item via estimação por Máxima Verossimilhança. Como ressaltado previamente o modelo que utilizaremos nesta dissertação será o ML3, que envolve os parâmetros θ , b , a e c , os quais representam o traço latente do respondente, a dificuldade, a discriminação e o acerto casual do item, respectivamente.

Na Seção 3.2 será apresentado o processo de estimação dos modelos da TRI para detectar o efeito-posição do item em duas etapas, enquanto na Seção 3.3 será abordado o processo de estimação em uma única etapa.

3.2 Estimação do EPI em duas etapas

Na Etapa 1 deste processo é necessário fazer a estimação dos parâmetros dos itens em cada caderno $t = 1, \dots, T$. Ou seja, serão obtidas as estimativas a_{it} , b_{it} e c_{it} dos parâmetros dos itens em cada caderno, ou seja, considerando apenas as respostas do respectivo caderno. Esta estimação normalmente é feita pelo método da máxima verossimilhança, baseada nas equações de estimação (2.5) a (2.7) apresentadas na Seção 2.1.1. Geralmente, considera-se que o parâmetro c_i , por ser uma probabilidade assintótica, não sofre influência da posição do item, e o mesmo não é considerado na etapa seguinte.

A Etapa 2 consiste em avaliar as diferenças entre as estimativas do parâmetro de dificuldade, b_{it} e b_{is} , entre os cadernos t e s , ($s, t = 1, \dots, T; s \neq t$), modelando-as como uma função da posição do item nos cadernos, tal como apresentado nas equações (2.10) e

(2.11). Essas diferenças também são avaliadas entre as estimativas a_{it} e a_{is} do parâmetro de discriminação. Assim, pode-se detectar se há efeito-posição do item nos parâmetros de dificuldade e de discriminação.

Outra abordagem na Etapa 2, que foi considerada neste trabalho, é obter as estimativas a_i e b_i dos parâmetros considerando os dados dos T cadernos conjuntamente. Em seguida, são avaliadas as diferenças entre as estimativas do parâmetro de dificuldade obtidas nos cadernos e as estimativas obtidas na estimação dos dados conjuntamente, $\delta_{it}^b = b_{it} - b_i$. O mesmo é feito para o parâmetro de discriminação, avaliando-se as diferenças $\delta_{it}^a = a_{it} - a_i$. O interesse é avaliar se essas diferenças δ_{it}^a e δ_{it}^b são significativas, ou seja, se há Efeito-posição do item. Neste sentido, alguns testes de hipóteses podem ser elaborados para detectar se existe evidências significativas da existência do tal efeito.

Assim, a hipótese nula H_0 de que não existe efeito-posição do item é confrontada com a hipótese alternativa H_1 de que existe efeito-posição. Isto deve ser feito para ambos os parâmetros de discriminação e de dificuldade, com as hipóteses formuladas com relação às diferenças. Para testar o efeito-posição no parâmetro de discriminação, testa-se as hipóteses

$$H_0 : \delta_{it}^a = 0 \text{ versus } H_1 : \delta_{it}^a \neq 0; \quad (3.1)$$

com $i = 1, \dots, I$ e $t = 1, \dots, T$. Para testar o efeito-posição no parâmetro de dificuldade, testa-se as hipóteses

$$H_0 : \delta_{it}^b = 0 \text{ versus } H_1 : \delta_{it}^b \neq 0; \quad (3.2)$$

com $i = 1, \dots, I$ e $t = 1, \dots, T$.

A decisão sobre a aceitação ou rejeição de uma hipótese H_0 pode ser baseada no intervalo de confiança (IC) construído para o parâmetro sob teste. No caso das hipóteses acima, pode-se construir o intervalo de confiança assintótico para cada diferença δ_{it}^a e δ_{it}^b , $i = 1, \dots, I$ e $t = 1, \dots, T$, e verificar se o valor zero pertence ou não ao intervalo, ou seja, se aceitamos ou rejeitamos H_0 . Os intervalos assintóticos para as diferenças são dados por

$$IC(\delta_{it}^a; \gamma\%) = [\widehat{\delta}_{it}^a - z_{\alpha/2}EP(\widehat{\delta}_{it}^a); \widehat{\delta}_{it}^a + z_{\alpha/2}EP(\widehat{\delta}_{it}^a)] \quad (3.3)$$

e

$$IC(\delta_{it}^b; \gamma\%) = [\hat{\delta}_{it}^b - z_{\alpha/2}EP(\hat{\delta}_{it}^b); \hat{\delta}_{it}^b + z_{\alpha/2}EP(\hat{\delta}_{it}^b)] \quad (3.4)$$

3.3 Estimação do EPI em uma etapa

Nesta seção será discutido o processo de estimação dos parâmetros em uma única etapa, via método da máxima verossimilhança. Neste processo os parâmetros associados ao efeito-posição do item já estão incluídos no modelo, como foi apresentado no Capítulo 2. Aqui as habilidades (proficiências) serão consideradas fixadas, e o interesse é estimar os parâmetros dos itens.

Nesta abordagem, δ_{ik}^a é um parâmetro que mede a mudança na discriminação do item i , dependendo da sua posição $k \in \{1, \dots, K\}$ dentro do caderno $t = 1, \dots, T$. A notação natural $k(t)$ será simplificada simplesmente por k , lembrando que há apenas T valores de k no conjunto $\{1, 2, \dots, K\}$, digamos k_1, k_2, \dots, k_T . De forma similar, δ_{ik}^b será o parâmetro de posição do modelo, associado à mudança na dificuldade do item entre a posição de referência e a posição k no caderno. Para resolver problemas de identificabilidade, essa modelagem requer que sejam estabelecidas restrições na parametrização. Nesse caso, as estimativas estarão sujeitas às seguintes restrições: $\sum_k \delta_{ik}^a = 0$ e $\sum_k \delta_{ik}^b = 0$.

O modelo completo utilizado para mensurar possíveis efeitos associados a posição do item será uma extensão do ML3, dado por

$$P_{jik} = P(Y_{jik} = 1 | \theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-D(a_i + \delta_{ik}^a)[\theta_j - (b_i + \delta_{ik}^b)]}}, \quad (3.5)$$

com $i = 1, 2, \dots, I$, $j = 1, 2, \dots, n$ e $k = 1, \dots, K$, e a variável Y_{jik} representa a variável dicotômica associada à resposta dada pelo indivíduo j ao item i , na posição k .

Para este modelo serão estimados os parâmetros dos itens $\zeta = (\zeta_1, \dots, \zeta_I)$, supondo as habilidades conhecidas. Neste caso, para cada item tem-se $2T + 3$ parâmetros a serem estimados, pois $\zeta_i = (a_i, b_i, c_i, \delta_{ik_1}^a, \dots, \delta_{ik_T}^a, \delta_{ik_1}^b, \dots, \delta_{ik_T}^b)$ para $i = 1, \dots, I$.

Adotando a independência entre as respostas de diferentes indivíduos (**S1**), assim como a independência local (**S2**), a função de verossimilhança pode ser expressa por

$$L(\mathbf{y}|\zeta) = P(\mathbf{Y}_{\dots} = \mathbf{y}_{\dots} | \boldsymbol{\theta}, \zeta) \quad (3.6)$$

$$= \prod_{k \in \{k_1, \dots, k_T\}} P(\mathbf{Y}_{..k} = \mathbf{y}_{..k} | \boldsymbol{\theta}_k, \zeta) \quad (3.7)$$

$$= \prod_{k \in \{k_1, \dots, k_T\}} \prod_{j=1}^{n_k} P(Y_{j.k} = y_{j.k} | \boldsymbol{\theta}_k, \zeta) \quad (3.8)$$

$$= \prod_{k \in \{k_1, \dots, k_T\}} \prod_{j=1}^{n_k} \prod_{i=1}^I P(Y_{jik} = y_{jik} | \theta_j, \zeta_i) \quad (3.9)$$

$$= \prod_{k \in \{k_1, \dots, k_T\}} \prod_{j=1}^{n_k} \prod_{i=1}^I P_{jik}^{y_{jik}} (1 - P_{jik})^{1-y_{jik}}, \quad (3.10)$$

com $\boldsymbol{\theta}_k$ representando o vetor de habilidades (proficiências) dos n_k indivíduos que responderam itens na posição k . Segue de (3.10) que a log-verossimilhança pode ser escrita como:

$$l(\mathbf{y}|\zeta) = \log L(\mathbf{y}|\zeta) \quad (3.11)$$

$$= \sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} \sum_{i=1}^I (y_{jik}) \log P_{jik} + (1 - y_{jik}) \log(1 - P_{jik}). \quad (3.12)$$

Os estimadores de máxima verossimilhança (EMV) de ζ_i , $i = 1, \dots, I$, são os valores que maximizam a função de verossimilhança ou, equivalentemente, são as soluções das equações de estimação

$$\frac{\partial l(\mathbf{y}|\zeta)}{\partial \zeta_i} = 0. \quad (3.13)$$

Assim, substituindo P_{jik} por sua expressão apresentada em (3.5), e resolvendo (3.13) individualmente para cada um dos parâmetros a_i , b_i , c_i , δ_{ik}^a e δ_{ik}^b , $t = 1, \dots, T$, as equações

de estimação podem ser expressas por

$$a_i : D(1 - c_i) \sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) [\theta_j - (b_i + \delta_{ik}^b)] W_{jik} = 0, \quad (3.14)$$

$$b_i : -D(1 - c_i) \sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) (a_i + \delta_{ik}^a) W_{jik} = 0, \quad (3.15)$$

$$c_i : \sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) \frac{W_{jik}}{P_{jik}^*} = 0, \quad (3.16)$$

$$\delta_{ik}^a : D(1 - c_i) \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) [\theta_j - (b_i + \delta_{ik}^b)] W_{jik} = 0, \quad \forall k \in \{k_1, \dots, k_T\}, \quad (3.17)$$

$$\delta_{ik}^b : -D(1 - c_i) \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) (a_i + \delta_{ik}^a) W_{jik} = 0, \quad \forall k \in \{k_1, \dots, k_T\}, \quad (3.18)$$

com $W_{jik} = \frac{P_{jik}^* Q_{jik}^*}{P_{jik} Q_{jik}}$, $P_{jik}^* = \{1 + e^{-D(a_i + \delta_{ik}^a)[\theta_j - (b_i + \delta_{ik}^b)]}\}^{-1}$ e $Q_{jik}^* = 1 - P_{jik}^*$. Para evitar o problema de falta de identificabilidade, as equações (3.17) e (3.18) estão sujeitas às restrições: $\sum_k \delta_{ik}^a = 0$ e $\sum_k \delta_{ik}^b = 0$.

De forma geral, as equações (3.14) a (3.18) não possuem solução explícita, sendo necessário utilizar algum método iterativo para a obtenção das soluções, tais como o método de Newton-Raphson (ANDRADE et al., 2000, pg.37) ou o Algoritmo EM (CASELLA e BERGER, 2011, pg. 291).

A implementação desses métodos iterativos requer o conhecimento do erro padrão dos estimadores, obtido a partir da inversa da matriz de informação de Fisher

$$I(\boldsymbol{\zeta}_i) = -E \left(\frac{\partial^2 \log L(\mathbf{Y}|\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \right), \quad (3.19)$$

que depende da segunda derivada do logaritmo da função de verossimilhança, dada por

$$\frac{\partial^2 \log L(\mathbf{y}|\boldsymbol{\zeta})}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} = \sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} \left\{ \left[\frac{\partial}{\partial \boldsymbol{\zeta}_i} \left(\frac{y_{jik} - P_{jik}}{P_{jik} Q_{jik}} \right) \right] \left(\frac{\partial P_{jik}}{\partial \boldsymbol{\zeta}_i} \right) + \left(\frac{y_{jik} - P_{jik}}{P_{jik} Q_{jik}} \right) \left(\frac{\partial^2 P_{jik}}{\partial \boldsymbol{\zeta}_i \partial \boldsymbol{\zeta}_i'} \right) \right\}. \quad (3.20)$$

Em (3.20) as derivadas $\frac{\partial P_{jik}}{\partial \boldsymbol{\zeta}_i}$ são dadas por

$$\frac{\partial P_{jik}}{\partial a_i} = \frac{\partial P_{jik}}{\partial \delta_{ik}^a} = D(1 - c_i)[\theta_j - (b_i + \delta_{ik}^b)]P_{jik}^* Q_{jik}^*, \quad (3.21)$$

$$\frac{\partial P_{jik}}{\partial b_i} = \frac{\partial P_{jik}}{\partial \delta_{ik}^b} = -D(1 - c_i)(a_i + \delta_{ik}^a)P_{jik}^* Q_{jik}^*, \quad (3.22)$$

$$\frac{\partial P_{jik}}{\partial c_i} = Q_{jik}^*, \quad (3.23)$$

$$(3.24)$$

e as derivadas de segunda ordem $\frac{\partial^2 P_{jik}}{\partial \zeta_i \partial \zeta_i}$ são dadas por

$$\frac{\partial^2 P_{jik}}{\partial a_i^2} = D^2(1 - c_i)[\theta_j - (b_i + \delta_{ik}^b)]^2 P_{jik}^* Q_{jik}^* (1 - 2P_{jik}^*), \quad (3.25)$$

$$\frac{\partial^2 P_{jik}}{\partial a_i \partial b_i} = -D(1 - c_i)P_{jik}^* Q_{jik}^* \{1 + D(a_i + \delta_{ik}^a)[\theta_j - (b_i + \delta_{ik}^b)](1 - 2P_{jik}^*)\}, \quad (3.26)$$

$$\frac{\partial^2 P_{jik}}{\partial a_i \partial c_i} = -D[\theta_j - (b_i + \delta_{ik}^b)]P_{jik}^* Q_{jik}^*, \quad (3.27)$$

$$\frac{\partial^2 P_{jik}}{\partial b_i^2} = D^2(1 - c_i)(a_i + \delta_{ik}^a)^2 P_{jik}^* Q_{jik}^* (1 - 2P_{jik}^*), \quad (3.28)$$

$$\frac{\partial^2 P_{jik}}{\partial b_i \partial c_i} = D(a_i + \delta_{ik}^a)P_{jik}^* Q_{jik}^*, \quad (3.29)$$

$$\frac{\partial^2 P_{jik}}{\partial c_i^2} = \frac{\partial Q_{jik}^*}{\partial c_i} = 0, \quad (3.30)$$

$$\frac{\partial^2 P_{jik}}{\partial (\delta_{ik}^a)^2} = \frac{\partial^2 P_{jik}}{\partial \delta_{ik}^a \partial a_i} = D^2(1 - c_i)[\theta_j - (b_i + \delta_{ik}^b)]^2 P_{jik}^* Q_{jik}^* (1 - 2P_{jik}^*), \quad (3.31)$$

$$\frac{\partial^2 P_{jik}}{\partial \delta_{ik}^a \partial b_i} = -D(1 - c_i)P_{jik}^* Q_{jik}^* \{1 + D(a_i + \delta_{ik}^a)[\theta_j - (b_i + \delta_{ik}^b)](1 - 2P_{jik}^*)\}, \quad (3.32)$$

$$\frac{\partial^2 P_{jik}}{\partial \delta_{ik}^a \partial c_i} = -D[\theta_j - (b_i + \delta_{ik}^b)]P_{jik}^* Q_{jik}^*, \quad (3.33)$$

$$\frac{\partial^2 P_{jik}}{\partial (\delta_{ik}^b)^2} = \frac{\partial^2 P_{jik}}{\partial \delta_{ik}^b \partial b_i} = D^2(1 - c_i)(a_i + \delta_{ik}^a)^2 P_{jik}^* Q_{jik}^* (1 - 2P_{jik}^*), \quad (3.34)$$

$$\frac{\partial^2 P_{jik}}{\partial \delta_{ik}^b \partial c_i} = D(a_i + \delta_{ik}^a)P_{jik}^* Q_{jik}^*. \quad (3.35)$$

Capítulo 4

Modelo Linear para o EPI: Estimação Bayesiana

No capítulo anterior a estimação dos parâmetros do ML3 foi apresentada através do método da Máxima Verossimilhança. No entanto, Andrade, Tavares e Valle (2000) discutem problemas com esse método na presença de escores nulos ou perfeitos, indicando a possibilidade de obter estimativas de máxima verossimilhança dos parâmetros dos itens fora do espaço paramétrico para itens que são respondidos incorretamente ou corretamente por todos os respondentes, bem como problemas na estimação das habilidade de indivíduos que erram ou acertam todos os itens. Essa limitação pode ser contornada com o uso da abordagem bayesiana, que será tratada neste capítulo.

4.1 Estimação Bayesiana

O processo de estimação no modelo (3.5) sob uma abordagem Bayesiana, consiste em estabelecer distribuições *a priori* para os parâmetros de interesse, construir uma nova função denominada distribuição *a posteriori* e estimar os parâmetros de interesse com base em alguma característica dessa distribuição (PAULINO et al., 2003), geralmente a moda ou a média da distribuição. Para isso, seja $\Psi = (\boldsymbol{\theta}, \boldsymbol{\zeta})$, o vetor contendo todos os parâmetros do modelo, onde $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ é o vetor de traços latentes dos n respondentes e $\boldsymbol{\zeta} = \{\boldsymbol{\zeta}_1, \dots, \boldsymbol{\zeta}_I\}$, com $\boldsymbol{\zeta}_i = (a_i, b_i, c_i, \delta_{ik}^a, \delta_{ik}^b)$, $i = 1, \dots, I$ e $k \in \{k_1, \dots, k_T\}$, o vetor de parâmetros de todos os I itens.

A função de verossimilhança obtida a partir do modelo é dada por

$$L(\mathbf{Y}|\Psi) = \prod_{k \in \{k_1, \dots, k_T\}} \prod_{j=1}^n \prod_{i \in I_{(j)}} p(y_{jik} | \theta_j, a_i, b_i, c_i, \delta_{ik}^a, \delta_{ik}^b), \quad (4.1)$$

onde $I_{(j)}$ representa o conjunto de itens respondidos pelo indivíduo j e $p(y_{jik} | \theta_j, a_i, b_i, c_i, \delta_{ik}^a, \delta_{ik}^b) = P(Y_{jik} = y_{jik} | \theta_j, a_i, b_i, c_i, \delta_{ik}^a, \delta_{ik}^b)$.

Assumindo as habilidades conhecidas, como foi feito no capítulo anterior, o interesse é apenas na estimação dos parâmetros dos itens. Assim, o vetor de parâmetros passa a ser $\Psi = \zeta$, sendo necessário estabelecer distribuições *a priori* para os seus componentes.

4.1.1 Distribuições *a priori*

As distribuições *a priori* para os parâmetros dos itens aqui adotadas levarão em conta o espaço paramétrico e os modelos geralmente utilizados na literatura (FOX, 2010). Assim, usualmente adota-se uma distribuição Log-normal (LN) para o parâmetro a_i , cujo espaço paramétrico é o conjunto dos reais positivos. Como os parâmetros b_i , δ_{ik}^a e δ_{ik}^b assumem valores no conjunto dos reais, em geral adota-se uma distribuição *a priori* Normal (N) para os mesmos. Já o parâmetro c_i , que representa uma probabilidade e só pode assumir valores no intervalo $[0,1]$, adota-se uma distribuição *a priori* Beta.

Assim, tem-se $a_i \sim LN(m_a, \sigma_a^2)$, cuja função densidade de probabilidade (fdp) *a priori* é dada por

$$p(a_i) = p(a_i|m_a, \sigma_a^2) = \frac{1}{\sqrt{2\pi}a_i\sigma_a} \exp\left[-\frac{1}{2\sigma_a^2}(\log a_i - m_a)^2\right], \quad i = 1, \dots, I, \quad (4.2)$$

o parâmetro $b_i \sim N(m_b, \sigma_b^2)$, com fdp *a priori*

$$p(b_i) = p(b_i|m_b, \sigma_b^2) = \frac{1}{\sqrt{2\pi}\sigma_b} \exp\left[-\frac{1}{2\sigma_b^2}(b_i - m_b)^2\right], \quad i = 1, \dots, I, \quad (4.3)$$

o parâmetro $c_i \sim \text{Beta}(\alpha - 1, \beta - 1)$, com fdp *a priori*

$$p(c_i) = p(c_i|\alpha, \beta) = \frac{\Gamma(\alpha + \beta - 2)}{\Gamma(\alpha - 1)\Gamma(\beta - 1)} c_i^{\alpha-2} (1 - c_i)^{\beta-2}, \quad i = 1, \dots, I. \quad (4.4)$$

Os parâmetros $\delta_{ik}^a \sim N(0, \sigma_{\delta^a}^2)$ e $\delta_{ik}^b \sim N(0, \sigma_{\delta^b}^2)$ têm fdp's *a priori* dadas, respectivamente, por

$$p(\delta^a) = p(\delta_{ik}^a|\sigma_{\delta^a}^2) = \frac{1}{\sqrt{2\pi}\sigma_{\delta^a}} \exp\left[-\frac{1}{2\sigma_{\delta^a}^2}(\delta_{ik}^a)^2\right], \quad i = 1, \dots, I, \quad k \in \{k_1, \dots, k_T\}, \quad (4.5)$$

e

$$p(\delta^b) = p(\delta_{ik}^b|\sigma_{\delta^b}^2) = \frac{1}{\sqrt{2\pi}\sigma_{\delta^b}} \exp\left[-\frac{1}{2\sigma_{\delta^b}^2}(\delta_{ik}^b)^2\right], \quad i = 1, \dots, I, \quad k \in \{k_1, \dots, k_T\}. \quad (4.6)$$

Assumindo independência entre os parâmetros dos itens, a distribuição *a priori* conjunta para o vetor de parâmetros é expressa pela fdp

$$p(\zeta) = \prod_{i=1}^I p(a_i)p(b_i)p(c_i)p(\delta^a)p(\delta^b). \quad (4.7)$$

4.1.2 Distribuição *a posteriori* conjunta para os parâmetros dos itens

A distribuição *a posteriori* conjunta dos parâmetros dos itens ζ é obtida através do Teorema de Bayes, com fdp conjunta *a posteriori* expressa por

$$p(\zeta|\mathbf{Y}) = \frac{L(\mathbf{Y}|\zeta) \times p(\zeta)}{\int \dots \int L(\mathbf{Y}|\zeta) \times p(\zeta) d\zeta}, \quad (4.8)$$

que pode ser simplificada para

$$p(\zeta|\mathbf{Y}) \propto L(\mathbf{Y}|\zeta) \times p(\zeta) \quad (4.9)$$

uma vez que o denominador em (4.8) não depende do vetor ζ .

Neste trabalho, o estimador Bayesiano para o vetor de parâmetros ζ será obtido pela moda da distribuição *a posteriori* conjunta, ou seja, serão os valores que maximizam $p(\zeta|\mathbf{Y})$ ou, equivalentemente, que maximizam $\log p(\zeta|\mathbf{Y}) \propto \log L(\mathbf{Y}|\zeta) + \log p(\zeta)$.

Logo, os estimadores Bayesianos são obtidos pela solução do sistema de equações

$$\frac{\partial \log p(\zeta|\mathbf{Y})}{\partial \zeta} = 0, \quad (4.10)$$

$$\frac{\partial \log L(\mathbf{Y}|\zeta)}{\partial \zeta} + \frac{\partial \log p(\zeta)}{\partial \zeta} = 0. \quad (4.11)$$

4.1.3 Equações de estimação dos parâmetros dos itens

Nesta seção serão apresentadas as equações de estimação para a obtenção dos estimadores bayesianos para cada componente do vetor $\zeta_i = (a_i, b_i, c_i, \delta_{ik}^a, \delta_{ik}^b)$, com base em (4.11):

$$\frac{\partial \log L(\mathbf{Y}|\zeta)}{\partial a_i} + \frac{\partial \log p(\zeta)}{\partial a_i} = 0, \quad (4.12)$$

$$\frac{\partial \log L(\mathbf{Y}|\zeta)}{\partial b_i} + \frac{\partial \log p(\zeta)}{\partial b_i} = 0, \quad (4.13)$$

$$\frac{\partial \log L(\mathbf{Y}|\zeta)}{\partial c_i} + \frac{\partial \log p(\zeta)}{\partial c_i} = 0, \quad (4.14)$$

$$\frac{\partial \log L(\mathbf{Y}|\zeta)}{\partial \delta_{ik}^a} + \frac{\partial \log p(\zeta)}{\partial \delta_{ik}^a} = 0, \quad (4.15)$$

$$\frac{\partial \log L(\mathbf{Y}|\zeta)}{\partial \delta_{ik}^b} + \frac{\partial \log p(\zeta)}{\partial \delta_{ik}^b} = 0, \quad (4.16)$$

As primeiras parcelas nas equações (4.12) a (4.16) foram obtidas na Seção 3.3, na estimação via Método da Máxima Verossimilhança, e são apresentadas no membro esquerdo das equações (3.14) a (3.18), respectivamente. A abordagem bayesiana adiciona uma nova parcela relativa a distribuição *a priori* em cada equação de estimação.

De (4.2) a (4.7), tem-se que

$$\frac{\partial \log p(\zeta)}{\partial a_i} = \frac{\partial \log p(a_i)}{\partial a_i} = -\frac{1}{a_i} \left[1 + \frac{\log a_i - m_a}{\sigma_a^2} \right], \quad (4.17)$$

$$\frac{\partial \log p(\zeta)}{\partial b_i} = \frac{\partial \log p(b_i)}{\partial b_i} = -\frac{b_i - m_b}{\sigma_b^2}, \quad (4.18)$$

$$\frac{\partial \log p(\zeta)}{\partial c_i} = \frac{\partial \log p(c_i)}{\partial c_i} = \frac{\alpha - 2}{c_i} - \frac{\beta - 2}{1 - c_i} \quad (4.19)$$

$$\frac{\partial \log p(\zeta)}{\partial \delta_{ik}^a} = \frac{\partial \log p(\delta^a)}{\partial \delta_{ik}^a} = -\frac{\delta_{ik}^a}{\sigma_{\delta^a}^2}, \quad (4.20)$$

$$\frac{\partial \log p(\zeta)}{\partial \delta_{ik}^b} = \frac{\partial \log p(\delta^b)}{\partial \delta_{ik}^b} = -\frac{\delta_{ik}^b}{\sigma_{\delta^b}^2}, \quad (4.21)$$

Assim, substituindo (3.14) a (3.18) e (4.17) a (4.21) no sistema de equações (4.12) a (4.16), as equações de estimação para a obtenção dos estimadores bayesianos dos parâmetros a_i , b_i , c_i , δ_{ik}^a e δ_{ik}^b são, respectivamente,

$$D(1 - c_i) \sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) [\theta_j - (b_i + \delta_{ik}^b)] W_{jik} - \frac{1}{a_i} \left[1 + \frac{\log a_i - m_a}{\sigma_a^2} \right] = 0, \quad (4.22)$$

$$-D(1 - c_i) \sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) (a_i + \delta_{ik}^a) W_{jik} - \frac{b_i - m_b}{\sigma_b^2} = 0, \quad (4.23)$$

$$\sum_{k \in \{k_1, \dots, k_T\}} \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) \frac{W_{jik}}{P_{jik}^*} + \frac{\alpha - 2}{c_i} - \frac{\beta - 2}{1 - c_i} = 0, \quad (4.24)$$

$$D(1 - c_i) \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) [\theta_j - (b_i + \delta_{ik}^b)] W_{jik} - \frac{\delta_{ik}^a}{\sigma_{\delta^a}^2} = 0, \quad (4.25)$$

$$D(1 - c_i) \sum_{j=1}^{n_k} (y_{jik} - P_{jik}) (a_i + \delta_{ik}^a) W_{jik} - \frac{\delta_{ik}^b}{\sigma_{\delta^b}^2} = 0. \quad (4.26)$$

O sistema (4.22) a (4.26) não tem solução analítica, e algum processo numérico é necessário para sua solução.

Após a obtenção das estimativas bayesianas, o próximo passo é verificar se os parâmetros δ_{ik}^a e δ_{ik}^b são significativos, ou seja, se existe efeito-posição nos parâmetros de discriminação

e/ou dificuldade de algum item em alguma posição. Para isso, testam-se as hipóteses (3.1) e (3.2), apresentadas na Seção 3.2. Na abordagem bayesiana, a decisão sobre a rejeição ou não das hipóteses nulas baseia-se nos intervalos de credibilidade (ICr) para os respectivos parâmetros, verificando se os mesmos contém ou não o valor zero.

Os intervalos com $\gamma\%$ de credibilidade são obtidos na distribuição posterior marginal dos parâmetros δ_{ik}^a e δ_{ik}^b através dos percentis de ordem $\frac{(1-\gamma)}{2}$ e $\frac{(1+\gamma)}{2}$.

Assim, o $ICr(\delta_{ik}^a; \gamma\%) = (l_{inf}^a, l_{sup}^a)$, onde l_{inf}^a e l_{sup}^a são, respectivamente, os percentis de ordem $\frac{(1-\gamma)}{2}$ e $\frac{(1+\gamma)}{2}$ na distribuição

$$p(\delta_{ik}^a | \mathbf{Y}) = \int p(\boldsymbol{\zeta} | \mathbf{Y}) d\boldsymbol{\zeta}_{(-\delta_{ik}^a)}. \quad (4.27)$$

E o $ICr(\delta_{ik}^b; \gamma\%) = (l_{inf}^b, l_{sup}^b)$, onde l_{inf}^b e l_{sup}^b são, respectivamente, os percentis de ordem $\frac{(1-\gamma)}{2}$ e $\frac{(1+\gamma)}{2}$ na distribuição

$$p(\delta_{ik}^b | \mathbf{Y}) = \int p(\boldsymbol{\zeta} | \mathbf{Y}) d\boldsymbol{\zeta}_{(-\delta_{ik}^b)}. \quad (4.28)$$

Em (4.27) o vetor $\boldsymbol{\zeta}_{(-\delta_{ik}^a)}$ representa o vetor $\boldsymbol{\zeta}$ sem o parâmetro δ_{ik}^a e em (4.28) o vetor $\boldsymbol{\zeta}_{(-\delta_{ik}^b)}$ representa o vetor $\boldsymbol{\zeta}$ sem o parâmetro δ_{ik}^b .

Capítulo 5

Resultados e Discussões

5.1 Aplicação a dados simulados

Nesta seção são apresentados os resultados do processo de estimação dos parâmetros dos itens, considerando as habilidades conhecidas, para um conjunto de dados (respostas) simulados. Todo o processo de estimação e simulação foi implementado no software R. Seguem abaixo algumas informações sobre a simulação dos dados e a construção do algoritmo de estimação:

a) Para o processo da estimação dos parâmetros, foi necessário a utilização de alguns pacotes específicos do R, que são: *mirt*, *irttoys*, *maxLik*, *readxl*, *plotrix*, *foreign*, *latex2exp* e *bbmle*;

b) Os dados (vetor de respostas) foram simulados considerando o número de cadernos $T = 4$, cada caderno com $I = 45$ itens, com $n = 10.000$ respondentes por caderno e o número de réplicas $R = 100$;

c) As habilidades, consideradas conhecidas, foram geradas de uma distribuição Normal com média zero e variância um, $N(0, 1)$;

d) As estimativas iniciais, utilizadas no processo de estimação dos parâmetros a_i , b_i e c_i , foram as estimativas obtidas considerando as respostas de todos os cadernos conjuntamente, e para os parâmetros δ_{ik}^a e δ_{ik}^b adotou-se estimativas iniciais iguais a zero;

e) Na solução das equações de estimação, via máxima verossimilhança e via estimação bayesiana, o método numérico utilizado foi o *algoritmo EM*.

Na Tabela 5.1, constam os valores (verdadeiros) dos parâmetros dos itens adotados para a simulação dos dados, ou seja, espera-se que as estimativas obtidas no processo de estimação sejam muito próximas desses valores. Caso isto aconteça pode-se afirmar que o processo de estimação está funcionando bem, produzindo erros pequenos de estimação.

A Tabela 5.1 mostra que para os parâmetros de discriminação dos itens foram considerados apenas 3 valores: 0,75, 1,00 e 2,00. Para os parâmetros de dificuldade foram adotados 5 valores: -2,00, -1,00, 0,00, 1,00 e 2,00. E para os parâmetros do acerto casual adotou-se 0,00, 0,10 e 0,20. Foram selecionados dezoito itens dentre os 45 para apresentarem efeito-posição do item, dos quais seis itens apresentaram o efeito-posição apenas no parâmetro de discriminação (δ^a), que foram os itens 1, 8, 15, 31, 38 e 45, seis itens apresentaram o efeito-posição no parâmetro de dificuldade (δ^b), são os itens 4, 11, 19, 26, 36 e 44, e seis itens apresentaram efeito-posição em ambos os parâmetros de discriminação e de dificuldade, que são os itens 5, 12, 20, 27, 37 e 43.

Os resultados do processo de estimação via máxima verossimilhança e via metodologia bayesiana são apresentados nas Seções 5.1.1 e 5.1.2, respectivamente.

5.1.1 Estimativas via máxima verossimilhança

A Tabela 5.2 apresenta as estimativas médias dos parâmetros dos itens obtidas nas 100 réplicas. Na Tabela 5.3 consta o desvio-padrão das estimativas dos parâmetros dos itens. Observa-se que os desvios-padrão apresentam valores pequenos, ou seja, quanto menor o desvio padrão, menor a dispersão das estimativas obtidas em torno da estimativa média.

As diferenças entre as estimativas médias e os valores (verdadeiros) adotados para os parâmetros dos itens, ou seja, os erros na estimação, são apresentados na Tabela 5.4. Observa-se que os erros foram pequenos, e a maior diferença em valor absoluto observada foi 0,149, associada a estimação do parâmetro b no item 1. Assim, tem-se que o processo de estimação via máxima verossimilhança está funcionando bem, pois produziu erros pequenos e estimativas com pouca variabilidade (com maior precisão).

Tabela 5.1 *Valores adotados para os parâmetros dos itens na simulação dos dados.*

ITEM	a	b	c	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	0,75	-2,00	0,0	-0,40	-0,30	0,50	0,20	0,00	0,00	0,00	0,00
Item 2	0,75	-2,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 3	0,75	-2,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 4	0,75	-1,00	0,0	0,00	0,00	0,00	0,00	0,40	-0,10	-0,40	0,10
Item 5	0,75	-1,00	0,1	-0,50	0,30	0,00	0,20	0,50	-0,30	0,00	-0,20
Item 6	0,75	-1,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 7	0,75	0,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 8	0,75	0,00	0,1	-0,40	-0,10	0,10	0,40	0,00	0,00	0,00	0,00
Item 9	0,75	0,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 10	0,75	1,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 11	0,75	1,00	0,1	0,00	0,00	0,00	0,00	0,20	-0,10	0,40	-0,50
Item 12	0,75	1,00	0,2	-0,10	0,30	-0,30	0,10	0,10	-0,30	0,30	-0,10
Item 13	0,75	2,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 14	0,75	2,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 15	0,75	2,00	0,2	-0,30	0,10	0,30	-0,10	0,00	0,00	0,00	0,00
Item 16	1,00	-2,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 17	1,00	-2,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 18	1,00	-2,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 19	1,00	-1,00	0,0	0,00	0,00	0,00	0,00	0,00	-0,20	0,20	0,00
Item 20	1,00	-1,00	0,1	-0,10	-0,30	0,30	0,10	0,10	0,30	-0,30	-0,10
Item 21	1,00	-1,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 22	1,00	0,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 23	1,00	0,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 24	1,00	0,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 25	1,00	1,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 26	1,00	1,00	0,1	0,00	0,00	0,00	0,00	-0,20	0,10	-0,10	0,20
Item 27	1,00	1,00	0,2	0,20	-0,20	0,40	-0,40	-0,20	0,20	-0,40	0,40
Item 28	1,00	2,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 29	1,00	2,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 30	1,00	2,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 31	2,00	-2,00	0,0	0,10	-0,20	-0,30	0,40	0,00	0,00	0,00	0,00
Item 32	2,00	-2,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 33	2,00	-2,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 34	2,00	-1,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 35	2,00	-1,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 36	2,00	-1,00	0,2	0,00	0,00	0,00	0,00	-0,40	0,30	-0,10	0,20
Item 37	2,00	0,00	0,0	0,40	-0,20	0,10	-0,30	-0,40	0,20	-0,10	0,30
Item 38	2,00	0,00	0,1	0,50	-0,30	-0,10	-0,10	0,00	0,00	0,00	0,00
Item 39	2,00	0,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 40	2,00	1,00	0,0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 41	2,00	1,00	0,1	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 42	2,00	1,00	0,2	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Item 43	2,00	2,00	0,0	0,20	0,10	-0,20	-0,10	-0,20	-0,10	0,20	0,10
Item 44	2,00	2,00	0,1	0,00	0,00	0,00	0,00	-0,30	-0,20	0,60	-0,10
Item 45	2,00	2,00	0,2	0,30	0,20	-0,60	0,10	0,00	0,00	0,00	0,00

Fonte: Dados Simulados.

Tabela 5.2 *Estimativas médias dos parâmetros dos itens para os dados simulados, via método da máxima verossimilhança.*

ITEM	a	b	c	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	0,767	-1,851	0,045	-0,407	-0,302	0,510	0,199	0,090	0,056	-0,080	-0,067
Item 2	0,762	-1,946	0,114	0,011	-0,007	-0,004	0,000	0,021	-0,018	-0,008	0,005
Item 3	0,758	-1,977	0,203	-0,003	0,003	0,001	0,000	-0,006	0,010	0,005	-0,009
Item 4	0,761	-0,950	0,018	0,000	0,001	-0,003	0,002	0,400	-0,093	-0,408	0,101
Item 5	0,748	-1,045	0,089	-0,500	0,305	-0,002	0,197	0,443	-0,273	0,013	-0,183
Item 6	0,753	-1,010	0,197	-0,005	0,001	0,004	0,000	-0,003	0,003	0,001	-0,002
Item 7	0,758	0,022	0,008	0,003	-0,003	0,002	-0,003	0,002	-0,001	-0,001	0,000
Item 8	0,754	-0,001	0,100	-0,398	-0,104	0,100	0,403	-0,001	-0,002	0,000	0,003
Item 9	0,746	-0,032	0,192	-0,001	-0,002	0,002	0,001	-0,003	0,001	0,001	0,002
Item 10	0,761	1,013	0,005	-0,002	0,003	-0,002	0,000	0,103	0,291	-0,293	-0,101
Item 11	0,745	0,997	0,097	0,006	0,001	-0,004	-0,002	0,192	-0,102	0,413	-0,502
Item 12	0,755	0,992	0,199	-0,101	0,303	-0,304	0,102	0,102	-0,296	0,297	-0,102
Item 13	0,762	1,997	0,002	0,001	0,001	0,000	-0,002	-0,002	-0,003	-0,001	0,006
Item 14	0,751	2,001	0,099	0,000	0,002	0,001	-0,003	0,002	0,003	-0,013	0,008
Item 15	0,751	2,002	0,199	-0,301	0,099	0,305	-0,102	0,002	0,007	-0,006	-0,002
Item 16	1,018	-1,937	0,032	0,002	-0,001	0,001	-0,001	-0,002	0,006	0,002	-0,006
Item 17	1,005	-1,987	0,106	0,001	-0,001	0,002	-0,003	0,007	-0,001	0,004	-0,009
Item 18	1,009	-1,985	0,201	-0,001	0,003	0,000	-0,002	-0,004	0,006	-0,005	0,003
Item 19	1,009	-0,978	0,010	0,002	-0,004	0,001	0,001	0,000	-0,201	0,200	0,000
Item 20	1,005	-0,995	0,102	-0,100	-0,305	0,302	0,102	0,001	0,001	-0,002	-0,001
Item 21	1,001	-1,007	0,197	-0,004	0,003	0,002	0,000	0,000	0,000	-0,001	0,002
Item 22	1,010	0,011	0,005	0,001	-0,004	-0,001	0,004	0,000	0,000	0,002	-0,001
Item 23	0,999	-0,005	0,098	-0,001	0,000	0,001	0,000	-0,002	0,000	0,003	-0,001
Item 24	1,001	0,000	0,200	0,000	-0,001	-0,001	0,001	-0,003	0,003	0,003	-0,003
Item 25	1,009	1,001	0,002	0,001	-0,003	0,003	-0,001	-0,001	0,002	0,000	-0,001
Item 26	0,998	0,995	0,098	0,000	-0,002	-0,003	0,005	-0,202	0,102	-0,096	0,196
Item 27	0,997	0,988	0,197	0,199	-0,197	0,397	-0,400	-0,198	0,199	-0,400	0,399
Item 28	1,011	1,995	0,001	0,002	0,000	-0,003	0,001	-0,007	0,005	0,010	-0,008
Item 29	1,001	1,998	0,099	-0,003	-0,002	0,003	0,002	0,004	0,003	-0,005	-0,002
Item 30	0,997	2,003	0,199	-0,004	0,007	0,007	-0,010	0,006	-0,011	-0,007	0,012
Item 31	2,040	-1,975	0,018	0,092	-0,208	-0,315	0,431	-0,003	-0,002	0,002	0,003
Item 32	2,034	-1,980	0,114	0,003	-0,001	-0,009	0,007	0,002	-0,006	0,001	0,003
Item 33	2,041	-1,977	0,212	-0,015	0,009	-0,001	0,007	-0,004	0,004	-0,001	0,001
Item 34	2,020	-0,993	0,004	-0,003	0,001	0,001	0,001	-0,001	0,000	0,001	0,000
Item 35	2,011	-0,999	0,101	0,004	0,000	-0,007	0,003	0,001	0,000	-0,002	0,002
Item 36	2,015	-1,000	0,199	0,004	-0,001	0,007	-0,011	-0,399	0,303	-0,101	0,197
Item 37	2,013	0,001	0,001	0,409	-0,201	0,095	-0,303	-0,401	0,201	-0,100	0,300
Item 38	2,008	0,000	0,100	-0,002	0,003	-0,006	0,005	0,498	-0,299	-0,099	-0,100
Item 39	2,008	0,000	0,200	-0,008	0,009	-0,003	0,002	-0,001	0,001	0,001	-0,001
Item 40	2,015	0,998	0,000	-0,001	0,000	0,000	0,001	-0,001	-0,001	0,002	0,000
Item 41	2,004	1,000	0,100	0,005	-0,003	-0,001	0,000	0,000	0,001	-0,002	0,001
Item 42	1,997	0,999	0,200	0,003	-0,001	0,008	-0,010	-0,003	0,000	0,001	0,002
Item 43	2,024	1,996	0,000	0,196	0,102	-0,203	-0,095	-0,199	-0,097	0,201	0,095
Item 44	2,017	1,996	0,100	-0,026	0,002	0,037	-0,013	-0,293	-0,202	0,594	-0,099
Item 45	2,006	2,001	0,200	0,004	-0,006	0,012	-0,010	0,306	0,200	-0,610	0,105

Fonte: Dados Simulados.

Tabela 5.3 *Desvios-padrão das estimativas de máxima verossimilhança dos parâmetros dos itens para os dados simulados.*

ITEM	a	b	c	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	0,021	0,159	0,044	0,021	0,025	0,034	0,027	0,141	0,093	0,105	0,082
Item 2	0,027	0,198	0,073	0,027	0,026	0,026	0,025	0,070	0,060	0,064	0,059
Item 3	0,032	0,243	0,080	0,029	0,029	0,030	0,028	0,070	0,080	0,077	0,062
Item 4	0,017	0,068	0,023	0,022	0,024	0,021	0,021	0,029	0,043	0,043	0,030
Item 5	0,028	0,186	0,044	0,024	0,028	0,024	0,029	0,227	0,096	0,070	0,088
Item 6	0,034	0,188	0,055	0,025	0,024	0,026	0,027	0,039	0,040	0,040	0,041
Item 7	0,016	0,038	0,011	0,020	0,020	0,019	0,019	0,026	0,024	0,027	0,027
Item 8	0,026	0,072	0,018	0,024	0,020	0,026	0,028	0,068	0,030	0,027	0,038
Item 9	0,034	0,119	0,031	0,025	0,024	0,027	0,024	0,033	0,030	0,034	0,032
Item 10	0,016	0,027	0,007	0,021	0,021	0,022	0,021	0,036	0,037	0,034	0,033
Item 11	0,034	0,061	0,020	0,031	0,027	0,028	0,028	0,046	0,037	0,051	0,039
Item 12	0,032	0,066	0,016	0,025	0,033	0,026	0,031	0,048	0,037	0,068	0,043
Item 13	0,019	0,029	0,004	0,024	0,024	0,023	0,027	0,061	0,058	0,061	0,067
Item 14	0,044	0,043	0,013	0,033	0,032	0,030	0,033	0,067	0,070	0,063	0,073
Item 15	0,047	0,052	0,011	0,038	0,039	0,056	0,035	0,115	0,071	0,075	0,085
Item 16	0,024	0,079	0,039	0,032	0,028	0,032	0,029	0,050	0,048	0,050	0,045
Item 17	0,029	0,132	0,064	0,034	0,030	0,032	0,032	0,056	0,051	0,057	0,057
Item 18	0,033	0,164	0,075	0,035	0,032	0,033	0,036	0,060	0,054	0,056	0,061
Item 19	0,017	0,035	0,015	0,020	0,023	0,021	0,023	0,026	0,026	0,022	0,025
Item 20	0,026	0,085	0,033	0,024	0,024	0,032	0,030	0,030	0,048	0,034	0,030
Item 21	0,035	0,104	0,039	0,026	0,027	0,031	0,029	0,031	0,034	0,034	0,031
Item 22	0,018	0,021	0,007	0,022	0,023	0,023	0,023	0,017	0,019	0,021	0,022
Item 23	0,026	0,042	0,016	0,031	0,027	0,028	0,025	0,022	0,024	0,024	0,020
Item 24	0,031	0,058	0,019	0,030	0,028	0,029	0,030	0,026	0,023	0,026	0,028
Item 25	0,019	0,017	0,003	0,026	0,025	0,024	0,025	0,029	0,026	0,024	0,028
Item 26	0,031	0,024	0,009	0,029	0,030	0,031	0,030	0,026	0,028	0,032	0,033
Item 27	0,037	0,036	0,011	0,038	0,034	0,045	0,032	0,030	0,042	0,027	0,059
Item 28	0,023	0,031	0,002	0,029	0,027	0,031	0,031	0,053	0,046	0,053	0,055
Item 29	0,040	0,030	0,006	0,042	0,041	0,036	0,041	0,062	0,059	0,054	0,062
Item 30	0,047	0,044	0,007	0,050	0,046	0,051	0,051	0,070	0,067	0,077	0,075
Item 31	0,033	0,028	0,024	0,063	0,053	0,051	0,065	0,032	0,036	0,032	0,028
Item 32	0,050	0,053	0,044	0,059	0,056	0,059	0,063	0,032	0,031	0,033	0,036
Item 33	0,058	0,063	0,047	0,055	0,060	0,065	0,076	0,037	0,036	0,035	0,037
Item 34	0,027	0,013	0,007	0,044	0,041	0,038	0,040	0,016	0,017	0,017	0,014
Item 35	0,032	0,021	0,014	0,050	0,047	0,043	0,039	0,018	0,019	0,018	0,016
Item 36	0,040	0,026	0,015	0,053	0,048	0,055	0,052	0,024	0,018	0,024	0,020
Item 37	0,023	0,007	0,002	0,036	0,034	0,038	0,027	0,011	0,012	0,011	0,013
Item 38	0,029	0,012	0,006	0,054	0,037	0,041	0,045	0,014	0,014	0,012	0,013
Item 39	0,039	0,016	0,008	0,045	0,046	0,048	0,045	0,016	0,014	0,013	0,013
Item 40	0,023	0,009	0,000	0,039	0,045	0,037	0,039	0,015	0,016	0,016	0,017
Item 41	0,039	0,011	0,004	0,055	0,058	0,061	0,058	0,016	0,015	0,015	0,018
Item 42	0,048	0,013	0,004	0,057	0,063	0,066	0,071	0,021	0,019	0,020	0,017
Item 43	0,033	0,019	0,000	0,057	0,053	0,060	0,054	0,028	0,028	0,038	0,033
Item 44	0,078	0,028	0,002	0,078	0,093	0,137	0,089	0,031	0,036	0,066	0,040
Item 45	0,080	0,030	0,003	0,153	0,121	0,104	0,132	0,054	0,047	0,038	0,045

Fonte: Dados Simulados.

Tabela 5.4 *Erros na estimação de máxima verossimilhança dos parâmetros dos itens para os dados simulados.*

ITEM	a	b	c	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	0,017	0,149	0,045	-0,007	-0,002	0,010	-0,001	0,090	0,056	-0,080	-0,066
Item 2	0,012	0,054	0,014	0,011	-0,007	-0,004	0,000	0,021	-0,018	-0,008	0,005
Item 3	0,008	0,023	0,003	-0,003	0,003	0,001	0,000	-0,006	0,010	0,005	-0,009
Item 4	0,011	0,050	0,017	0,000	0,001	-0,003	0,002	0,000	0,007	-0,008	0,001
Item 5	-0,002	-0,045	-0,011	0,000	0,005	-0,002	-0,003	-0,057	0,027	0,013	0,017
Item 6	0,003	-0,010	-0,003	-0,005	0,001	0,004	0,000	-0,003	0,003	0,001	-0,002
Item 7	0,008	0,022	0,008	0,003	-0,003	0,002	-0,003	0,002	-0,001	-0,001	0,000
Item 8	0,004	-0,001	0,000	0,002	-0,004	0,000	0,003	-0,001	-0,002	0,000	0,003
Item 9	-0,004	-0,032	-0,008	-0,001	-0,002	0,002	0,001	-0,003	0,001	0,001	0,002
Item 10	0,011	0,013	0,005	-0,002	0,003	-0,002	0,000	0,003	-0,009	0,007	-0,001
Item 11	-0,005	-0,003	-0,003	0,006	0,001	-0,004	-0,002	-0,008	-0,002	0,013	-0,002
Item 12	0,005	-0,008	-0,001	-0,001	0,003	-0,004	0,002	0,002	0,004	-0,003	-0,002
Item 13	0,012	-0,003	0,002	0,001	0,001	0,000	-0,002	-0,002	-0,003	-0,001	0,006
Item 14	0,001	0,001	-0,001	0,000	0,002	0,001	-0,003	0,002	0,003	-0,013	0,008
Item 15	0,001	0,002	-0,001	-0,001	-0,001	0,005	-0,002	0,002	0,007	-0,006	-0,002
Item 16	0,018	0,063	0,032	0,002	-0,001	0,001	-0,001	-0,002	0,006	0,002	-0,006
Item 17	0,005	0,013	0,006	0,001	-0,001	0,002	-0,003	0,007	-0,001	0,004	-0,009
Item 18	0,009	0,015	0,001	-0,001	0,003	0,000	-0,002	-0,004	0,006	-0,005	0,003
Item 19	0,009	0,022	0,010	0,002	-0,004	0,001	0,001	0,000	-0,001	0,000	0,000
Item 20	0,005	0,005	0,002	0,000	-0,005	0,002	0,002	0,001	0,001	-0,002	-0,001
Item 21	0,001	-0,007	-0,003	-0,004	0,003	0,002	0,000	0,000	0,000	-0,001	0,002
Item 22	0,010	0,011	0,005	0,001	-0,004	-0,001	0,004	0,000	0,000	0,002	-0,001
Item 23	-0,001	-0,005	-0,002	-0,001	0,000	0,001	0,000	-0,002	0,000	0,003	-0,001
Item 24	0,001	0,000	0,000	0,000	-0,001	-0,001	0,001	-0,003	0,003	0,003	-0,003
Item 25	0,009	0,001	0,002	0,001	-0,003	0,003	-0,001	-0,001	0,002	0,000	-0,001
Item 26	-0,002	-0,005	-0,002	0,000	-0,002	-0,003	0,005	-0,002	0,002	0,004	-0,004
Item 27	-0,003	-0,012	-0,003	-0,001	0,003	-0,003	0,000	0,002	-0,001	0,000	-0,001
Item 28	0,011	-0,005	0,001	0,002	0,000	-0,003	0,001	-0,007	0,005	0,010	-0,008
Item 29	0,001	-0,002	-0,001	-0,003	-0,002	0,003	0,002	0,004	0,003	-0,005	-0,002
Item 30	-0,003	0,003	-0,001	-0,004	0,007	0,007	-0,010	0,006	-0,011	-0,007	0,012
Item 31	0,040	0,025	0,018	-0,008	-0,008	-0,015	0,031	-0,003	-0,002	0,002	0,003
Item 32	0,034	0,020	0,014	0,003	-0,001	-0,009	0,007	0,002	-0,006	0,001	0,003
Item 33	0,041	0,023	0,012	-0,015	0,009	-0,001	0,007	-0,004	0,004	-0,001	0,001
Item 34	0,020	0,007	0,004	-0,003	0,001	0,001	0,001	-0,001	0,000	0,001	0,000
Item 35	0,011	0,001	0,001	0,004	0,000	-0,007	0,003	0,001	0,000	-0,002	0,002
Item 36	0,015	0,000	-0,001	0,004	-0,001	0,007	-0,011	0,001	0,003	-0,001	-0,003
Item 37	0,013	0,001	0,001	0,009	-0,001	-0,005	-0,003	-0,001	0,001	0,000	0,000
Item 38	0,008	0,000	0,000	-0,002	0,003	-0,006	0,005	-0,002	0,001	0,001	0,000
Item 39	0,008	0,000	0,000	-0,008	0,009	-0,003	0,002	-0,001	0,001	0,001	-0,001
Item 40	0,015	-0,002	0,000	-0,001	0,000	0,000	0,001	-0,001	-0,001	0,002	0,000
Item 41	0,004	0,000	0,000	0,005	-0,003	-0,001	0,000	0,000	0,001	-0,002	0,001
Item 42	-0,003	-0,001	0,000	0,003	-0,001	0,008	-0,010	-0,003	0,000	0,001	0,002
Item 43	0,024	-0,004	0,000	-0,004	0,002	-0,003	0,005	0,001	0,003	0,001	-0,005
Item 44	0,017	-0,004	0,000	-0,026	0,002	0,037	-0,013	0,007	-0,002	-0,006	0,001
Item 45	0,006	0,001	0,000	0,004	-0,006	0,012	-0,010	0,006	0,000	-0,010	0,005

Fonte: Dados Simulados.

5.1.2 Estimativas bayesianas

No processo de estimação bayesiana dos parâmetros para os dados simulados, as prioris utilizadas foram:

- $a_i \sim LN(m_a, \sigma_a^2)$, com $m_a = 1$ e $\sigma_a^2 = e^{0.5}$.
- $b_i \sim N(m_b, \sigma_b^2)$, com $m_b = 0$ e $\sigma_b^2 = 2$.
- $c_i \sim \text{Beta}(\alpha - 1, \beta - 1)$, com $\alpha = 5$ e $\beta = 17$.
- $\delta_{ik}^a \sim N(0, \sigma_{\delta^a}^2)$, com $\sigma_{\delta^a}^2 = 0.1$.
- $\delta_{ik}^b \sim N(0, \sigma_{\delta^b}^2)$, com $\sigma_{\delta^b}^2 = 0.1$.

As prioris acima especificadas são frequentemente adotadas na literatura (BAKER e KIM, 2004).

A Tabela 5.5 apresenta as estimativas médias dos parâmetros dos itens obtidas nas 100 réplicas, via metodologia bayesiana. Na Tabela 5.6 consta o desvio-padrão das estimativas bayesianas dos parâmetros dos itens. Aqui também observa-se que os desvios-padrão são pequenos.

As diferenças entre as estimativas médias e os valores (verdadeiros) adotados para os parâmetros dos itens, ou seja, os erros na estimação, são apresentados na Tabela 5.7. Observa-se que os erros foram pequenos, e a maior diferença em valor absoluto observada foi 0,304, associada a estimação do parâmetro b no item 1.

Assim, tem-se que o processo de estimação bayesiano também está funcionando bem, pois produziu estimativas precisas e pequenos erros de estimação.

Em ambos os processos de estimação, por máxima verossimilhança e bayesiano, o tempo total de execução do algoritmo para as 100 réplicas foi cerca de 15 horas.

Tabela 5.5 *Estimativas médias dos parâmetros dos itens para os dados simulados, via método bayesiano.*

ITEM	a	b	c	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	0,781	-1,696	0,092	-0,410	-0,305	0,508	0,208	0,190	0,096	-0,170	-0,117
Item 2	0,764	-1,907	0,133	-0,001	0,005	0,000	-0,004	-0,002	0,009	-0,002	-0,005
Item 3	0,742	-2,095	0,169	0,000	-0,004	0,004	0,000	-0,003	-0,007	0,010	0,000
Item 4	0,786	-0,834	0,059	0,009	0,000	-0,010	0,001	0,408	-0,098	-0,414	0,104
Item 5	0,755	-0,985	0,104	-0,505	0,303	-0,003	0,205	0,511	-0,305	-0,008	-0,197
Item 6	0,744	-1,057	0,184	0,002	-0,001	0,001	-0,002	0,003	0,003	-0,003	-0,003
Item 7	0,786	0,104	0,035	0,003	0,001	-0,001	-0,003	-0,004	0,001	-0,001	0,004
Item 8	0,756	0,018	0,105	-0,403	-0,101	0,101	0,404	0,002	0,006	0,000	-0,007
Item 9	0,744	-0,030	0,192	0,002	0,001	-0,005	0,002	-0,002	0,002	-0,002	0,001
Item 10	0,788	1,049	0,021	0,008	0,004	-0,007	-0,004	0,091	0,295	-0,290	-0,096
Item 11	0,756	1,010	0,103	0,001	-0,001	0,004	-0,004	0,208	-0,102	0,392	-0,499
Item 12	0,742	0,974	0,193	-0,101	0,294	-0,292	0,099	0,107	-0,294	0,280	-0,093
Item 13	0,799	1,998	0,013	-0,002	0,002	0,000	0,000	0,000	-0,003	0,002	0,001
Item 14	0,744	2,006	0,098	0,002	-0,004	0,000	0,002	-0,006	0,006	0,004	-0,004
Item 15	0,745	1,995	0,197	-0,297	0,090	0,309	-0,102	-0,006	0,009	-0,014	0,011
Item 16	1,034	-1,840	0,085	0,002	-0,004	0,002	0,000	-0,004	-0,003	0,007	-0,001
Item 17	1,016	-1,935	0,133	0,004	-0,001	0,003	-0,006	0,011	-0,002	-0,002	-0,008
Item 18	0,998	-2,029	0,187	0,003	0,005	-0,008	0,000	0,004	0,007	-0,013	0,002
Item 19	1,032	-0,905	0,043	-0,004	-0,005	0,010	-0,002	-0,003	-0,201	0,205	-0,001
Item 20	1,014	-0,971	0,110	-0,102	-0,302	0,302	0,101	0,001	0,011	-0,010	-0,003
Item 21	1,001	-1,012	0,195	0,000	-0,002	0,000	0,002	-0,002	-0,004	0,003	0,003
Item 22	1,033	0,053	0,022	-0,002	0,005	-0,004	0,001	-0,001	0,000	0,002	-0,001
Item 23	0,999	0,003	0,101	0,001	-0,002	-0,001	0,002	-0,001	0,001	0,001	0,000
Item 24	1,000	-0,006	0,199	-0,001	-0,005	0,006	0,000	-0,003	-0,002	-0,001	0,006
Item 25	1,038	1,020	0,011	-0,003	-0,001	0,002	0,002	0,006	0,000	-0,002	-0,004
Item 26	1,005	0,997	0,100	0,003	-0,005	-0,002	0,004	-0,200	0,105	-0,101	0,196
Item 27	1,002	0,991	0,199	0,203	-0,201	0,396	-0,397	-0,200	0,199	-0,395	0,395
Item 28	1,043	1,986	0,006	0,003	-0,004	0,003	-0,003	-0,007	0,005	0,000	0,002
Item 29	1,004	1,996	0,099	0,003	-0,004	-0,002	0,002	-0,007	0,002	0,003	0,001
Item 30	0,992	1,996	0,197	0,002	-0,001	0,002	-0,003	-0,003	0,009	-0,007	0,001
Item 31	2,061	-1,940	0,056	0,100	-0,214	-0,312	0,426	-0,002	0,006	0,005	-0,009
Item 32	2,032	-1,978	0,117	-0,009	0,003	-0,002	0,007	-0,001	0,001	-0,001	0,001
Item 33	2,030	-1,998	0,196	0,008	0,000	-0,005	-0,003	0,004	0,000	-0,004	0,000
Item 34	2,037	-0,976	0,018	0,000	0,003	0,001	-0,004	0,000	0,000	0,001	0,000
Item 35	2,015	-0,995	0,102	0,008	-0,002	-0,002	-0,004	0,001	-0,002	0,000	0,001
Item 36	2,006	-1,004	0,197	0,005	-0,002	-0,008	0,005	-0,400	0,301	-0,104	0,203
Item 37	2,029	0,010	0,006	0,404	-0,193	0,096	-0,307	-0,401	0,202	-0,101	0,300
Item 38	2,008	0,000	0,100	0,001	-0,003	0,002	0,000	0,500	-0,301	-0,098	-0,101
Item 39	2,002	-0,003	0,198	-0,004	0,001	-0,003	0,007	0,001	0,000	0,000	-0,001
Item 40	2,034	1,000	0,002	0,001	0,000	-0,002	0,000	0,004	-0,001	0,000	-0,002
Item 41	2,001	1,001	0,100	-0,001	0,001	0,004	-0,004	0,001	-0,001	0,000	-0,001
Item 42	2,003	0,999	0,200	0,002	-0,003	0,003	-0,002	-0,002	0,000	-0,001	0,003
Item 43	2,055	1,990	0,001	0,200	0,095	-0,195	-0,100	-0,198	-0,098	0,194	0,102
Item 44	2,002	2,003	0,099	-0,002	-0,005	0,015	-0,007	-0,304	-0,199	0,600	-0,096
Item 45	1,983	2,007	0,199	-0,016	0,006	0,019	-0,010	0,311	0,198	-0,608	0,099

Fonte: Dados Simulados.

Tabela 5.6 *Desvios-padrão das estimativas bayesianas dos parâmetros dos itens para os dados simulados.*

ITEM	a	b	c	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	0,018	0,085	0,023	0,023	0,022	0,032	0,029	0,095	0,079	0,059	0,059
Item 2	0,021	0,106	0,035	0,026	0,029	0,030	0,029	0,062	0,064	0,070	0,060
Item 3	0,019	0,118	0,039	0,028	0,026	0,029	0,026	0,071	0,068	0,083	0,072
Item 4	0,016	0,048	0,016	0,022	0,025	0,024	0,024	0,030	0,034	0,038	0,029
Item 5	0,018	0,107	0,025	0,019	0,031	0,024	0,027	0,141	0,060	0,051	0,063
Item 6	0,026	0,132	0,037	0,025	0,027	0,025	0,026	0,043	0,046	0,046	0,046
Item 7	0,017	0,030	0,009	0,021	0,023	0,020	0,023	0,027	0,023	0,028	0,023
Item 8	0,025	0,073	0,019	0,023	0,023	0,022	0,026	0,063	0,035	0,034	0,037
Item 9	0,032	0,104	0,028	0,026	0,023	0,024	0,025	0,033	0,033	0,033	0,033
Item 10	0,017	0,024	0,005	0,025	0,020	0,021	0,025	0,036	0,038	0,031	0,035
Item 11	0,030	0,046	0,014	0,030	0,025	0,029	0,028	0,045	0,035	0,047	0,035
Item 12	0,034	0,076	0,017	0,026	0,035	0,030	0,031	0,050	0,044	0,076	0,044
Item 13	0,021	0,035	0,004	0,024	0,029	0,025	0,029	0,057	0,064	0,055	0,060
Item 14	0,035	0,049	0,011	0,030	0,031	0,035	0,030	0,073	0,075	0,081	0,068
Item 15	0,044	0,056	0,010	0,037	0,036	0,051	0,033	0,099	0,075	0,076	0,081
Item 16	0,019	0,050	0,023	0,031	0,032	0,027	0,031	0,039	0,049	0,042	0,048
Item 17	0,023	0,084	0,039	0,032	0,030	0,030	0,028	0,049	0,044	0,049	0,049
Item 18	0,026	0,106	0,045	0,036	0,033	0,033	0,033	0,059	0,055	0,057	0,057
Item 19	0,017	0,030	0,011	0,023	0,026	0,026	0,023	0,024	0,031	0,023	0,027
Item 20	0,024	0,065	0,025	0,025	0,024	0,036	0,030	0,033	0,042	0,032	0,031
Item 21	0,027	0,080	0,030	0,031	0,027	0,032	0,029	0,033	0,034	0,036	0,032
Item 22	0,017	0,017	0,006	0,023	0,026	0,024	0,023	0,020	0,020	0,021	0,019
Item 23	0,025	0,038	0,014	0,029	0,028	0,029	0,027	0,023	0,019	0,020	0,019
Item 24	0,031	0,054	0,018	0,032	0,030	0,028	0,029	0,026	0,025	0,023	0,024
Item 25	0,018	0,017	0,003	0,028	0,027	0,027	0,026	0,029	0,027	0,027	0,024
Item 26	0,029	0,027	0,009	0,029	0,031	0,030	0,030	0,029	0,031	0,028	0,029
Item 27	0,035	0,041	0,011	0,040	0,032	0,040	0,033	0,034	0,042	0,027	0,058
Item 28	0,021	0,028	0,002	0,034	0,030	0,028	0,030	0,049	0,047	0,046	0,046
Item 29	0,040	0,034	0,005	0,038	0,039	0,041	0,040	0,052	0,051	0,057	0,055
Item 30	0,049	0,037	0,007	0,049	0,051	0,050	0,042	0,065	0,071	0,070	0,067
Item 31	0,036	0,022	0,015	0,064	0,051	0,043	0,067	0,031	0,031	0,029	0,028
Item 32	0,041	0,036	0,031	0,064	0,061	0,059	0,054	0,037	0,035	0,033	0,034
Item 33	0,050	0,049	0,036	0,065	0,070	0,070	0,075	0,038	0,042	0,037	0,043
Item 34	0,027	0,010	0,005	0,038	0,042	0,049	0,042	0,015	0,016	0,015	0,015
Item 35	0,043	0,024	0,015	0,047	0,038	0,045	0,041	0,016	0,016	0,017	0,017
Item 36	0,038	0,027	0,017	0,054	0,047	0,052	0,051	0,022	0,020	0,021	0,019
Item 37	0,022	0,007	0,002	0,042	0,033	0,039	0,037	0,010	0,011	0,012	0,013
Item 38	0,036	0,011	0,005	0,050	0,041	0,043	0,038	0,015	0,014	0,013	0,012
Item 39	0,036	0,014	0,006	0,050	0,052	0,052	0,051	0,015	0,014	0,015	0,013
Item 40	0,026	0,008	0,001	0,032	0,035	0,038	0,042	0,015	0,013	0,015	0,017
Item 41	0,039	0,011	0,003	0,058	0,069	0,055	0,054	0,017	0,016	0,017	0,017
Item 42	0,048	0,015	0,004	0,058	0,069	0,064	0,063	0,018	0,019	0,017	0,019
Item 43	0,033	0,018	0,000	0,057	0,058	0,061	0,060	0,025	0,027	0,037	0,035
Item 44	0,072	0,028	0,002	0,089	0,088	0,133	0,085	0,036	0,037	0,068	0,037
Item 45	0,089	0,030	0,003	0,142	0,133	0,108	0,125	0,062	0,049	0,034	0,054

Fonte: Dados Simulados.

Tabela 5.7 *Erros na estimação bayesiana dos parâmetros dos itens para os dados simulados.*

ITEM	a	b	c	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	0,031	0,304	0,092	-0,010	-0,005	0,008	0,008	0,190	0,096	-0,170	-0,117
Item 2	0,014	0,093	0,033	-0,001	0,005	0,000	-0,004	-0,002	0,009	-0,002	-0,005
Item 3	-0,008	-0,095	-0,031	0,000	-0,004	0,004	0,000	-0,003	-0,007	0,010	0,000
Item 4	0,036	0,166	0,059	0,009	0,000	-0,010	0,001	0,008	0,002	-0,014	0,004
Item 5	0,005	0,015	0,004	-0,005	0,003	-0,003	0,005	0,011	-0,005	-0,008	0,003
Item 6	-0,006	-0,057	-0,016	0,002	-0,001	0,001	-0,001	-0,002	0,003	-0,003	-0,003
Item 7	0,036	0,104	0,035	0,003	0,001	-0,001	-0,003	-0,004	0,001	-0,001	0,004
Item 8	0,006	0,018	0,005	-0,003	-0,001	0,001	0,004	0,002	0,006	0,000	-0,007
Item 9	-0,006	-0,030	-0,008	0,002	0,001	-0,004	0,002	-0,002	0,002	-0,002	0,001
Item 10	0,038	0,049	0,021	0,008	0,004	-0,007	-0,004	-0,009	-0,005	0,010	0,004
Item 11	0,006	0,010	0,003	0,001	-0,001	0,004	-0,004	0,008	-0,002	-0,008	0,001
Item 12	-0,008	-0,026	-0,007	-0,001	-0,006	0,008	-0,001	0,007	0,006	-0,020	0,007
Item 13	0,049	-0,002	0,013	-0,002	0,002	0,000	0,000	0,000	-0,003	0,002	0,001
Item 14	-0,006	0,006	-0,002	0,002	-0,004	0,000	0,002	-0,006	0,006	0,004	-0,004
Item 15	-0,005	-0,005	-0,003	0,003	-0,010	0,009	-0,002	-0,006	0,009	-0,014	0,011
Item 16	0,034	0,160	0,085	0,002	-0,004	0,002	0,000	-0,004	-0,003	0,007	-0,001
Item 17	0,016	0,065	0,033	0,004	-0,001	0,003	-0,006	0,011	-0,002	-0,002	-0,008
Item 18	-0,002	-0,029	-0,013	0,003	0,005	-0,007	0,000	0,004	0,007	-0,013	0,002
Item 19	0,032	0,095	0,043	-0,004	-0,005	0,010	-0,002	-0,003	-0,001	0,005	-0,001
Item 20	0,014	0,029	0,010	-0,002	-0,002	0,002	0,001	0,001	0,011	-0,010	-0,003
Item 21	0,001	-0,012	-0,005	0,000	-0,002	0,000	0,002	-0,002	-0,004	0,003	0,003
Item 22	0,033	0,053	0,022	-0,002	0,005	-0,004	0,001	-0,001	0,000	0,002	-0,001
Item 23	-0,001	0,003	0,001	0,001	-0,002	-0,001	0,002	-0,001	0,001	0,001	0,000
Item 24	0,000	-0,006	-0,001	-0,001	-0,005	0,006	0,000	-0,003	-0,002	-0,001	0,006
Item 25	0,038	0,020	0,011	-0,003	-0,001	0,002	0,002	0,006	0,000	-0,002	-0,004
Item 26	0,005	-0,003	0,000	0,003	-0,005	-0,002	0,004	0,000	0,005	-0,001	-0,004
Item 27	0,002	-0,009	-0,001	0,003	-0,001	-0,004	0,003	0,000	-0,001	0,005	-0,005
Item 28	0,043	-0,014	0,006	0,003	-0,004	0,003	-0,003	-0,007	0,005	0,000	0,002
Item 29	0,004	-0,004	-0,001	0,003	-0,004	-0,002	0,002	-0,007	0,002	0,003	0,001
Item 30	-0,008	-0,004	-0,003	0,002	-0,001	0,002	-0,003	-0,003	0,009	-0,007	0,001
Item 31	0,061	0,060	0,056	0,000	-0,014	-0,012	0,026	-0,002	0,006	0,005	-0,009
Item 32	0,032	0,022	0,017	-0,009	0,003	-0,002	0,007	-0,001	0,001	-0,001	0,001
Item 33	0,030	0,002	-0,004	0,008	0,000	-0,005	-0,003	0,004	0,000	-0,004	0,000
Item 34	0,037	0,024	0,018	0,000	0,003	0,001	-0,004	0,000	0,000	0,001	0,000
Item 35	0,015	0,005	0,002	0,008	-0,002	-0,002	-0,004	0,001	-0,002	0,000	0,001
Item 36	0,006	-0,004	-0,003	0,005	-0,002	-0,008	0,005	0,000	0,001	-0,004	0,003
Item 37	0,029	0,010	0,006	0,004	0,007	-0,004	-0,007	-0,001	0,002	-0,001	0,000
Item 38	0,008	0,000	0,000	0,001	-0,003	0,002	0,000	0,000	-0,001	0,002	-0,001
Item 39	0,002	-0,003	-0,002	-0,004	0,001	-0,003	0,007	0,001	0,000	0,000	-0,001
Item 40	0,034	0,000	0,002	0,001	0,000	-0,002	0,000	0,004	-0,001	0,000	-0,002
Item 41	0,001	0,001	0,000	-0,001	0,001	0,004	-0,004	0,001	-0,001	0,000	-0,001
Item 42	0,003	-0,001	0,000	0,002	-0,003	0,003	-0,002	-0,002	0,000	-0,001	0,003
Item 43	0,055	-0,010	0,001	0,000	-0,005	0,005	0,000	0,002	0,002	-0,006	0,002
Item 44	0,002	0,003	-0,001	-0,002	-0,005	0,015	-0,007	-0,004	0,001	0,000	0,004
Item 45	-0,017	0,007	-0,001	-0,016	0,006	0,019	-0,010	0,011	-0,002	-0,008	-0,001

Fonte: Dados Simulados.

5.2 Aplicação aos dados do ENEM-2014

Nesta seção serão apresentados os resultados do processo de estimação utilizando dados reais, o banco de dados utilizado foi o microdados do ENEM-2014, que foi extraído do site do INEP. Para a modelagem do EPI foi utilizado o processo de estimação em duas etapas, com estimativas obtidas via máxima verossimilhança, e o processo com modelagem do EPI em uma etapa, sendo as estimativas obtidas via máxima verossimilhança e via metodologia bayesiana.

5.2.1 Estrutura do exame e da base de dados

O ENEM é composto por quatro cadernos diferenciados pela cor (AZUL, AMARELO, BRANCO E ROSA). No exame, os candidatos são submetidos a provas objetivas referentes a 4 áreas do conhecimento, *Ciências da Natureza e suas tecnologias*, *Ciências Humanas e suas tecnologias*, *Linguagens, códigos e suas tecnologias* e *Matemática e suas tecnologias*. Em cada área o candidato deve responder 45 itens, totalizando 180 itens, além de elaborar uma Redação.

Segundo o INEP, o ENEM-2014 teve 8.722.248 candidatos inscritos. Para evitar um esforço computacional (com relação ao tempo de execução do algoritmo) muito grande na implementação do processo de estimação, os resultados foram obtidos apenas para os itens da prova de *Ciências da Natureza e suas tecnologias*, considerando uma amostra de candidatos. Para a base de dados do estudo foi feita a seleção de uma amostra entre os candidatos que participaram da prova de Ciências da Natureza (CN). Na Tabela 5.8 são apresentados o total (população) de candidatos que realizaram a prova de CN e o tamanho amostral selecionado, por caderno de prova.

Tabela 5.8 *População de candidatos e tamanho amostral por caderno de prova de Ciências da Natureza, ENEM-2014.*

Caderno	População	Amostra
Azul	1.606.456	10.000
Amarelo	1.519.145	10.000
Branco	1.559.347	10.000
Rosa	1.473.610	10.000
Total	6.158.558	40.000

Fonte: Microdados ENEM 2014/INEP.

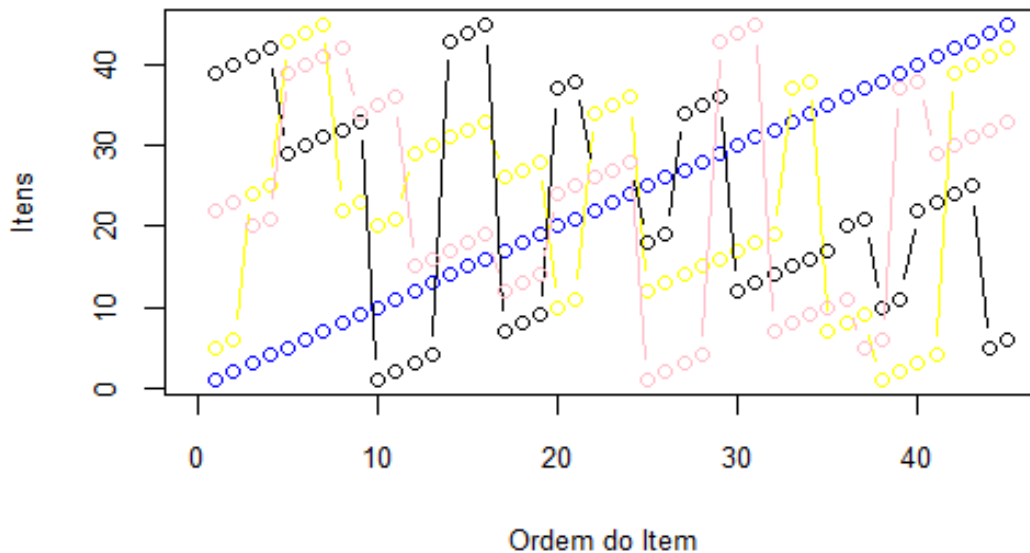
Cada caderno da prova de CN contém os mesmos 45 itens, porém dispostos em posições distintas (ver a Tabela 5.9). Observa-se na Figura 5.1 a posição do item nos cadernos com relação a ordem do item do caderno Azul, que será considerado o caderno de referência. Por exemplo, temos que o item 1 aparece na 1ª posição do caderno azul, este mesmo item aparece na 38ª posição do caderno amarelo, já no caderno branco na 10ª posição e no caderno rosa na 25ª posição.

Tabela 5.9 *Distribuição dos 45 itens da prova de Ciências da Natureza por caderno, ENEM-2014.*

Posição do item	Azul	Amarelo	Branco	Rosa
1ª	1	5	39	22
2ª	2	6	40	23
3ª	3	24	41	20
4ª	4	25	42	21
5ª	5	43	29	39
6ª	6	44	30	40
7ª	7	45	31	41
8ª	8	22	32	42
9ª	9	23	33	34
10ª	10	20	1	35
11ª	11	21	2	36
12ª	12	29	3	15
13ª	13	30	4	16
14ª	14	31	43	17
15ª	15	32	44	18
16ª	16	33	45	19
17ª	17	26	7	12
18ª	18	27	8	13
19ª	19	28	9	14
20ª	20	10	37	24
21ª	21	11	38	25
22ª	22	34	26	26
23ª	23	35	27	27
24ª	24	36	28	28
25ª	25	12	18	1
26ª	26	13	19	2
27ª	27	14	34	3
28ª	28	15	35	4
29ª	29	16	36	43
30ª	30	17	12	44
31ª	31	18	13	45
32ª	32	19	14	7
33ª	33	37	15	8
34ª	34	38	16	9
35ª	35	7	17	10
36ª	36	8	20	11
37ª	37	9	21	5
38ª	38	1	10	6
39ª	39	2	11	37
40ª	40	3	22	38
41ª	41	4	23	29
42ª	42	39	24	30
43ª	43	40	25	31
44ª	44	41	5	32
45ª	45	42	6	33

Fonte: Microdados ENEM 2014/INEP.

Figura 5.1 *Distribuição dos itens da prova de Ciências da Natureza nos cadernos azul, amarelo, branco e rosa, segundo a ordem destes itens, ENEM-2014.*



A leitura dos microdados do ENEM-2014 foi realizada pelo software SPSS e o processo de estimação foi todo implementado no software R. A fim de obter as estimativas iniciais dos parâmetros para o processo numérico de estimação, procedeu-se a estimação dos parâmetros a_i , b_i e c_i considerando conjuntamente as respostas de todos os cadernos e adotando o modelo ML3, ou seja, sem parâmetros associados ao EPI. Para os parâmetros δ_{ik}^a e δ_{ik}^b as estimativas iniciais foram assumidas iguais a zero.

Nessa etapa inicial de ajuste do ML3 para a obtenção de estimativas iniciais, os itens 6, 8, 17, 20 e 25 apresentaram problemas na convergência do processo de estimação, produzindo estimativas instáveis (que não convergiram). Uma avaliação gráfica do comportamento destes itens, através das suas curvas características empíricas (representação das proporções de acerto ao item versus as habilidades dos respondentes) apresentadas nas Figuras 5.2 a 5.6, comprovou um comportamento atípico. O que espera-se nessas curvas é um comportamento crescente, ou seja, que a proporção de acertos aumente com o aumento da habilidade, e que isto ocorra de forma semelhante nos 4 cadernos de prova.

Observa-se as curvas dos itens 6 e 8 não apresentam um comportamento crescente bem definido, já os itens 17, 20 e 25 parecem ser itens muito difíceis, pois a proporção de

acerto aumenta um pouco só para habilidades muito altas, e no caso do item 20 isso não ocorre no caderno rosa. Diante dessas análises, estes cinco itens foram retirados do estudo para não prejudicar a convergência no processo de estimação.

Figura 5.2 *Curva Característica Empírica do Item 6 por caderno de prova.*

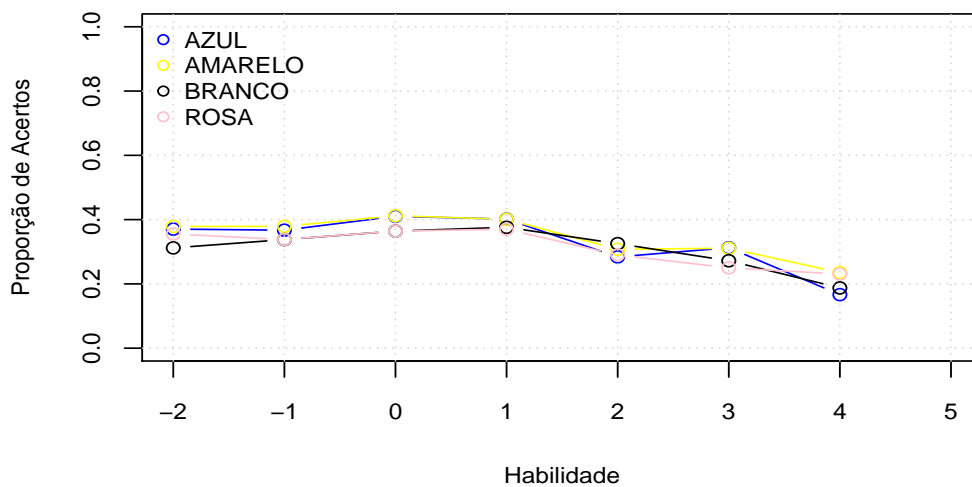


Figura 5.3 *Curva Característica Empírica do Item 8 por caderno de prova.*

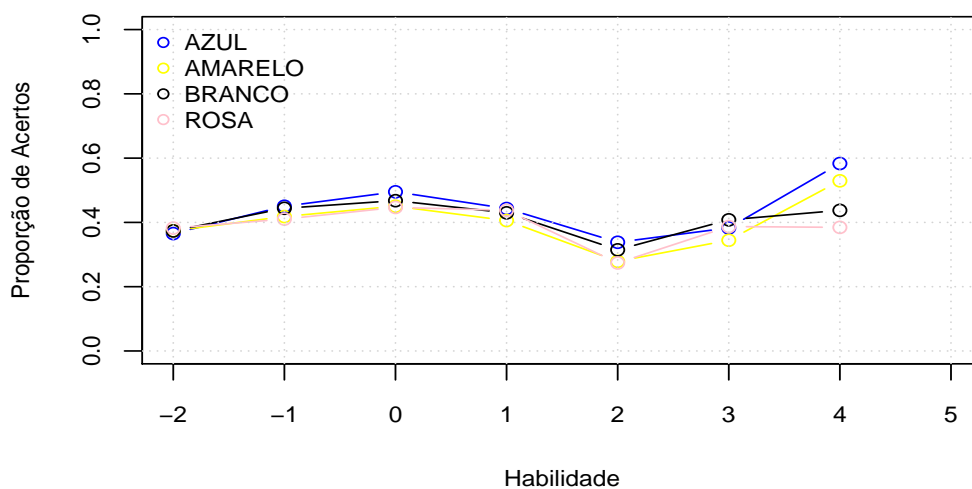


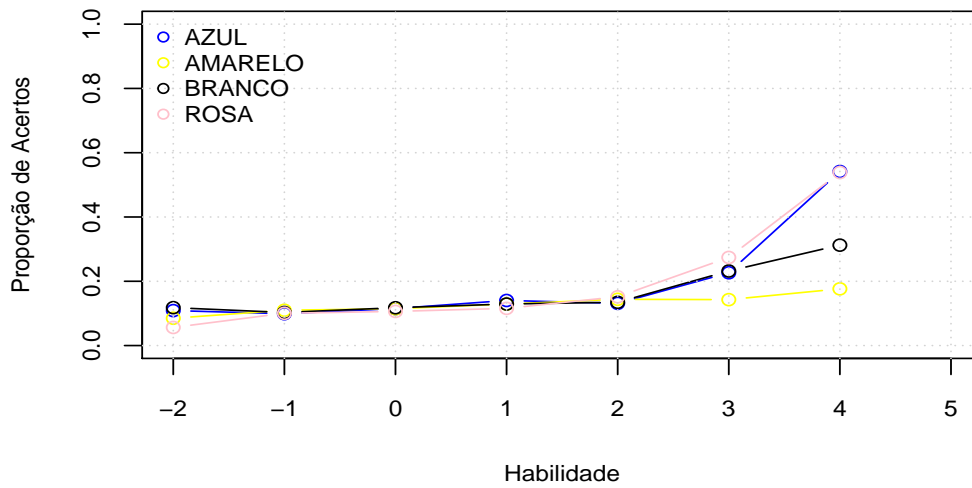
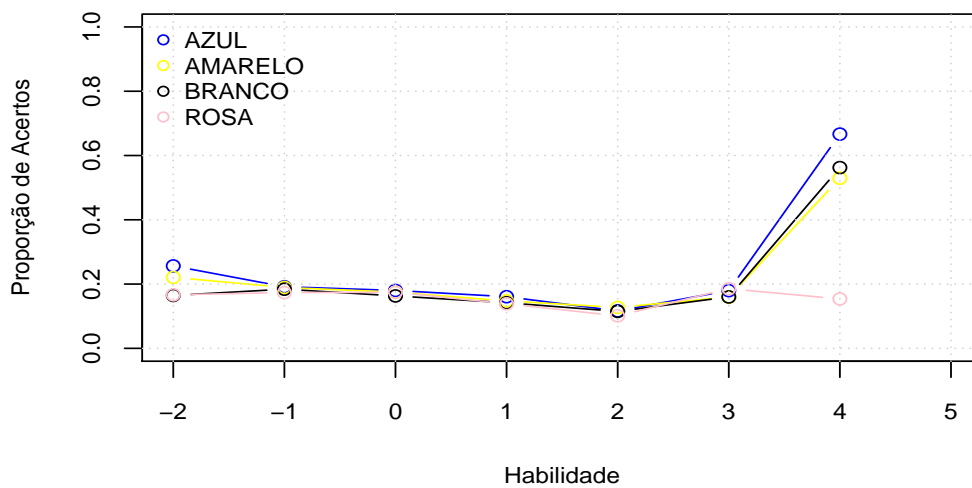
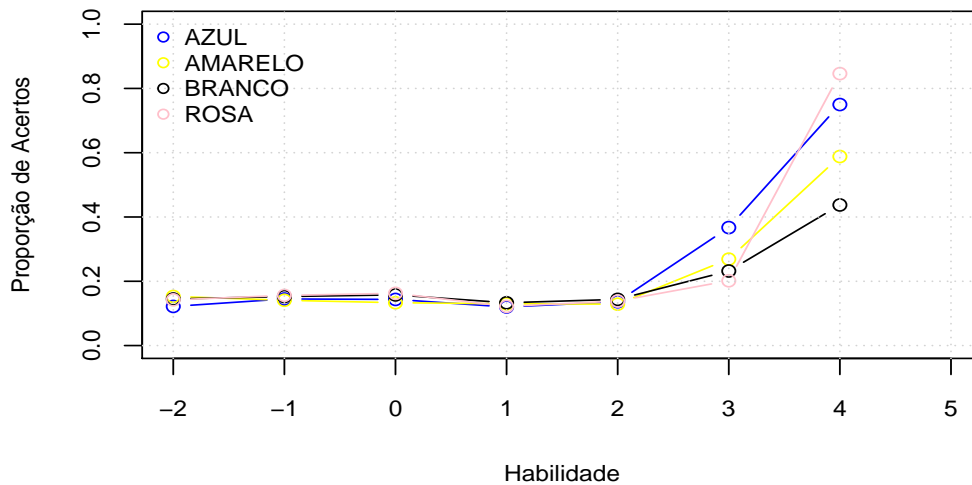
Figura 5.4 *Curva Característica Empírica do Item 17 por caderno de prova.*Figura 5.5 *Curva Característica Empírica do Item 20 por caderno de prova.*

Figura 5.6 *Curva Característica Empírica do Item 25 por caderno de prova.*

5.2.2 Distribuição dos Escores

As Figuras 5.7, 5.8, 5.9 e 5.10 correspondem aos histogramas dos escores (total de acertos) dos 10.000 respondentes dos cadernos de prova azul, amarelo, branco e rosa, respectivamente. Observa-se uma distribuição muito parecida entre os cadernos de prova, com uma leve assimetria à direita e predominância de 6 a 12 acertos entre os 40 itens.

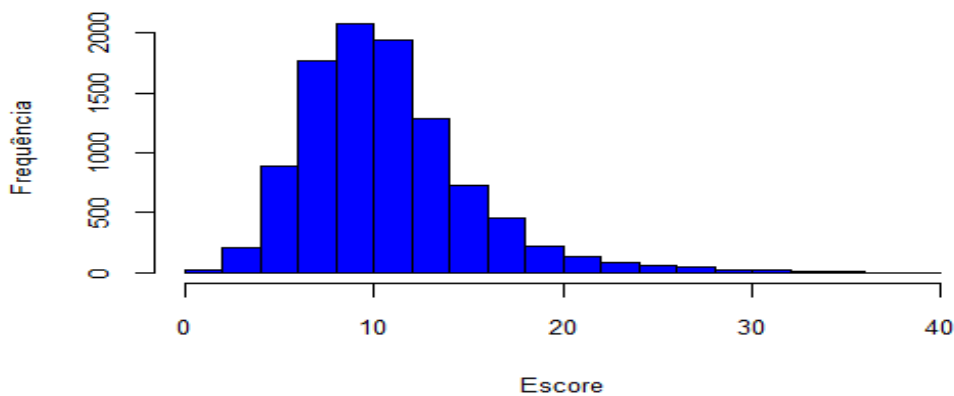
Figura 5.7 *Histograma do escore dos respondentes do caderno azul.*

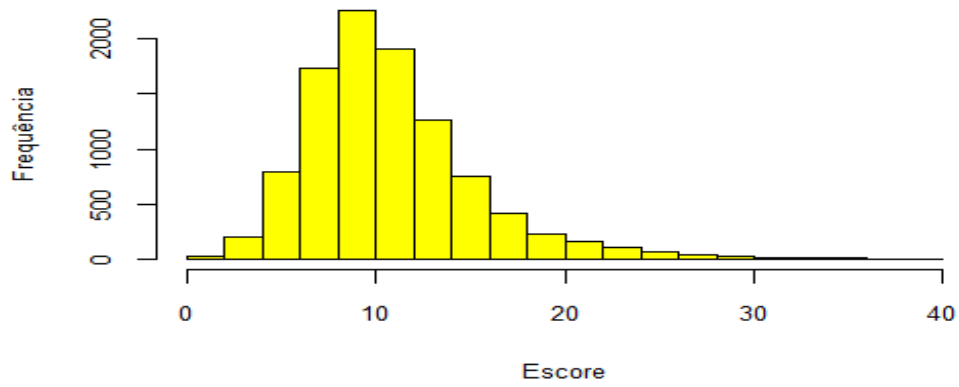
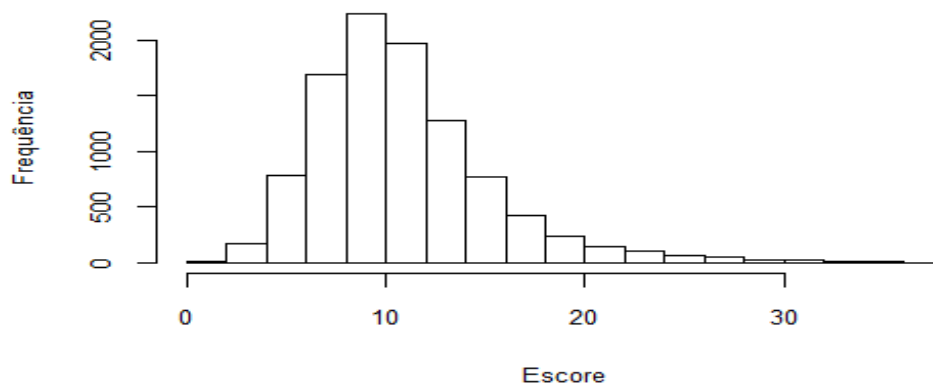
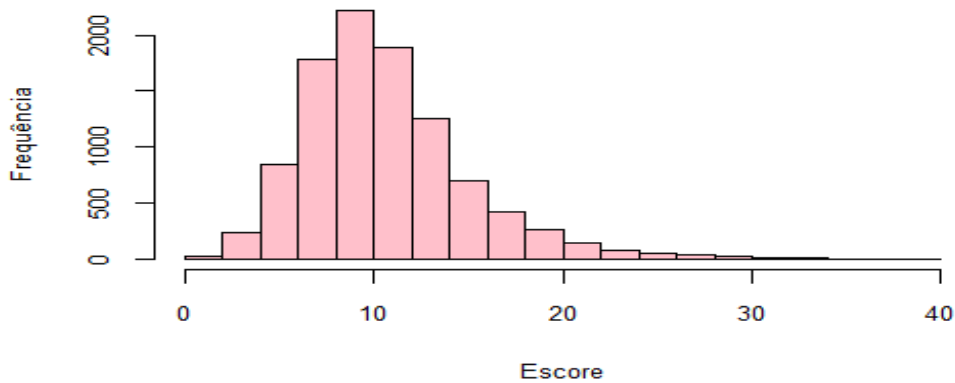
Figura 5.8 *Histograma do escore dos respondentes do caderno amarelo.*Figura 5.9 *Histograma do escore dos respondentes do caderno branco.*

Figura 5.10 *Histograma do escore dos respondentes do caderno rosa.*

5.2.3 Proporções de acertos

A Tabela 5.10 apresenta as proporções de acerto do item por caderno de prova. Nota-se que as proporções em cada caderno são muito semelhantes em todos os 40 itens. Por exemplo, no item 1 a proporção máxima foi 0,162 e a mínima 0,146, observa-se uma diferença pequena entre o máximo e o mínimo.

As Figuras 5.11 a 5.14 apresentam o gráfico de dispersão entre a proporção de acertos e a dificuldade do item nos cadernos azul, amarelo, branco e rosa, respectivamente. Observa-se nas 4 figuras que quando a dificuldade do item aumenta diminui a proporção de acertos ao item.

Tabela 5.10 *Proporção de acerto, seus valores mínimo e máximo e a diferença, por item e por caderno de prova, ENEM-2014.*

Item	Azul	Amarelo	Branco	Rosa	Mínimo	Máximo	Diferença
Item 1	0,150	0,162	0,146	0,159	0,146	0,162	0,016
Item 2	0,217	0,215	0,210	0,210	0,210	0,217	0,007
Item 3	0,196	0,201	0,207	0,197	0,196	0,207	0,011
Item 4	0,424	0,382	0,408	0,405	0,382	0,424	0,042
Item 5	0,275	0,282	0,259	0,258	0,258	0,282	0,024
Item 7	0,426	0,405	0,418	0,408	0,405	0,426	0,021
Item 9	0,291	0,234	0,251	0,245	0,234	0,291	0,057
Item 10	0,292	0,289	0,285	0,284	0,284	0,292	0,008
Item 11	0,282	0,278	0,254	0,254	0,254	0,282	0,029
Item 12	0,263	0,275	0,265	0,270	0,263	0,275	0,012
Item 13	0,213	0,214	0,209	0,212	0,209	0,214	0,005
Item 14	0,177	0,176	0,175	0,181	0,175	0,181	0,007
Item 15	0,316	0,318	0,320	0,343	0,316	0,343	0,028
Item 16	0,297	0,310	0,309	0,313	0,297	0,313	0,017
Item 18	0,215	0,209	0,215	0,222	0,209	0,222	0,014
Item 19	0,216	0,214	0,215	0,203	0,203	0,216	0,013
Item 21	0,174	0,162	0,169	0,163	0,162	0,174	0,012
Item 22	0,137	0,143	0,148	0,137	0,137	0,148	0,012
Item 23	0,161	0,160	0,168	0,148	0,148	0,168	0,020
Item 24	0,490	0,515	0,462	0,495	0,462	0,515	0,053
Item 26	0,289	0,300	0,295	0,285	0,285	0,300	0,016
Item 27	0,261	0,278	0,257	0,261	0,257	0,278	0,021
Item 28	0,189	0,211	0,206	0,195	0,189	0,211	0,022
Item 29	0,257	0,256	0,252	0,246	0,246	0,257	0,011
Item 30	0,386	0,422	0,449	0,376	0,376	0,449	0,073
Item 31	0,480	0,498	0,522	0,479	0,479	0,522	0,043
Item 32	0,248	0,231	0,244	0,241	0,231	0,248	0,016
Item 33	0,348	0,341	0,351	0,321	0,321	0,351	0,031
Item 34	0,203	0,205	0,215	0,191	0,191	0,215	0,024
Item 35	0,354	0,378	0,354	0,407	0,354	0,407	0,053
Item 36	0,320	0,318	0,326	0,299	0,299	0,326	0,028
Item 37	0,129	0,127	0,127	0,125	0,125	0,129	0,004
Item 38	0,246	0,250	0,247	0,240	0,240	0,250	0,010
Item 39	0,205	0,220	0,216	0,222	0,205	0,222	0,018
Item 40	0,561	0,550	0,569	0,561	0,550	0,569	0,019
Item 41	0,384	0,401	0,458	0,451	0,384	0,458	0,075
Item 42	0,397	0,402	0,412	0,402	0,397	0,412	0,015
Item 43	0,239	0,199	0,223	0,218	0,199	0,239	0,040
Item 44	0,151	0,135	0,140	0,144	0,135	0,151	0,016
Item 45	0,326	0,378	0,363	0,333	0,326	0,378	0,051

Fonte: Microdados ENEM 2014/INEP.

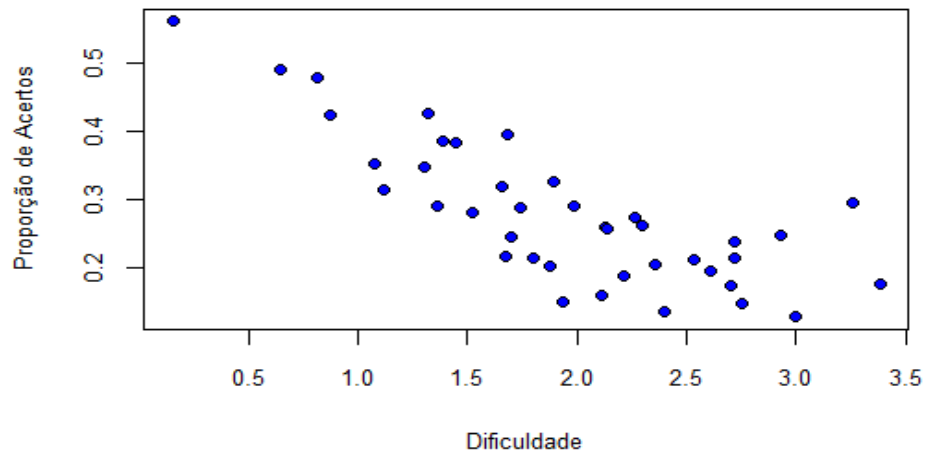
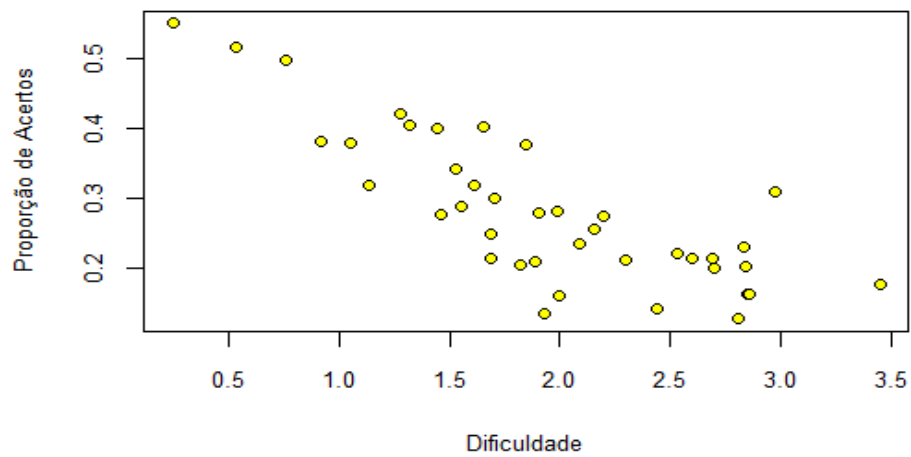
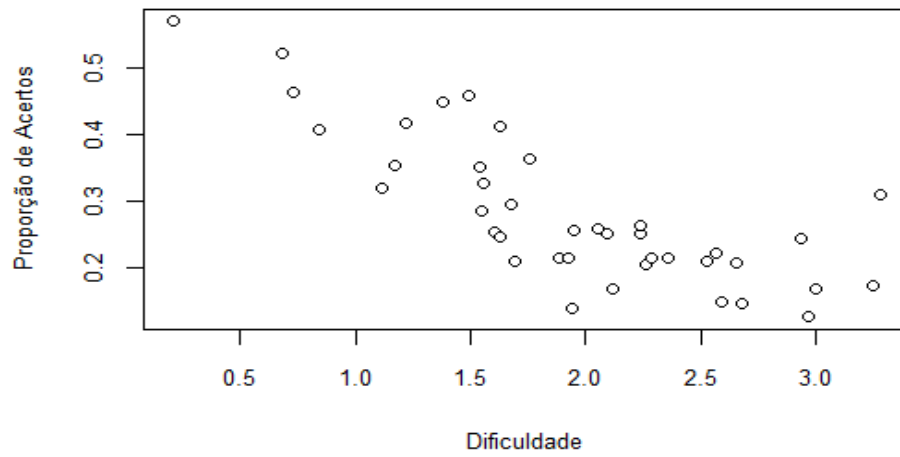
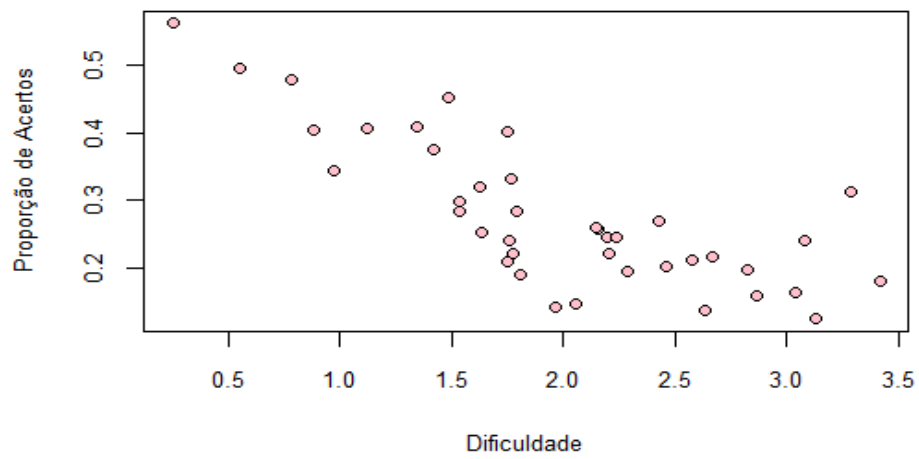
Figura 5.11 *Proporção de acertos segundo a dificuldade do item no caderno azul.*Figura 5.12 *Proporção de acertos segundo a dificuldade do item no caderno amarelo.*

Figura 5.13 *Proporção de acertos segundo a dificuldade do item no caderno branco.*Figura 5.14 *Proporção de acertos segundo a dificuldade do item no caderno rosa.*

5.2.4 A Modelagem do EPI em 2 Etapas

Nesta seção será apresentada a modelagem do EPI em duas etapas para a amostra selecionada dos 4 cadernos do ENEM-2014, considerando apenas 40 itens. O método utilizado na estimação dos parâmetros dos itens será o método da máxima verossimilhança.

5.2.4.1 Etapa 1: calibração dos itens

A Etapa 1 consistiu em realizar a calibração, estimação dos parâmetros a_i , b_i e c_i , para cada caderno de prova. Essas estimativas dos parâmetros de discriminação, dificuldade e acerto casual são apresentadas nas Tabelas 5.11, 5.12 e 5.13, respectivamente. Além das estimativas, as tabelas apresentam também a maior e a menor estimativa e a diferença entre esses valores. Nas 3 tabelas as diferenças são pequenas, sugerindo pouca variação das estimativas entre os cadernos de prova.

Posteriormente, foi feita uma segunda calibração utilizando conjuntamente as respostas de todos os cadernos de prova. Essas estimativas são apresentadas na Tabela 5.14.

Tabela 5.11 *Estimativas do parâmetro de discriminação a_i por caderno, seus valores máximos e mínimos e a diferença entre eles.*

ITEM	Estimativa do Parâmetro a_i				Máximo	Mínimo	Diferença
	Azul	Amarelo	Branco	Rosa			
Item 1	2,416	2,270	2,428	2,248	2,428	2,248	0,180
Item 2	1,750	1,830	1,756	1,650	1,830	1,650	0,180
Item 3	2,333	2,139	2,180	2,017	2,333	2,017	0,316
Item 4	1,485	1,396	1,468	1,503	1,503	1,396	0,107
Item 5	1,145	1,046	1,286	1,197	1,286	1,046	0,240
Item 7	1,796	1,262	1,354	1,297	1,796	1,262	0,534
Item 9	1,145	1,835	1,334	1,406	1,835	1,145	0,690
Item 10	2,431	2,218	2,595	2,239	2,595	2,218	0,377
Item 11	1,551	1,522	1,739	1,604	1,739	1,522	0,217
Item 12	1,676	1,593	1,804	1,252	1,804	1,252	0,552
Item 13	2,174	1,874	2,222	1,971	2,222	1,874	0,348
Item 14	1,800	1,439	1,607	1,619	1,800	1,439	0,361
Item 15	2,757	2,781	2,643	2,585	2,781	2,585	0,196
Item 16	1,477	1,715	1,445	1,450	1,715	1,445	0,270
Item 18	1,748	1,965	1,911	1,818	1,965	1,748	0,217
Item 19	1,599	1,662	1,811	1,923	1,923	1,599	0,324
Item 21	1,897	1,699	1,776	1,823	1,897	1,699	0,198
Item 22	2,778	2,479	2,462	2,354	2,778	2,354	0,424
Item 23	2,494	2,564	2,473	2,319	2,564	2,319	0,245
Item 24	1,731	1,496	1,766	1,501	1,766	1,496	0,270
Item 26	0,973	0,993	1,077	1,057	1,077	0,973	0,104
Item 27	1,292	1,510	1,656	1,335	1,656	1,292	0,364
Item 28	2,750	2,372	2,437	2,356	2,750	2,356	0,394
Item 29	1,232	1,156	1,343	1,243	1,343	1,156	0,187
Item 30	1,589	1,469	1,643	1,767	1,767	1,469	0,298
Item 31	0,992	0,926	0,812	0,951	0,992	0,812	0,180
Item 32	1,740	1,703	1,595	1,633	1,740	1,595	0,145
Item 33	1,096	1,095	1,000	1,155	1,155	1,000	0,155
Item 34	2,274	2,405	2,164	2,327	2,405	2,164	0,241
Item 35	1,229	1,255	1,262	1,245	1,262	1,229	0,033
Item 36	1,857	2,183	1,960	1,864	2,183	1,857	0,326
Item 37	1,925	1,960	1,842	1,814	1,960	1,814	0,146
Item 38	2,833	2,673	2,802	2,985	2,985	2,673	0,312
Item 39	1,446	1,432	1,438	1,245	1,446	1,245	0,201
Item 40	0,937	0,855	0,861	0,911	0,937	0,855	0,082
Item 41	1,432	1,280	1,615	1,937	1,937	1,280	0,657
Item 42	1,365	1,569	1,043	1,205	1,569	1,043	0,526
Item 43	1,910	1,917	1,963	2,048	2,048	1,910	0,138
Item 44	3,357	3,336	3,215	3,424	3,424	3,215	0,209
Item 45	2,718	2,718	2,718	2,718	2,718	2,718	0,000

Fonte: Microdados ENEM 2014/INEP.

Tabela 5.12 *Estimativas do parâmetro de dificuldade b_i por caderno, seus valores máximos e mínimos e a diferença entre eles.*

ITEM	Estimativa do Parâmetro b_i				Máximo	Mínimo	Diferença
	Azul	Amarelo	Branco	Rosa			
Item 1	2,721	2,814	2,638	2,838	2,838	2,638	0,200
Item 2	1,659	1,674	1,681	1,736	1,736	1,659	0,077
Item 3	2,577	2,812	2,619	2,795	2,812	2,577	0,235
Item 4	0,880	0,923	0,841	0,880	0,923	0,841	0,083
Item 5	2,239	1,979	2,026	2,132	2,239	1,979	0,260
Item 7	1,315	1,313	1,228	1,339	1,339	1,228	0,111
Item 9	1,962	2,069	2,207	2,170	2,207	1,962	0,244
Item 10	1,352	1,543	1,535	1,528	1,543	1,352	0,191
Item 11	1,512	1,453	1,593	1,616	1,616	1,453	0,163
Item 12	2,270	2,178	2,213	2,401	2,401	2,178	0,223
Item 13	2,499	2,669	2,507	2,547	2,669	2,499	0,171
Item 14	3,341	3,399	3,219	3,381	3,399	3,219	0,180
Item 15	1,115	1,136	1,110	0,968	1,136	0,968	0,168
Item 16	3,223	2,943	3,248	3,268	3,268	2,943	0,324
Item 18	1,785	1,869	1,865	1,762	1,869	1,762	0,107
Item 19	2,684	2,567	2,331	2,430	2,684	2,331	0,353
Item 21	2,669	2,822	2,953	3,009	3,009	2,669	0,340
Item 22	2,360	2,411	2,561	2,609	2,609	2,360	0,249
Item 23	2,085	1,972	2,096	2,039	2,096	1,972	0,125
Item 24	0,638	0,544	0,721	0,556	0,721	0,544	0,177
Item 26	1,726	1,692	1,658	1,779	1,779	1,658	0,120
Item 27	2,110	1,887	1,929	2,125	2,125	1,887	0,238
Item 28	2,187	2,280	2,237	2,264	2,280	2,187	0,093
Item 29	2,116	2,138	2,071	2,221	2,221	2,071	0,150
Item 30	1,377	1,268	1,366	1,413	1,413	1,268	0,145
Item 31	0,826	0,760	0,682	0,775	0,826	0,682	0,143
Item 32	2,887	2,793	2,902	3,051	3,051	2,793	0,258
Item 33	1,297	1,517	1,546	1,613	1,613	1,297	0,316
Item 34	1,858	1,805	1,908	1,788	1,908	1,788	0,120
Item 35	1,068	1,057	1,156	1,108	1,156	1,057	0,100
Item 36	1,640	1,600	1,541	1,523	1,640	1,523	0,117
Item 37	2,959	2,771	2,925	3,104	3,104	2,771	0,332
Item 38	1,679	1,676	1,614	1,743	1,743	1,614	0,129
Item 39	2,329	2,494	2,257	2,183	2,494	2,183	0,311
Item 40	0,152	0,250	0,204	0,247	0,250	0,152	0,099
Item 41	1,429	1,423	1,481	1,470	1,481	1,423	0,058
Item 42	1,678	1,638	1,620	1,733	1,733	1,620	0,112
Item 43	2,680	2,674	2,533	2,625	2,680	2,533	0,147
Item 44	1,913	1,914	1,920	1,945	1,945	1,913	0,032
Item 45	1,821	1,799	1,722	1,699	1,821	1,699	0,122

Fonte: Microdados ENEM 2014/INEP.

Tabela 5.13 *Estimativas do parâmetro de acerto casual c_i por caderno, seus valores máximos e mínimos e a diferença entre eles.*

ITEM	Estimativa do Parâmetro c_i				Máximo	Mínimo	Diferença
	Azul	Amarelo	Branco	Rosa			
Item 1	0,140	0,153	0,135	0,150	0,153	0,135	0,018
Item 2	0,120	0,124	0,115	0,115	0,124	0,115	0,009
Item 3	0,182	0,191	0,192	0,184	0,192	0,182	0,010
Item 4	0,208	0,154	0,175	0,182	0,208	0,154	0,054
Item 5	0,192	0,161	0,170	0,169	0,192	0,161	0,031
Item 7	0,313	0,246	0,258	0,258	0,313	0,246	0,067
Item 9	0,183	0,184	0,181	0,178	0,184	0,178	0,006
Item 10	0,190	0,206	0,212	0,200	0,212	0,190	0,022
Item 11	0,158	0,141	0,152	0,145	0,158	0,141	0,017
Item 12	0,221	0,223	0,224	0,209	0,224	0,209	0,015
Item 13	0,195	0,195	0,192	0,191	0,195	0,191	0,004
Item 14	0,170	0,163	0,162	0,172	0,172	0,162	0,010
Item 15	0,184	0,191	0,184	0,178	0,191	0,178	0,013
Item 16	0,282	0,295	0,294	0,298	0,298	0,282	0,016
Item 18	0,133	0,146	0,149	0,142	0,149	0,133	0,016
Item 19	0,188	0,184	0,180	0,176	0,188	0,176	0,012
Item 21	0,155	0,142	0,155	0,151	0,155	0,142	0,013
Item 22	0,122	0,126	0,136	0,124	0,136	0,122	0,014
Item 23	0,129	0,121	0,136	0,109	0,136	0,109	0,027
Item 24	0,253	0,250	0,238	0,223	0,253	0,223	0,030
Item 26	0,126	0,138	0,141	0,142	0,142	0,126	0,016
Item 27	0,180	0,197	0,188	0,186	0,197	0,180	0,017
Item 28	0,166	0,188	0,181	0,170	0,188	0,166	0,022
Item 29	0,171	0,163	0,172	0,169	0,172	0,163	0,009
Item 30	0,263	0,279	0,338	0,266	0,338	0,263	0,075
Item 31	0,223	0,228	0,233	0,206	0,233	0,206	0,027
Item 32	0,231	0,212	0,224	0,226	0,231	0,212	0,019
Item 33	0,152	0,179	0,180	0,176	0,180	0,152	0,028
Item 34	0,150	0,151	0,163	0,132	0,163	0,132	0,031
Item 35	0,132	0,164	0,154	0,210	0,210	0,132	0,078
Item 36	0,238	0,243	0,237	0,199	0,243	0,199	0,044
Item 37	0,118	0,113	0,114	0,116	0,118	0,113	0,005
Item 38	0,191	0,192	0,185	0,193	0,193	0,185	0,008
Item 39	0,150	0,176	0,157	0,141	0,176	0,141	0,035
Item 40	0,181	0,183	0,205	0,205	0,205	0,181	0,024
Item 41	0,255	0,260	0,362	0,368	0,368	0,255	0,113
Item 42	0,296	0,312	0,272	0,292	0,312	0,272	0,040
Item 43	0,221	0,181	0,202	0,201	0,221	0,181	0,040
Item 44	0,120	0,103	0,107	0,115	0,120	0,103	0,017
Item 45	0,286	0,337	0,316	0,282	0,337	0,282	0,055

Fonte: Microdados ENEM 2014/INEP.

Tabela 5.14 *Estimativas dos Parâmetros a_i , b_i e c_i , considerando os dados conjuntamente.*

ITEM	a_i	b_i	c_i
Item 1	3,089	2,670	0,146
Item 2	1,564	1,756	0,108
Item 3	2,840	2,643	0,191
Item 4	1,299	0,870	0,155
Item 5	0,937	2,242	0,151
Item 7	1,312	1,370	0,268
Item 9	1,234	2,251	0,177
Item 10	2,381	1,556	0,206
Item 11	1,424	1,606	0,137
Item 12	1,493	2,388	0,221
Item 13	2,621	2,533	0,198
Item 14	1,979	3,301	0,171
Item 15	2,670	1,133	0,187
Item 16	1,657	3,305	0,297
Item 18	1,738	1,902	0,139
Item 19	1,770	2,583	0,184
Item 21	2,002	2,881	0,154
Item 22	3,131	2,462	0,129
Item 23	2,644	2,096	0,125
Item 24	1,585	0,678	0,250
Item 26	0,817	1,716	0,089
Item 27	1,285	2,125	0,181
Item 28	2,976	2,257	0,180
Item 29	1,000	2,289	0,148
Item 30	1,572	1,438	0,292
Item 31	0,782	0,675	0,180
Item 32	1,911	2,936	0,228
Item 33	0,872	1,481	0,127
Item 34	2,328	1,905	0,151
Item 35	1,051	1,053	0,123
Item 36	1,930	1,650	0,233
Item 37	2,189	2,915	0,117
Item 38	3,159	1,740	0,196
Item 39	1,190	2,480	0,148
Item 40	0,778	0,057	0,140
Item 41	1,480	1,543	0,317
Item 42	1,163	1,784	0,293
Item 43	2,407	2,615	0,206
Item 44	3,951	1,962	0,112
Item 45	2,718	1,864	0,310

Fonte: Microdados ENEM 2014/INEP.

5.2.4.2 Etapa 2: teste de hipóteses para a significância do EPI

A Etapa 2 consistiu em obter para cada item as diferenças entre a estimativa do parâmetro a_i obtida no caderno azul e a estimativa obtida com o uso conjunto das respostas, ou seja, as estimativas para δ_1^a . Essas diferenças também foram obtidas para as estimativas do parâmetro b_i , obtendo-se as estimativas para δ_1^b .

O mesmo foi feito para o caderno amarelo, obtendo-se estimativas para δ_2^a e δ_2^b , para o caderno branco, obtendo-se estimativas para δ_3^a e δ_3^b e para o caderno rosa, obtendo-se estimativas para δ_4^a e δ_4^b . Esses resultados são apresentados na Tabela 5.15.

Tabela 5.15 *Estimativas dos parâmetros de EPI no processo de estimação em 2 Etapas.*

ITEM	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	-0,673	-0,819	-0,661	-0,841	0,051	0,144	-0,032	0,167
Item 2	0,186	0,266	0,192	0,086	-0,097	-0,082	-0,075	-0,020
Item 3	-0,507	-0,701	-0,660	-0,823	-0,066	0,168	-0,024	0,152
Item 4	0,186	0,097	0,169	0,204	0,010	0,053	-0,029	0,010
Item 5	0,208	0,109	0,349	0,260	-0,003	-0,263	-0,217	-0,110
Item 7	0,484	-0,050	0,042	-0,015	-0,056	-0,057	-0,142	-0,031
Item 9	-0,089	0,601	0,100	0,172	-0,289	-0,182	-0,044	-0,081
Item 10	0,050	-0,163	0,214	-0,142	-0,204	-0,013	-0,021	-0,028
Item 11	0,127	0,098	0,315	0,180	-0,094	-0,153	-0,013	0,010
Item 12	0,183	0,100	0,311	-0,241	-0,118	-0,210	-0,174	0,013
Item 13	-0,447	-0,747	-0,399	-0,650	-0,035	0,136	-0,026	0,014
Item 14	-0,179	-0,540	-0,372	-0,360	0,040	0,098	-0,082	0,080
Item 15	0,087	0,111	-0,027	-0,085	-0,018	0,002	-0,023	-0,166
Item 16	-0,180	0,058	-0,212	-0,207	-0,082	-0,361	-0,057	-0,037
Item 18	0,010	0,227	0,173	0,080	-0,117	-0,033	-0,037	-0,140
Item 19	-0,171	-0,108	0,041	0,153	0,101	-0,016	-0,252	-0,153
Item 21	-0,105	-0,303	-0,226	-0,179	-0,212	-0,059	0,072	0,128
Item 22	-0,353	-0,652	-0,669	-0,777	-0,102	-0,050	0,099	0,147
Item 23	-0,150	-0,080	-0,171	-0,325	-0,011	-0,125	0,000	-0,058
Item 24	0,146	-0,089	0,181	-0,084	-0,040	-0,134	0,043	-0,122
Item 26	0,156	0,176	0,260	0,240	0,010	-0,024	-0,058	0,063
Item 27	0,007	0,225	0,371	0,050	-0,015	-0,237	-0,195	0,001
Item 28	-0,226	-0,604	-0,539	-0,620	-0,070	0,023	-0,021	0,007
Item 29	0,232	0,156	0,343	0,243	-0,173	-0,151	-0,218	-0,068
Item 30	0,017	-0,103	0,071	0,195	-0,061	-0,169	-0,071	-0,025
Item 31	0,210	0,144	0,030	0,169	0,150	0,085	0,007	0,100
Item 32	-0,171	-0,208	-0,316	-0,278	-0,049	-0,143	-0,034	0,115
Item 33	0,224	0,223	0,128	0,283	-0,184	0,036	0,065	0,132
Item 34	-0,054	0,077	-0,164	-0,001	-0,047	-0,100	0,003	-0,117
Item 35	0,178	0,204	0,211	0,194	0,015	0,003	0,103	0,055
Item 36	-0,073	0,253	0,030	-0,066	-0,010	-0,051	-0,109	-0,127
Item 37	-0,264	-0,229	-0,347	-0,375	0,045	-0,143	0,010	0,189
Item 38	-0,326	-0,486	-0,357	-0,174	-0,061	-0,064	-0,126	0,003
Item 39	0,256	0,242	0,248	0,055	-0,151	0,014	-0,223	-0,297
Item 40	0,159	0,077	0,083	0,133	0,095	0,194	0,148	0,190
Item 41	-0,048	-0,200	0,135	0,457	-0,113	-0,119	-0,061	-0,072
Item 42	0,202	0,406	-0,120	0,042	-0,107	-0,146	-0,164	-0,051
Item 43	-0,497	-0,490	-0,444	-0,359	0,065	0,059	-0,082	0,010
Item 44	-0,594	-0,615	-0,736	-0,527	-0,049	-0,048	-0,042	-0,017
Item 45	0,000	0,000	0,000	0,000	-0,043	-0,065	-0,142	-0,164

Fonte: Microdados ENEM 2014/INEP.

Como o objetivo principal não é a estimação em si, mas sim fazer afirmações sobre a significância ou não das estimativas obtidas, foram realizados testes de hipóteses para testar a hipótese nula de que não existe EPI contra a hipótese alternativa de que existe EPI. Esses testes ocorreram individualmente para cada item em cada caderno, verificando

a significância ou não do EPI no parâmetro de discriminação e de dificuldade com as hipóteses expressas, respectivamente, por

$$H_0 : \delta_{it}^a = 0 \quad \text{versus} \quad H_1 : \delta_{it}^a \neq 0, \quad (5.1)$$

e

$$H_0 : \delta_{it}^b = 0 \quad \text{versus} \quad H_1 : \delta_{it}^b \neq 0, \quad (5.2)$$

para $i = 1, 2, \dots, 40$ e $t = 1, 2, 3$ e 4 .

A decisão sobre a rejeição ou não das hipótese nulas acima baseou-se no Intervalo de Confiança (IC). Foram construídos intervalos assintóticos com 95% de confiança para os parâmetros δ_{i1}^a , δ_{i2}^a , δ_{i3}^a , δ_{i4}^a , δ_{i1}^b , δ_{i2}^b , δ_{i3}^b e δ_{i4}^b , e verificado se o valor zero estava ou não contido nos intervalos, indicando a não-rejeição ou a rejeição, respectivamente, da hipótese H_0 ao nível de significância de 5%. Assim, para cada item foram construídos 8 intervalos, que são apresentados nas Tabelas A.1 e A.2 no Apêndice A.

Os itens que apresentaram EPI significativo, por caderno de prova e por parâmetro, são apresentados na Tabela 5.16. Observa-se que entre os 40 itens da prova de Ciências da Natureza do ENEM-2014 considerados no estudo, não apresentaram EPI os seguintes itens: 1, 3, 12, 13, 14, 16, 19, 21, 22, 28, 32, 37, 43 e 44.

Dentre os itens que apresentaram EPI apenas no parâmetro de discriminação (a), os itens 26 e 35 apresentaram efeito em todos os cadernos, o item 33 apresentou nos cadernos azul, amarelo e rosa, e os itens 31 e 40 apresentaram efeito nos cadernos azul e rosa. Em relação aos itens que apresentaram EPI apenas no parâmetro de dificuldade (b), apenas o item 24 apresentou efeito em mais de um caderno (amarelo e rosa). Os itens 7, 9, 11 e 27 apresentaram EPI significativo em ambos os parâmetros.

Tabela 5.16 *Itens com EPI nos parâmetros de discriminação e de dificuldade por caderno na estimação em duas etapas.*

Caderno	Itens com EPI no parâmetro a	Itens com EPI no parâmetro b
Azul	7, 26, 31, 33, 35 e 40	9 e 10
Amarelo	2, 9, 26, 33, 35 e 42	11, 23, 24, 27 e 30
Branco	5, 11, 26, 27, 29 e 35	7 e 38
Rosa	4, 5, 26, 31, 33, 35, 40, 41 e 45	15, 18, 24, 34, 36 e 39

Fonte: Microdados ENEM 2014/INEP.

Na próxima seção, os dados do ENEM-2014 serão analisados com a modelagem do EPI em uma etapa, ou seja, com a incorporação de parâmetros associados ao EPI no ML3.

5.2.5 A Modelagem do EPI em uma Etapa

Nesta seção serão apresentados os resultados obtidos no processo de estimação via método da máxima (EMV) verossimilhança e via estimação bayesiana (EB), considerando a modelagem do EPI em uma etapa, para a amostra de dados do ENEM-2014 discutida na Seção 5.2.4. O modelo em (3.5) será utilizado em toda esta seção para o ajuste dos dados.

5.2.5.1 Estimativas de máxima verossimilhança e bayesianas

Aqui também foi utilizado o software R para a implementação do processo de estimação, com a aplicação do algoritmo EM na solução numérica das equações de estimação de máxima verossimilhança e bayesianas.

De forma similar ao já discutido anteriormente, as estimativas iniciais adotadas em ambos os processos, EMV e EB, para os parâmetros a_i , b_i e c_i foram as estimativas obtidas considerando as respostas de todos os cadernos conjuntamente, e para os parâmetros δ_{ik}^a e δ_{ik}^b adotou-se estimativas iniciais iguais a zero. No processo de estimação bayesiana foram utilizadas as prioris já apresentadas na Seção 5.1.2.

As Tabelas (5.17) e (5.18) apresentam as estimativas dos parâmetros dos itens, via estimação por máxima verossimilhança e via estimação bayesiana, respectivamente. Pode-se observar que as estimativas obtidas nos dois processos de estimação estão muito próximas. Isso se deve ao fato da função de verossimilhança, por ser proveniente de uma amostra muito grande, influenciar a distribuição posterior mais do que as prioris.

Tabela 5.17 *Estimativas de máxima verossimilhança dos parâmetros dos itens na modelagem do EPI em uma etapa.*

ITEM	a_i	b_i	c_i	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	2,885	3,117	0,146	-0,094	-0,048	0,060	0,081	0,035	0,000	-0,057	0,022
Item 2	1,225	1,920	0,090	-0,005	0,021	0,022	-0,038	-0,035	-0,031	0,001	0,065
Item 3	2,646	3,095	0,192	-0,102	0,378	-0,192	-0,084	-0,030	0,023	-0,056	0,064
Item 4	1,093	0,409	0,000	-0,020	0,011	0,016	-0,007	-0,087	0,106	-0,022	0,002
Item 5	0,681	2,713	0,132	-0,096	0,006	0,062	0,027	0,257	-0,225	-0,078	0,046
Item 7	0,783	0,553	0,019	-0,016	0,007	0,025	-0,017	-0,066	0,053	-0,038	0,051
Item 9	0,837	2,569	0,141	-0,089	0,132	-0,049	0,006	-0,269	0,019	0,150	0,100
Item 10	1,857	1,580	0,182	0,191	-0,104	-0,032	-0,055	-0,126	0,036	0,045	0,044
Item 11	0,886	1,315	0,000	-0,010	0,036	-0,013	-0,013	-0,093	-0,111	0,098	0,106
Item 12	1,077	2,677	0,201	0,027	0,003	0,074	-0,104	0,038	-0,119	-0,036	0,116
Item 13	3,105	2,926	0,203	-0,214	-0,626	0,114	0,726	0,004	0,097	-0,048	-0,053
Item 14	3,809	3,556	0,175	1,000	-0,291	-0,148	-0,561	-0,004	0,122	-0,051	-0,067
Item 15	2,293	1,016	0,148	0,041	-0,017	-0,051	0,027	0,033	0,044	0,025	-0,102
Item 16	1,182	4,286	0,294	0,212	0,401	-0,237	-0,375	-0,149	-0,711	0,246	0,614
Item 18	1,207	2,177	0,118	0,028	0,008	-0,048	0,012	-0,047	0,071	0,073	-0,097
Item 19	1,466	3,095	0,185	-0,280	-0,146	0,031	0,395	0,293	0,091	-0,200	-0,184
Item 21	1,383	3,496	0,148	-0,025	0,027	-0,146	0,144	-0,201	-0,078	0,212	0,067
Item 22	2,750	2,874	0,128	0,399	-0,294	-0,115	0,010	-0,101	-0,058	0,069	0,090
Item 23	1,682	2,338	0,104	-0,089	0,136	-0,220	0,173	0,050	-0,109	0,089	-0,030
Item 24	1,648	0,561	0,230	0,050	-0,130	0,140	-0,060	0,001	-0,098	0,112	-0,016
Item 26	0,773	1,277	0,000	-0,002	-0,004	0,007	-0,001	0,024	-0,050	-0,029	0,055
Item 27	0,578	1,898	0,000	-0,006	0,021	0,008	-0,023	0,050	-0,188	0,040	0,098
Item 28	2,311	2,575	0,176	0,487	-0,293	-0,159	-0,035	-0,037	0,019	-0,002	0,021
Item 29	0,577	2,013	0,000	-0,001	0,007	0,008	-0,014	-0,041	-0,052	-0,018	0,111
Item 30	1,149	1,239	0,220	0,087	-0,036	-0,189	0,138	0,085	-0,075	-0,166	0,157
Item 31	0,836	0,004	0,000	0,030	-0,004	-0,046	0,020	0,080	-0,015	-0,152	0,087
Item 32	1,298	3,568	0,222	-0,133	0,346	-0,116	-0,097	0,060	-0,273	0,028	0,185
Item 33	0,811	0,918	0,000	0,067	-0,015	-0,028	-0,025	-0,108	0,006	-0,048	0,150
Item 34	1,619	2,066	0,127	-0,040	0,059	-0,248	0,229	0,020	-0,042	0,071	-0,049
Item 35	1,030	0,588	0,000	0,039	0,021	-0,014	-0,046	0,080	-0,034	0,106	-0,153
Item 36	1,127	1,544	0,148	-0,083	-0,031	-0,027	0,141	0,034	0,025	-0,062	0,002
Item 37	1,827	3,439	0,116	-0,028	0,046	-0,078	0,060	0,081	-0,191	0,010	0,100
Item 38	2,733	1,827	0,187	-0,028	-0,171	-0,056	0,255	-0,001	-0,007	-0,055	0,064
Item 39	0,500	2,742	0,000	0,003	-0,061	-0,006	0,063	0,105	0,261	0,003	-0,369
Item 40	0,946	-0,331	0,000	0,052	-0,026	-0,029	0,003	0,004	0,049	-0,050	-0,003
Item 41	0,732	0,919	0,112	0,082	0,034	-0,059	-0,056	0,211	0,138	-0,203	-0,146
Item 42	0,601	0,696	0,000	0,001	0,004	0,011	-0,016	0,044	0,006	-0,078	0,028
Item 43	1,850	3,092	0,202	-0,641	0,410	0,012	0,220	0,244	-0,002	-0,155	-0,087
Item 44	3,475	2,183	0,110	-0,231	0,143	-0,125	0,213	-0,025	-0,004	0,012	0,017
Item 45	0,558	1,179	0,000	0,014	-0,038	0,008	0,016	0,173	-0,166	-0,122	0,114

Fonte: Microdados ENEM 2014/INEP.

Tabela 5.18 *Estimativas bayesianas dos parâmetros dos itens na modelagem do EPI em uma etapa.*

ITEM	a_i	b_i	c_i	δ_1^a	δ_2^a	δ_3^a	δ_4^a	δ_1^b	δ_2^b	δ_3^b	δ_4^b
Item 1	2,880	3,118	0,146	-0,089	-0,048	0,062	0,076	0,036	0,000	-0,058	0,023
Item 2	1,225	1,919	0,090	-0,004	0,021	0,022	-0,038	-0,035	-0,031	0,001	0,065
Item 3	2,655	3,094	0,192	-0,110	0,385	-0,202	-0,073	-0,029	0,023	-0,056	0,062
Item 4	1,093	0,409	0,000	-0,020	0,011	0,016	-0,007	-0,087	0,106	-0,022	0,002
Item 5	0,680	2,714	0,132	-0,095	0,006	0,062	0,028	0,255	-0,223	-0,077	0,045
Item 7	0,783	0,552	0,019	-0,016	0,007	0,025	-0,017	-0,066	0,053	-0,038	0,050
Item 9	0,837	2,569	0,141	-0,088	0,133	-0,050	0,006	-0,270	0,016	0,155	0,099
Item 10	1,857	1,580	0,182	0,191	-0,105	-0,032	-0,055	-0,126	0,037	0,045	0,044
Item 11	0,886	1,315	0,000	-0,010	0,036	-0,013	-0,013	-0,093	-0,111	0,098	0,106
Item 12	1,076	2,677	0,201	0,027	0,003	0,074	-0,104	0,039	-0,119	-0,036	0,116
Item 13	3,096	2,926	0,203	-0,191	-0,604	0,113	0,683	0,004	0,095	-0,047	-0,052
Item 14	3,750	3,557	0,175	1,000	-0,370	-0,204	-0,426	-0,002	0,127	-0,048	-0,077
Item 15	2,292	1,016	0,148	0,040	-0,018	-0,050	0,027	0,033	0,044	0,025	-0,102
Item 16	1,202	4,257	0,295	0,216	0,408	-0,239	-0,385	-0,147	-0,693	0,234	0,607
Item 18	1,206	2,177	0,118	0,028	0,007	-0,048	0,012	-0,047	0,072	0,073	-0,098
Item 19	1,465	3,096	0,185	-0,284	-0,144	0,034	0,393	0,299	0,088	-0,203	-0,184
Item 21	1,383	3,496	0,148	-0,028	0,024	-0,143	0,147	-0,198	-0,075	0,209	0,064
Item 22	2,750	2,874	0,128	0,389	-0,294	-0,114	0,019	-0,101	-0,056	0,068	0,089
Item 23	1,683	2,337	0,104	-0,088	0,136	-0,220	0,173	0,050	-0,109	0,089	-0,030
Item 24	1,648	0,561	0,230	0,050	-0,130	0,139	-0,060	0,001	-0,098	0,112	-0,016
Item 26	0,773	1,277	0,000	-0,002	-0,004	0,007	-0,001	0,024	-0,050	-0,029	0,055
Item 27	0,578	1,898	0,000	-0,005	0,021	0,008	-0,023	0,048	-0,189	0,042	0,099
Item 28	2,309	2,575	0,176	0,480	-0,291	-0,157	-0,032	-0,036	0,019	-0,002	0,020
Item 29	0,577	2,013	0,000	-0,001	0,007	0,008	-0,014	-0,041	-0,052	-0,018	0,111
Item 30	1,149	1,239	0,220	0,087	-0,036	-0,189	0,138	0,085	-0,075	-0,166	0,157
Item 31	0,836	0,004	0,000	0,030	-0,004	-0,046	0,020	0,080	-0,015	-0,152	0,087
Item 32	1,290	3,578	0,221	-0,133	0,348	-0,114	-0,101	0,062	-0,281	0,028	0,191
Item 33	0,811	0,918	0,000	0,067	-0,015	-0,028	-0,025	-0,108	0,006	-0,048	0,150
Item 34	1,619	2,066	0,127	-0,038	0,059	-0,248	0,228	0,019	-0,042	0,071	-0,049
Item 35	1,030	0,588	0,000	0,039	0,021	-0,014	-0,046	0,080	-0,034	0,106	-0,153
Item 36	1,128	1,544	0,148	-0,083	-0,031	-0,027	0,141	0,034	0,025	-0,062	0,002
Item 37	1,827	3,439	0,116	-0,031	0,045	-0,074	0,060	0,083	-0,190	0,008	0,099
Item 38	2,736	1,827	0,187	-0,031	-0,171	-0,057	0,259	-0,001	-0,007	-0,056	0,064
Item 39	0,500	2,742	0,000	0,004	-0,062	-0,006	0,064	0,098	0,270	0,004	-0,372
Item 40	0,946	-0,331	0,000	0,052	-0,026	-0,029	0,003	0,004	0,049	-0,050	-0,003
Item 41	0,731	0,918	0,111	0,081	0,033	-0,059	-0,056	0,212	0,138	-0,203	-0,146
Item 42	0,601	0,696	0,000	0,001	0,004	0,011	-0,016	0,043	0,006	-0,078	0,028
Item 43	1,850	3,093	0,202	-0,646	0,412	0,011	0,223	0,248	-0,003	-0,157	-0,089
Item 44	3,473	2,183	0,110	-0,228	0,139	-0,127	0,216	-0,025	-0,003	0,012	0,017
Item 45	0,558	1,179	0,000	0,014	-0,038	0,007	0,016	0,173	-0,166	-0,121	0,114

Fonte: Microdados ENEM 2014/INEP.

5.2.5.2 Testes de hipóteses para a significância do EPI em uma etapa

A fim de testar a significância dos parâmetros δ_{it}^a e δ_{it}^b , $t = 1, 2, 3, 4$, foram realizados testes de hipóteses para verificar se havia presença de EPI em algum item do ENEM-2014. Esses testes ocorreram individualmente para cada item em cada caderno, verificando a significância ou não do EPI no parâmetro de discriminação (conforme hipóteses em (5.1)) e de dificuldade (hipóteses em (5.2)).

Na abordagem por máxima verossimilhança, a decisão sobre a existência ou não

de EPI significativa baseou-se nos intervalos de confiança assintóticos. Já na abordagem bayesiana a decisão baseou-se nos intervalos de credibilidade.

Os intervalos assintóticos com 95% de confiança para os parâmetros δ_{i1}^a , δ_{i2}^a , δ_{i3}^a e δ_{i4}^a são apresentados na Tabela A.3. Para os parâmetros δ_{i1}^b , δ_{i2}^b , δ_{i3}^b e δ_{i4}^b , os intervalos assintóticos com 95% de confiança são apresentados na Tabela A.4. Observa-se que os seguintes itens não apresentaram EPI significativo: 1, 2, 3, 12, 13, 16, 18, 19, 21, 22, 26, 29, 32, 37 e 44. A Tabela 5.19 apresenta os itens que apresentaram EPI significativo nos parâmetros de discriminação (a_i) e/ou dificuldade (b_i), com base nos IC assintóticos construídos a partir das estimativas de máxima verossimilhança.

Entre os itens que apresentaram EPI no parâmetro a_i , nenhum item apresentou efeito em todos os cadernos, apenas o item 41 apresentou EPI em 3 cadernos (azul, branco e rosa), 5 itens apresentaram EPI em dois cadernos, o item 24 nos cadernos amarelo e branco, os itens 30 e 34 nos cadernos branco e rosa, o item 36 nos cadernos azul e rosa e o item 39 nos cadernos amarelo e rosa.

Em relação aos itens que apresentaram EPI no parâmetro b_i , os itens 11, 30, 41 e 45 apresentaram efeito nos 4 cadernos, 3 itens apresentaram efeito em 3 cadernos, o item 15 nos cadernos azul, amarelo e rosa, os itens 31 e 35 nos cadernos azul, branco e rosa, 7 itens apresentaram efeito em 2 cadernos, os itens 4 e 7 nos cadernos azul e amarelo, o item 33 nos cadernos azul e rosa, os itens 24 e 40 nos cadernos amarelo e branco, o item 38 nos cadernos branco e rosa e o item 39 nos cadernos amarelo e rosa.

Alguns itens apresentaram EPI em ambos os parâmetros. No caderno azul foram os itens 10, 33 e 41, no caderno amarelo os itens 24 e 39, no caderno branco os itens 24, 30 e 41 e no caderno rosa os itens 30, 39 e 41.

Tabela 5.19 *Itens com EPI nos parâmetros a_i , b_i e em ambos, por caderno, segundo os intervalos de confiança assintóticos.*

CADERNO	Itens com EPI no parâmetro a	Itens com EPI no parâmetro b	Itens com EPI nos parâmetros a e b
Azul	5, 10, 14, 28, 33, 36, 40, 41 e 43	4, 7, 9, 10, 11, 15, 30, 31, 33, 35, 41 e 45	10, 33 e 41
Amarelo	9, 24 e 39	4, 7, 11, 15, 23, 24, 27, 30, 39, 40, 41 e 45	24 e 39
Branco	23, 24, 30, 34 e 41	11, 24, 30, 31, 35, 38, 40, 41, 42 e 45	24, 30 e 41
Rosa	30, 34, 36, 39 e 41	11, 15, 30, 31, 33, 35, 38, 39, 41 e 45	30, 39 e 41

Fonte: Microdados ENEM 2014/INEP.

Os intervalos com 95% de credibilidade para os parâmetros δ_{i1}^a , δ_{i2}^a , δ_{i3}^a e δ_{i4}^a são apresentados na Tabela A.5. Para os parâmetros δ_{i1}^b , δ_{i2}^b , δ_{i3}^b e δ_{i4}^b , os intervalos com 95%

de credibilidade são apresentados na Tabela A.6. Observa-se que os seguintes itens não apresentaram EPI significativo: 2, 3, 7, 9, 13, 14, 16, 18, 21, 22, 23, 24, 28, 29, 30, 32, 33, 34, 35, 39, 40, 41 e 42. A Tabela 5.20 apresenta os itens que apresentaram EPI significativo nos parâmetros de discriminação (a_i) e/ou dificuldade (b_i), com base nos ICr construídos a partir das estimativas bayesianas.

Dentre os itens que apresentaram EPI apenas no parâmetro de discriminação (a), os itens 1, 12, 15, 27 e 37 apresentaram efeito em todos os cadernos, o item 5 apresentou nos cadernos azul, amarelo e rosa, o item 31 apresentou nos cadernos amarelo, branco e rosa e o item 43 apresentou efeito nos cadernos amarelo, branco e rosa.

Em relação aos itens que apresentaram EPI apenas no parâmetro de dificuldade (b), os itens 4, 10, 11, 12, 26, 27, 36, 37, 38, 43, 44 apresentaram efeito em todos os cadernos, o item 1 apresentou nos cadernos azul, branco e rosa, o item 5 apresentou nos cadernos azul, amarelo e rosa, o item 45 apresentou nos cadernos azul, amarelo e branco e o item 19 apresentou apenas nos cadernos amarelo e branco.

Alguns itens apresentaram EPI em ambos os parâmetros, no caderno azul foram os itens 1,5,12,27,37 e 43, no caderno amarelo os itens 5, 12, 27 e 37, no caderno branco os itens 1, 12, 27, 37 e 43, e no caderno rosa os itens 1, 5, 12, 27, 37 e 43.

Tabela 5.20 *Itens com EPI nos parâmetros a_i , b_i e em ambos, por caderno, segundo os intervalos de credibilidade.*

CADERNO	Itens com EPI no parâmetro a	Itens com EPI no parâmetro b	Itens com EPI nos parâmetros a e b
Azul	1, 5, 12, 15, 27, 37 e 43	1, 4, 5, 10, 11, 12, 26, 27, 36, 37, 38, 43, 44 e 45	1, 5, 12, 27, 37 e 43
Amarelo	1, 5, 12, 15, 27, 31 e 37	4, 5, 10, 11, 12, 19, 26, 27, 36, 37, 38, 43, 44 e 45	5, 12, 27 e 37
Branco	1, 12, 15, 27, 31, 37 e 43	1, 4, 10, 11, 12, 19, 26, 27, 36, 37, 38, 43, 44 e 45	1, 12, 27, 37 e 43
Rosa	1, 5, 12, 15, 27, 31, 37 e 43	1, 4, 5, 10, 11, 12, 26, 27, 36, 37, 38, 43 e 44	1, 5, 12, 27, 37 e 43

Fonte: Microdados ENEM 2014/INEP.

Capítulo 6

Considerações finais e recomendações para trabalhos futuros

6.1 Considerações finais

Avaliações educacionais e concursos públicos têm feito parte do cotidiano de uma parcela considerável da população brasileira. A preocupação com a segurança tem sido cada vez maior, dada a importância de alguns testes. O ENEM passou ao status de um teste de grande impacto, pois é o principal instrumento para seleção de candidatos às universidades brasileiras, principalmente as públicas. Uma forma de melhorar a segurança de um teste é preparar diferentes cadernos de teste, alterando as posições dos itens. No caso do ENEM são preparados 4 diferentes cadernos diferenciados pelas posições dos itens, com impressões em 4 cores distintas. Este trabalho visa responder a uma pergunta frequente: será que as características de cada item são robustas à posição em que ele aparece no teste? Ou seja, considerando o modelo logístico de 3 parâmetros da TRI para cada item, se obtivermos estimativas dos parâmetros de discriminação, dificuldade e acerto casual para cada caderno de prova, podemos dizer que elas serão estatisticamente iguais?

Neste trabalho considerou-se apenas possíveis diferenças nos parâmetros de discriminação e dificuldade. Foi apresentada uma revisão da literatura da área sobre o assunto e implementada computacionalmente no software R a devida adaptação ao caso do ENEM na edição 2014, últimos microdados disponíveis quando este trabalho foi iniciado. Em ambos os casos buscou-se, para cada item, identificar o efeito de cada caderno, que equivale à diferença entre o valor da estimativa no caderno e a estimativa obtida quando usamos o conjunto de dados total, formado pelos 4 cadernos.

Relativamente aos métodos disponíveis, foram adotados o processo de estimação dos efeitos em 2 etapas (E2) e o processo de estimação em 1 etapa (E1). Neste último, os efeitos são parâmetros do modelo são estimados diretamente com o conjunto de dados

total. A estimação foi implementada no R, seguida de um estudo de simulação para verificar se havia recuperação dos parâmetros usados na simulação para geração dos dados. Nesta etapa de simulação foram adotados padrões característicos do ENEM, tais como 4 cadernos com 45 itens cada. Adotou-se 10.000 indivíduos em cada um dos cadernos, totalizando 40.000 no conjunto da dados total. Como opção do processo temos o uso de prioris para os parâmetros, caracterizando a estimação bayesiana. A conclusão neste caso é de que o processo recuperou os parâmetros com alta precisão.

Relativamente à aplicação aos dados do ENEM 2014, verificou-se a presença de efeito em alguns itens. Um conjunto de 8 itens apresentaram efeito significativo no parâmetro de discriminação, enquanto 15 apresentaram efeito no parâmetro de dificuldade, e 6 apresentaram efeito significativo em ambos. Em 2 itens verificou-se efeito somente no parâmetro de discriminação, enquanto em 9 verificou-se efeito restrito ao parâmetro de dificuldade.

Uma questão a ser levantada seria quais os fatores que podem acarretar em efeito posição de parâmetros de itens. Considerando que não houve problemas de impressão dos diferentes cadernos, e que a distribuição dos cadernos é aleatória aos candidatos, indicando que os grupos de respondentes são equivalentes, restam alternativas como efeito cansaço, fadiga ou similares. Modelos mais específicos considerando um efeito linear ou quadrático, por exemplo, dos efeitos como função das posições dos itens pode ser mais conclusivo quanto aos fatores que levam à ocorrência de efeito posição nos itens.

6.2 Recomendações para trabalhos futuros

Recomenda-se para trabalhos futuros:

1. Modelagem do efeito fadiga em uma etapa, onde a equação funcional passa a compor a equação de verossimilhança cujos coeficientes passam a ser estimados diretamente.
2. Aplicar esta metodologia aos dados da Prova Brasil, onde cada aluno responde a duas disciplinas, com 26 itens em cada, sendo que metade deles respondem Matemática primeiro, e a outra metade começa com Língua Portuguesa.
3. Realizar a extensão do modelo E1 para o caso multidimensional, onde os itens têm parâmetros relativos a cada dimensão.

Bibliografia

- [1] Albano, A. D. (2013). Multilevel modeling of item position effects. *Journal of Educational Measurement*, 50(4), 408-426.
- [2] Alexandrowicz, R. e Matschinger, H. (2008). Estimation of item location effects by means of the generalized logistic regression model: A simulation study and an application. *Psychology Science Quarterly*, 50, 64-74.
- [3] Andrade, D.F., Tavares, H.R. e Valle, R.C. (2000). Teoria da Resposta ao Item: Conceitos e Aplicações. Associação Brasileira de Estatística: São Paulo.
- [4] Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In F.M. Lord & M.R. Novick. *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley
- [5] Brennan, R. L. (1992). The context of context effects. *Applied Measurement in Education*, 5, 225-264.
- [6] Casella, G. e Berger, R. L. (2011). Inferência estatística-tradução da 2ª edição norteamericana. Cengage Learning.
- [7] Davey, T. e Lee, Y. H. (2011). Potential impact of context effects on the scoring and equating of the multistage GRE R revised general test (Research Report 11-26). Princeton, NJ: Educational Testing Service.
- [8] Debeer, D. e Janssen, R. (2013). Modeling item-position effects within an IRT framework. *Journal of Educational Measurement*, 50, 164-185.
- [9] De Boeck, P. e Wilson, M. (2004). Explanatory item response models: A generalized linear and nonlinear approach. New York, NY: Springer.
- [10] De Boeck, P. (2008). Random item IRT models. *Psychometrika*, 73(4), 533-559.

-
- [11] De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F. e Partchev, I. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software*, 39, 1–28.
- [12] Dorans, N. J. e Lawrence, I. M. (1990). Checking the statistical equivalence of nearly identical test editions. *Applied Measurement in Education*, 3, 245–254.
- [13] Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 37, 359–374.
- [14] Fox, J. P. (2010). Bayesian item response modeling: Theory and applications. Springer Science & Business Media.
- [15] Hanson, B. A. (1996). Testing for differences in test score distributions using loglinear models. *Applied Measurement in Education*, 9, 305–321.
- [16] Hohensinn, C., Kubinger, K. D., Reif, M., Holoher-Ertl, S., Khorramdel, L., e Frebort, M. (2008). Examining item-position effects in large-scale assessment using the Linear Logistic Test Model. *Psychology Science*, 50(3), 391.
- [17] Kang, C. (2014). Linear and Nonlinear Modeling of Item Position Effects. Tese. Nebraska: University of Nebraska.
- [18] Kingston, N. M., & Dorans, N. J. (1984). Item location effects and their implications for IRT equating and adaptive testing. *Applied Psychological Measurement*, 8, 147–154.
- [19] Leary, L. F., e Dorans, N. J. (1985). Implications for altering the context in which test items appear: A historical perspective on an immediate concern. *Review of Educational research*, 55(3), 387-413.
- [20] Lord, F. M. (1952). A theory of test scores (No. 7). Psychometric Monograph.
- [21] MacNicol, K. (1956). Effects of varying order of item difficulty in an unspeeeded verbal test. Unpublished manuscript, Educational Testing Service, Princeton, NJ.
- [22] Meulders, M. e Xie, Y. (2004). Person-by-item predictors. In P. De Boeck & M. Wilson (Eds.), *Explanatory item response models* (pp. 213-246). New York, NY: Springer.
- [23] Meyers, J. L., Murphy, S., Goodman, J., e Turhan, A. (2012). The impact of item

position change on item parameters and common equating results under the 3PL model. In Annual Meeting of the National Council on Measurement in Education, Vancouver, BC.

[24] Molenberghs, G., e Verbeke, G. (2004). An introduction to (generalized (non) linear mixed models. In Explanatory item response models (pp. 111-153). Springer New York.

[25] Mollenkopf, W. G. (1950). An experimental study of the effects on item analysis data of changing item placement and test-time limit. *Psychometrika*, 15, 291–315.

[26] Moses, I., Yang, W., e Wilson, C. (2007). Using kernel equating to assess item order effects on test scores. *Journal of Educational Measurement*, 44, 157–178.

[27] Neto, J. L. H., Junqueira, R. D., e de Oliveira, A. S. (2016). Do Saeb ao Sinaeb: prolongamentos críticos da avaliação da educação básica. *Em Aberto*, 29(96).

[28] Paulino, C. D. M., Turkman, M. A. A. e Murteira, B. (2003). Estatística bayesiana.

[29] Pommerich, M., e Harris, D. J. (2003). Context Effects in Pretesting: Impact on Item Statistics and Examinee Scores.

[30] Pomplun, M. e Ritchie, T. (2004). An investigation of context effects for item randomization within testlets. *Journal of Educational Computing Research*, 30, 243–254.

[31] Rasch, G. (1960). Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen : Danish Institute for Educational Research.

[32] Schweizer, K., Schreiner, M., e Gold, A. (2009). The confirmatory investigation of APM items with loadings as a function of the position and easiness of items: A two-dimensional model of APM. *Psychology Science Quarterly*, 51(1), 47.

[33] Weirich, S., Hecht, M., e Böhme, K. (2014). Modeling item position effects using generalized linear mixed models. *Applied Psychological Measurement*, 38(7), 535-548.

[34] Whitely, E. e Dawis, R. (1976). The influence of test context on item difficulty. *Educational and Psychological Measurement*, 36, 329-337.

[35] Yen, W.M (1980). The extent, causes and importance of context effects on item parameters for two latent trait models. *Journal of Educational Measurement*, 17, 297-311.

Apêndice A

**Intervalos de Confiança e
Credibilidade**

Tabela A.1 *Limites inferior (l_{inf}^a) e superior (l_{sup}^a) dos intervalos de confiança para os parâmetros δ_1^a , δ_2^a , δ_3^a e δ_4^a na modelagem do EPI em 2 etapas.*

ITEM	δ_1^a		δ_2^a		δ_3^a		δ_4^a	
	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a
Item 1	-2,180	0,916	-2,393	0,859	-2,148	0,900	-2,559	0,993
Item 2	-0,050	0,438	0,022	0,534	-0,054	0,428	-0,134	0,344
Item 3	-1,886	0,938	-2,392	1,098	-2,029	0,761	-2,271	0,727
Item 4	-0,021	0,373	-0,074	0,284	-0,015	0,357	0,018	0,402
Item 5	-0,053	0,489	-0,119	0,345	0,090	0,614	0,004	0,522
Item 7	0,156	0,824	-0,314	0,256	-0,255	0,317	-0,298	0,280
Item 9	-0,301	0,155	0,265	0,931	-0,156	0,372	-0,092	0,446
Item 10	-0,296	0,416	-0,513	0,217	-0,194	0,644	-0,487	0,243
Item 11	-0,100	0,352	-0,105	0,329	0,067	0,559	-0,055	0,405
Item 12	-0,262	0,672	-0,320	0,556	-0,166	0,838	-0,616	0,194
Item 13	-1,652	0,822	-1,817	0,429	-1,655	0,921	-1,835	0,575
Item 14	-0,949	0,649	-1,304	0,274	-1,393	0,719	-2,047	1,401
Item 15	-0,258	0,446	-0,233	0,483	-0,363	0,317	-0,382	0,266
Item 16	-1,312	0,996	-1,383	1,541	-1,319	0,939	-1,355	1,007
Item 18	-0,250	0,304	-0,073	0,559	-0,138	0,486	-0,194	0,374
Item 19	-0,728	0,426	-0,653	0,479	-0,486	0,608	-0,433	0,793
Item 21	-0,846	0,690	-0,994	0,452	-1,112	0,686	-1,133	0,833
Item 22	-1,618	0,950	-1,738	0,526	-1,912	0,670	-2,003	0,549
Item 23	-0,764	0,518	-0,682	0,510	-0,789	0,525	-0,856	0,270
Item 24	-0,084	0,392	-0,306	0,140	-0,038	0,432	-0,298	0,134
Item 26	0,035	0,275	0,051	0,301	0,118	0,394	0,099	0,375
Item 27	-0,263	0,297	-0,057	0,523	0,064	0,680	-0,231	0,337
Item 28	-1,329	0,951	-1,612	0,498	-1,534	0,496	-1,578	0,390
Item 29	-0,016	0,494	-0,075	0,405	0,075	0,601	-0,016	0,504
Item 30	-0,263	0,303	-0,371	0,179	-0,223	0,407	-0,110	0,490
Item 31	0,045	0,359	-0,011	0,305	-0,119	0,185	0,018	0,318
Item 32	-1,267	0,961	-1,189	0,823	-1,287	0,713	-1,602	1,132
Item 33	0,077	0,369	0,064	0,386	-0,028	0,274	0,108	0,452
Item 34	-0,467	0,389	-0,330	0,548	-0,568	0,306	-0,400	0,422
Item 35	0,046	0,306	0,055	0,347	0,068	0,356	0,034	0,360
Item 36	-0,410	0,288	-0,125	0,637	-0,299	0,395	-0,368	0,278
Item 37	-1,378	0,890	-1,212	0,804	-1,356	0,706	-1,513	0,839
Item 38	-0,900	0,288	-1,017	0,137	-0,896	0,244	-0,800	0,524
Item 39	-0,069	0,583	-0,106	0,608	-0,071	0,575	-0,209	0,365
Item 40	0,051	0,267	-0,026	0,182	-0,032	0,194	0,016	0,248
Item 41	-0,317	0,213	-0,440	0,076	-0,189	0,437	0,084	0,794
Item 42	-0,095	0,507	0,079	0,725	-0,379	0,161	-0,234	0,350
Item 43	-1,583	0,663	-1,501	0,619	-1,475	0,677	-1,480	0,822
Item 44	-1,519	0,349	-1,520	0,310	-1,586	0,156	-1,491	0,489
Item 45	-0,132	0,354	-0,112	0,408	-0,089	0,407	0,089	0,613

Fonte: Microdados ENEM 2014/INEP.

Tabela A.2 Limites inferior (l_{inf}^b) e superior (l_{sup}^b) dos intervalos de confiança para os parâmetros δ_1^b , δ_2^b , δ_3^b e δ_4^b na modelagem do EPI em 2 etapas.

ITEM	δ_1^b		δ_2^b		δ_3^b		δ_4^b	
	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b
Item 1	-0,147	0,226	-0,076	0,335	-0,220	0,140	-0,048	0,345
Item 2	-0,223	0,027	-0,207	0,037	-0,201	0,051	-0,155	0,112
Item 3	-0,257	0,107	-0,050	0,346	-0,228	0,168	-0,079	0,346
Item 4	-0,119	0,133	-0,069	0,183	-0,150	0,098	-0,103	0,140
Item 5	-0,356	0,355	-0,590	0,036	-0,523	0,095	-0,442	0,214
Item 7	-0,168	0,065	-0,202	0,099	-0,290	-0,011	-0,178	0,116
Item 9	-0,552	-0,044	-0,415	0,034	-0,333	0,242	-0,358	0,187
Item 10	-0,283	-0,127	-0,101	0,076	-0,101	0,060	-0,115	0,058
Item 11	-0,223	0,035	-0,279	-0,028	-0,138	0,110	-0,121	0,141
Item 12	-0,400	0,145	-0,491	0,047	-0,437	0,065	-0,373	0,383
Item 13	-0,218	0,137	-0,117	0,354	-0,214	0,133	-0,181	0,198
Item 14	-0,186	0,227	-1,066	1,232	-0,564	0,339	-0,302	0,410
Item 15	-0,076	0,044	-0,057	0,063	-0,081	0,041	-0,225	-0,106
Item 16	-1,178	0,989	-1,198	0,440	-1,347	1,198	-1,313	1,177
Item 18	-0,262	0,024	-0,179	0,105	-0,180	0,106	-0,286	-0,007
Item 19	-0,338	0,526	-0,431	0,376	-0,620	0,090	-0,521	0,185
Item 21	-0,526	0,069	-0,458	0,304	-0,314	0,456	-0,267	0,487
Item 22	-0,228	0,017	-0,209	0,080	-0,060	0,241	-0,017	0,290
Item 23	-0,138	0,109	-0,247	-0,010	-0,126	0,121	-0,191	0,063
Item 24	-0,134	0,065	-0,255	-0,017	-0,042	0,153	-0,242	-0,008
Item 26	-0,200	0,217	-0,234	0,188	-0,259	0,148	-0,152	0,274
Item 27	-0,293	0,256	-0,481	-0,006	-0,431	0,038	-0,282	0,278
Item 28	-0,187	0,034	-0,120	0,139	-0,151	0,107	-0,133	0,139
Item 29	-0,459	0,115	-0,447	0,135	-0,494	0,056	-0,379	0,239
Item 30	-0,191	0,069	-0,302	-0,034	-0,197	0,065	-0,147	0,101
Item 31	-0,190	0,482	-0,272	0,466	-0,426	0,464	-0,224	0,450
Item 32	-0,478	0,358	-0,587	0,262	-0,548	0,439	-0,336	0,520
Item 33	-0,367	0,003	-0,160	0,244	-0,166	0,280	-0,054	0,332
Item 34	-0,169	0,071	-0,218	0,010	-0,122	0,128	-0,235	-0,005
Item 35	-0,120	0,152	-0,150	0,143	-0,035	0,250	-0,101	0,219
Item 36	-0,139	0,123	-0,172	0,068	-0,232	0,013	-0,251	-0,006
Item 37	-0,291	0,367	-0,479	0,155	-0,357	0,361	-0,217	0,549
Item 38	-0,145	0,025	-0,154	0,018	-0,210	-0,046	-0,081	0,086
Item 39	-0,473	0,157	-0,364	0,392	-0,536	0,073	-0,621	-0,009
Item 40	-0,197	0,392	-0,114	0,504	-0,209	0,509	-0,145	0,534
Item 41	-0,263	0,036	-0,279	0,049	-0,217	0,086	-0,215	0,061
Item 42	-0,309	0,087	-0,330	0,039	-0,420	0,078	-0,267	0,172
Item 43	-0,215	0,329	-0,235	0,314	-0,344	0,149	-0,255	0,260
Item 44	-0,127	0,027	-0,127	0,025	-0,121	0,036	-0,096	0,055
Item 45	-0,273	0,221	-0,321	0,185	-0,397	0,078	-0,369	0,073

Fonte: Microdados ENEM 2014/INEP.

Tabela A.3 Limites inferior (l_{inf}^a) e superior (l_{sup}^a) dos intervalos de confiança para os parâmetros δ_1^a , δ_2^a , δ_3^a e δ_4^a na modelagem do EPI em 1 etapa.

ITEM	δ_1^a		δ_2^a		δ_3^a		δ_4^a	
	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a
Item 1	-0,803	0,615	-0,854	0,758	-0,735	0,856	-0,988	1,150
Item 2	-0,088	0,079	-0,066	0,107	-0,077	0,121	-0,132	0,057
Item 3	-0,902	0,697	-0,405	1,160	-0,999	0,615	-1,301	1,134
Item 4	-0,071	0,031	-0,039	0,062	-0,038	0,069	-0,056	0,042
Item 5	-0,171	-0,021	-0,073	0,086	-0,013	0,137	-0,052	0,106
Item 7	-0,058	0,026	-0,031	0,046	-0,014	0,064	-0,058	0,025
Item 9	-0,183	0,005	0,028	0,235	-0,132	0,035	-0,091	0,103
Item 10	0,035	0,347	-0,240	0,032	-0,174	0,110	-0,199	0,089
Item 11	-0,058	0,037	-0,011	0,083	-0,062	0,037	-0,060	0,034
Item 12	-0,109	0,164	-0,130	0,136	-0,090	0,238	-0,248	0,039
Item 13	-1,015	0,587	-1,342	0,090	-0,779	1,007	-0,278	1,730
Item 14	0,063	1,937	-1,232	0,651	-1,201	0,905	-2,336	1,214
Item 15	-0,097	0,179	-0,143	0,109	-0,186	0,084	-0,110	0,164
Item 16	-0,347	0,770	-0,123	0,924	-0,696	0,222	-1,022	0,272
Item 18	-0,083	0,139	-0,098	0,114	-0,170	0,074	-0,099	0,123
Item 19	-0,582	0,021	-0,438	0,146	-0,286	0,348	-0,032	0,823
Item 21	-0,337	0,286	-0,297	0,350	-0,443	0,151	-0,326	0,615
Item 22	-0,158	0,956	-0,894	0,306	-0,679	0,449	-0,785	0,805
Item 23	-0,261	0,084	-0,076	0,348	-0,402	-0,038	-0,061	0,407
Item 24	-0,061	0,161	-0,222	-0,038	0,037	0,242	-0,152	0,033
Item 26	-0,046	0,041	-0,051	0,042	-0,042	0,057	-0,046	0,044
Item 27	-0,046	0,035	-0,026	0,068	-0,035	0,050	-0,073	0,026
Item 28	0,008	0,966	-0,652	0,067	-0,566	0,247	-0,484	0,413
Item 29	-0,045	0,043	-0,040	0,054	-0,035	0,052	-0,057	0,029
Item 30	-0,006	0,179	-0,107	0,034	-0,275	-0,103	0,032	0,244
Item 31	-0,012	0,072	-0,046	0,039	-0,095	0,002	-0,020	0,060
Item 32	-0,452	0,186	-0,116	0,807	-0,460	0,228	-0,464	0,270
Item 33	0,024	0,111	-0,058	0,028	-0,065	0,010	-0,061	0,012
Item 34	-0,202	0,123	-0,084	0,202	-0,380	-0,117	0,068	0,389
Item 35	-0,008	0,086	-0,021	0,063	-0,058	0,029	-0,092	0,001
Item 36	-0,164	-0,001	-0,114	0,052	-0,102	0,047	0,052	0,229
Item 37	-0,656	0,599	-0,491	0,583	-0,587	0,431	-0,661	0,781
Item 38	-0,317	0,260	-0,409	0,067	-0,374	0,262	-0,059	0,570
Item 39	-0,040	0,046	-0,101	-0,021	-0,047	0,036	0,021	0,106
Item 40	0,004	0,099	-0,072	0,020	-0,072	0,014	-0,041	0,047
Item 41	0,033	0,130	-0,021	0,088	-0,111	-0,007	-0,103	-0,009
Item 42	-0,039	0,041	-0,036	0,044	-0,026	0,048	-0,058	0,026
Item 43	-0,986	-0,297	-0,255	1,074	-0,381	0,404	-0,392	0,832
Item 44	-0,640	0,178	-0,290	0,576	-0,455	0,205	-0,288	0,714
Item 45	-0,026	0,053	-0,079	0,004	-0,030	0,045	-0,023	0,056

Fonte: Microdados ENEM 2014/INEP

Tabela A.4 *Limites inferior (l_{inf}^b) e superior (l_{sup}^b) dos intervalos de confiança para os parâmetros δ_1^b , δ_2^b , δ_3^b e δ_4^b na modelagem do EPI em 1 etapa.*

ITEM	δ_1^b		δ_2^b		δ_3^b		δ_4^b	
	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b
Item 1	-0,207	0,277	-0,267	0,267	-0,270	0,157	-0,262	0,306
Item 2	-0,125	0,055	-0,119	0,058	-0,098	0,099	-0,040	0,170
Item 3	-0,289	0,229	-0,245	0,290	-0,326	0,214	-0,304	0,431
Item 4	-0,124	-0,050	0,064	0,148	-0,060	0,017	-0,032	0,036
Item 5	-0,041	0,555	-0,493	0,043	-0,304	0,148	-0,213	0,305
Item 7	-0,118	-0,013	0,004	0,102	-0,088	0,012	-0,006	0,107
Item 9	-0,483	-0,055	-0,175	0,213	-0,079	0,379	-0,124	0,324
Item 10	-0,179	-0,072	-0,023	0,095	-0,013	0,104	-0,021	0,110
Item 11	-0,162	-0,024	-0,181	-0,040	0,016	0,180	0,037	0,174
Item 12	-0,194	0,271	-0,335	0,098	-0,272	0,201	-0,153	0,385
Item 13	-0,191	0,199	-0,099	0,292	-0,252	0,157	-0,222	0,115
Item 14	-0,600	0,592	-0,623	0,868	-0,745	0,642	-1,202	1,068
Item 15	0,003	0,064	0,011	0,077	-0,005	0,055	-0,130	-0,074
Item 16	-0,844	0,546	-1,444	0,022	-0,587	1,080	-0,685	1,913
Item 18	-0,176	0,083	-0,060	0,202	-0,075	0,221	-0,216	0,022
Item 19	-0,115	0,701	-0,259	0,441	-0,496	0,096	-0,471	0,102
Item 21	-0,623	0,221	-0,484	0,328	-0,233	0,658	-0,458	0,591
Item 22	-0,269	0,067	-0,233	0,118	-0,096	0,233	-0,138	0,318
Item 23	-0,085	0,185	-0,217	-0,002	-0,053	0,231	-0,151	0,090
Item 24	-0,033	0,035	-0,135	-0,061	0,079	0,145	-0,051	0,019
Item 26	-0,051	0,098	-0,131	0,031	-0,117	0,058	-0,026	0,136
Item 27	-0,076	0,175	-0,324	-0,053	-0,095	0,176	-0,063	0,260
Item 28	-0,170	0,095	-0,137	0,175	-0,147	0,142	-0,134	0,176
Item 29	-0,196	0,115	-0,209	0,104	-0,156	0,121	-0,038	0,260
Item 30	0,016	0,154	-0,138	-0,013	-0,243	-0,089	0,086	0,227
Item 31	0,043	0,118	-0,061	0,032	-0,203	-0,102	0,042	0,131
Item 32	-0,427	0,546	-0,832	0,287	-0,492	0,548	-0,493	0,863
Item 33	-0,154	-0,063	-0,048	0,061	-0,100	0,004	0,090	0,210
Item 34	-0,084	0,124	-0,134	0,049	-0,043	0,185	-0,128	0,030
Item 35	0,041	0,119	-0,071	0,004	0,063	0,149	-0,190	-0,115
Item 36	-0,042	0,111	-0,055	0,106	-0,137	0,013	-0,072	0,076
Item 37	-0,352	0,514	-0,605	0,222	-0,430	0,449	-0,467	0,668
Item 38	-0,059	0,057	-0,061	0,046	-0,107	-0,003	0,004	0,123
Item 39	-0,123	0,333	0,021	0,501	-0,212	0,219	-0,556	-0,182
Item 40	-0,039	0,046	0,009	0,090	-0,093	-0,007	-0,046	0,040
Item 41	0,128	0,294	0,054	0,222	-0,281	-0,125	-0,222	-0,070
Item 42	-0,026	0,113	-0,063	0,076	-0,140	-0,017	-0,042	0,099
Item 43	-0,070	0,557	-0,327	0,323	-0,396	0,087	-0,379	0,205
Item 44	-0,091	0,041	-0,072	0,065	-0,055	0,079	-0,056	0,090
Item 45	0,072	0,274	-0,266	-0,066	-0,209	-0,034	0,030	0,199

Fonte: Microdados ENEM 2014/INEP.

Tabela A.5 Limites inferior (l_{inf}^a) e superior (l_{sup}^a) dos intervalos de credibilidade para os parâmetros δ_1^a , δ_2^a , δ_3^a e δ_4^a na modelagem do EPI em 1 etapa.

ITEM	δ_1^a		δ_2^a		δ_3^a		δ_4^a	
	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a	l_{inf}^a	l_{sup}^a
Item 1	-0,455	-0,370	-0,354	-0,265	0,447	0,576	0,152	0,273
Item 2	-0,052	0,053	-0,050	0,071	-0,065	0,052	-0,058	0,050
Item 3	-0,050	0,056	-0,055	0,044	-0,060	0,057	-0,052	0,058
Item 4	-0,033	0,057	-0,051	0,052	-0,054	0,038	-0,045	0,046
Item 5	-0,540	-0,469	0,240	0,360	-0,052	0,041	0,151	0,253
Item 7	-0,034	0,045	-0,045	0,047	-0,032	0,043	-0,051	0,041
Item 9	-0,050	0,049	-0,045	0,046	-0,055	0,042	-0,041	0,053
Item 10	-0,037	0,059	-0,043	0,038	-0,044	0,035	-0,044	0,043
Item 11	-0,065	0,055	-0,050	0,053	-0,060	0,059	-0,064	0,051
Item 12	-0,161	-0,050	0,234	0,365	-0,351	-0,236	0,046	0,158
Item 13	-0,052	0,049	-0,055	0,067	-0,050	0,057	-0,049	0,061
Item 14	-0,057	0,053	-0,062	0,056	-0,067	0,068	-0,063	0,058
Item 15	-0,373	-0,223	0,028	0,173	0,203	0,407	-0,162	-0,030
Item 16	-0,061	0,067	-0,064	0,050	-0,056	0,049	-0,053	0,072
Item 18	-0,065	0,074	-0,051	0,072	-0,071	0,046	-0,069	0,068
Item 19	-0,043	0,041	-0,059	0,041	-0,036	0,070	-0,048	0,050
Item 21	-0,058	0,065	-0,048	0,052	-0,066	0,057	-0,055	0,055
Item 22	-0,043	0,050	-0,041	0,060	-0,045	0,041	-0,049	0,047
Item 23	-0,050	0,063	-0,058	0,044	-0,055	0,057	-0,051	0,061
Item 24	-0,060	0,064	-0,063	0,049	-0,045	0,066	-0,061	0,057
Item 26	-0,054	0,056	-0,057	0,056	-0,066	0,051	-0,062	0,064
Item 27	0,134	0,282	-0,265	-0,137	0,321	0,477	-0,461	-0,335
Item 28	-0,060	0,074	-0,061	0,059	-0,046	0,059	-0,062	0,068
Item 29	-0,062	0,083	-0,069	0,078	-0,081	0,075	-0,075	0,078
Item 30	-0,092	0,105	-0,098	0,109	-0,100	0,112	-0,081	0,081
Item 31	-0,028	0,235	-0,311	-0,115	-0,392	-0,231	0,318	0,570
Item 32	-0,130	0,129	-0,111	0,140	-0,135	0,116	-0,097	0,110
Item 33	-0,102	0,129	-0,158	0,134	-0,157	0,107	-0,136	0,175
Item 34	-0,066	0,077	-0,077	0,083	-0,087	0,102	-0,083	0,074
Item 35	-0,075	0,117	-0,069	0,078	-0,095	0,082	-0,088	0,086
Item 36	-0,103	0,110	-0,088	0,093	-0,101	0,090	-0,084	0,106
Item 37	0,328	0,491	-0,257	-0,130	0,029	0,183	-0,381	-0,238
Item 38	-0,085	0,122	-0,082	0,070	-0,077	0,091	-0,087	0,071
Item 39	-0,104	0,091	-0,102	0,097	-0,093	0,106	-0,083	0,111
Item 40	-0,067	0,067	-0,065	0,066	-0,089	0,068	-0,093	0,078
Item 41	-0,111	0,112	-0,102	0,147	-0,112	0,110	-0,107	0,094
Item 42	-0,115	0,127	-0,144	0,120	-0,104	0,130	-0,111	0,130
Item 43	0,102	0,302	-0,013	0,205	-0,322	-0,063	-0,229	-0,006
Item 44	-0,160	0,197	-0,177	0,168	-0,239	0,237	-0,184	0,163
Item 45	-0,285	0,292	-0,225	0,339	-0,183	0,248	-0,283	0,220

Fonte: Microdados ENEM 2014/INEP.

Tabela A.6 *Limites inferior (l_{inf}^b) e superior (l_{sup}^b) dos intervalos de credibilidade para os parâmetros δ_1^b , δ_2^b , δ_3^b e δ_4^b na modelagem do EPI em 1 etapa.*

ITEM	δ_1^b		δ_2^b		δ_3^b		δ_4^b	
	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b	l_{inf}^b	l_{sup}^b
Item 1	0,004	0,354	-0,083	0,233	-0,307	-0,064	-0,234	-0,002
Item 2	-0,109	0,125	-0,115	0,147	-0,152	0,123	-0,125	0,124
Item 3	-0,132	0,149	-0,155	0,099	-0,173	0,177	-0,141	0,156
Item 4	0,351	0,464	-0,165	-0,028	-0,491	-0,343	0,042	0,159
Item 5	0,301	0,809	-0,432	-0,214	-0,113	0,083	-0,325	-0,084
Item 7	-0,058	0,050	-0,036	0,052	-0,050	0,052	-0,041	0,052
Item 9	-0,058	0,054	-0,060	0,059	-0,063	0,063	-0,060	0,068
Item 10	0,017	0,156	0,224	0,364	-0,347	-0,231	-0,162	-0,029
Item 11	0,117	0,304	-0,174	-0,039	0,311	0,508	-0,565	-0,439
Item 12	0,009	0,206	-0,388	-0,220	0,129	0,428	-0,194	-0,003
Item 13	-0,112	0,135	-0,131	0,124	-0,116	0,109	-0,121	0,112
Item 14	-0,127	0,189	-0,147	0,160	-0,143	0,164	-0,127	0,162
Item 15	-0,212	0,195	-0,167	0,131	-0,150	0,145	-0,157	0,149
Item 16	-0,081	0,071	-0,096	0,080	-0,083	0,087	-0,091	0,093
Item 18	-0,099	0,117	-0,089	0,111	-0,129	0,091	-0,120	0,099
Item 19	-0,058	0,039	-0,258	-0,141	0,165	0,249	-0,062	0,047
Item 21	-0,062	0,055	-0,072	0,071	-0,068	0,063	-0,062	0,072
Item 22	-0,037	0,046	-0,034	0,036	-0,043	0,042	-0,037	0,045
Item 23	-0,051	0,038	-0,042	0,033	-0,040	0,039	-0,036	0,040
Item 24	-0,047	0,041	-0,055	0,039	-0,037	0,044	-0,047	0,050
Item 26	-0,252	-0,145	0,053	0,165	-0,153	-0,049	0,142	0,249
Item 27	-0,259	-0,123	0,108	0,280	-0,446	-0,343	0,293	0,500
Item 28	-0,106	0,094	-0,083	0,087	-0,078	0,091	-0,084	0,096
Item 29	-0,106	0,094	-0,115	0,090	-0,108	0,124	-0,101	0,105
Item 30	-0,117	0,155	-0,108	0,181	-0,135	0,130	-0,145	0,128
Item 31	-0,067	0,059	-0,051	0,066	-0,055	0,063	-0,065	0,045
Item 32	-0,085	0,067	-0,067	0,066	-0,075	0,068	-0,070	0,060
Item 33	-0,075	0,070	-0,097	0,077	-0,080	0,066	-0,091	0,083
Item 34	-0,029	0,031	-0,037	0,027	-0,026	0,030	-0,032	0,032
Item 35	-0,032	0,040	-0,035	0,026	-0,035	0,033	-0,027	0,037
Item 36	-0,438	-0,357	0,267	0,342	-0,149	-0,067	0,165	0,240
Item 37	-0,419	-0,382	0,182	0,222	-0,124	-0,079	0,273	0,319
Item 38	0,473	0,527	-0,329	-0,273	-0,124	-0,071	-0,128	-0,079
Item 39	-0,025	0,031	-0,027	0,028	-0,029	0,029	-0,027	0,022
Item 40	-0,025	0,031	-0,029	0,027	-0,024	0,029	-0,037	0,026
Item 41	-0,028	0,035	-0,030	0,031	-0,032	0,038	-0,031	0,031
Item 42	-0,043	0,032	-0,039	0,040	-0,034	0,032	-0,030	0,041
Item 43	-0,247	-0,151	-0,147	-0,037	0,120	0,282	0,040	0,177
Item 44	-0,377	-0,243	-0,262	-0,123	0,508	0,757	-0,156	-0,022
Item 45	0,190	0,417	0,105	0,301	-0,676	-0,548	-0,007	0,222

Fonte: Microdados ENEM 2014/INEP.