

**JOSÉ ROBERTO SILVA DOS SANTOS**

**Um Modelo de Resposta ao Item para Grupos Múltiplos com  
Distribuições Normais Assimétricas Centralizadas**

Durante a elaboração deste trabalho o autor recebeu auxílio financeiro da CAPES

**Campinas  
2012**



UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO  
CIENTÍFICA  
DEPARTAMENTO DE ESTATÍSTICA

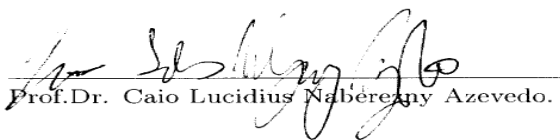
JOSÉ ROBERTO SILVA DOS SANTOS

Um Modelo de Resposta ao Item para Grupos Múltiplos com  
Distribuições Normais Assimétricas Centralizadas

Dissertação de mestrado apresentada ao  
Instituto de Matemática, Estatística e  
Computação Científica da UNICAMP  
para obtenção do título de mestre em Es-  
tatística.

**Orientador:** Prof. Dr. Caio Lucidius Naberezny Azevedo.

Este exemplar corresponde à versão final da dissertação defendida pelo aluno José Roberto Silva dos Santos, e orientada pelo Prof.Dr. Caio Lucidius Naberezny Azevedo.

  
Prof.Dr. Caio Lucidius Naberezny Azevedo.

Campinas, 2012

FICHA CATALOGRÁFICA ELABORADA POR  
MARIA FABIANA BEZERRA MULLER-CRB8/6162  
BIBLIOTECA DO INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E  
COMPUTAÇÃO CIENTÍFICA-UNICAMP

|       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Sa59m | <p>Santos, José Roberto Silva dos, 1984-</p> <p>Um modelo de resposta ao item para grupos múltiplos com distribuições normais assimétricas centralizadas / José Roberto Silva dos Santos. – Campinas, SP: [s.n], 2012.</p> <p>Orientador: Caio Lucidius Naberezny Azevedo.<br/>Dissertação (Mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.</p> <p>1. Teoria da resposta ao item. 2. Distribuição normal assimétrica. 3. Métodos MCMC. I. Azevedo, Caio Lucidius Naberezny, 1979-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título</p> |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Informações para Biblioteca Digital

**Título em Inglês:** A multiple group IRT model with skew-normal latent trait distribution under the centred parametrization

**Palavras-chave em Inglês:**

Item response theory

skew-normal distribution

MCMC methods

**Área de concentração:** Estatística

**Titulação:** Mestre em Estatística

**Banca examinadora:**

Caio Lucidius Naberezny Azevedo [orientador]

Hélio dos Santos Migon

Heleno Bolfarine

**Data de defesa:** 19-04-2012

**Programa de Pos-Graduação:** Estatística

**Dissertação de Mestrado defendida em 19 de abril de 2012 e aprovada**

**Pela Banca Examinadora composta pelos Profs. Drs.**



---

**Prof.(a). Dr(a). CAIO LUCIDIUS NABEREZNY AZEVEDO**



---

**Prof.(a). Dr(a). HÉLIO DOS SANTOS MIGON**



---

**Prof.(a). Dr(a). HELENE BOLFARINE**



*Dedico este trabalho à*

*Helena, Ribamar e Weskley  
com amor.*



*"A mente que se abre a uma nova ideia  
jamais voltará ao seu tamanho original".*

*Albert Einstein*



*Para minha amada mãe Helena*

*"(...)*

*Mãe significa amor  
Seu mérito é o perdão,  
Se o filho comete um erro  
Lhe corta o coração,  
Mas ela para defendê-lo  
Enfrenta um batalhão.*

*Deus lhe deu sabedoria  
E o dom da perfeição,  
Ler dos filhos os pensamentos  
Sabe deles o coração,  
Acredita no amor  
Como a única solução.*

*Tão sensível coração  
E tão forte quando é preciso,  
Dá a vida em prol do filho  
No momento decisivo,  
Por isso mesmo ser mãe  
É padecer no paraíso.*

*Como mulher és exemplo  
Como mãe és uma santa,  
Teu carinho, teu amor  
O mundo inteiro encanta,  
Agora vou encerrar  
Sinto a voz embargar e ficar  
presa à garganta."*

*Edisio Araújo*



*Para minha namorada Sandra*

*"Fonte de mel  
Nos olhos de gueixa  
Kabuki, máscara  
Choque entre o azul  
E o cacho de acácias  
Luz das acácias  
Você é mãe do sol  
A sua coisa é toda tão certa  
Beleza esperta  
Você me deixa a rua deserta  
Quando atravessa  
E não olha pra trás*

*Linda  
E sabe viver  
Você me faz feliz...  
Você é linda  
Mais que demais  
Você é linda sim  
Onda do mar do amor  
Que bateu em mim*

*Você é forte  
Dentes e músculos  
Peitos e lábios  
Você é forte  
Letras e músicas  
Todas as músicas  
Que ainda hei de ouvir  
No Abaeté  
Areias e estrelas  
Não são mais belas  
Do que você  
Mulher das estrelas  
Mina de estrelas  
Diga o que você quer"  
(...)*

*Caetano Veloso*



# Agradecimentos

Nas próximas linhas expresso meu profundo agradecimento àqueles que foram mais do que fundamentais para o desenvolvimento deste trabalho. Dificilmente, conseguiria citar tão brevemente todos aqueles que, de alguma forma, contribuíram para conclusão desta etapa. Porém, as pessoas aqui lembradas representam todos a quem devo minha eterna gratidão. Desta forma, agradeço com respeito e sinceridade à:

Deus, causa primária de todas as coisas, pelo seu inimaginável amor, por ter me concedido saúde e paz para desenvolver este trabalho.

À meus pais, Helena e Ribamar, pelo amor incondicional, pelo apoio em qualquer situação, pelos profundos ensinamentos e orientações, por conduzirem a nossa vidas com sabedoria e dedicação. Em especial à minha mãe pela alegria e conforto de seu sorriso, por todas as palavras e atitudes de amor, pela segurança que sempre me passou e que me fez chegar até aqui.

Ao meu irmão Weskley, por todo o carinho e brincadeiras e pelo apoio a qualquer momento.

À toda minha família, pelo carinho, amor e por compreenderem os momentos de ausência.

À minha namorada Sandra, por todo carinho e ternura que me tem dado nesse um ano de convivência. Pelo seu grande apoio e ajuda durante a elaboração deste trabalho e pelos lindos sonhos compartilhados.

À todos os meus amigos de Fortaleza-CE, especialmente à Tiago, William ("Berada"), Junior ("Chumbim"), Rafael ("Jabbs") e Roberto, pela longa amizade e respeito, por todos os momentos felizes (desde a infância) e pelas calorosas recepções a cada reencontro.

À todos os amigos, professores e funcionários do DEMA-UFC. Em especial aos amigos Dival,

William, Mota, Diego, George, Mardônio, Marcelo, Vladiana ("Vla"), Bruno, Eliardo, Francielton, Patricia, Andréa, Manuela ("Manú"), pela ajuda, apoio e torcida, pelas brincadeiras e momentos de descontração. Aos professores Maurício, Juvêncio, Ailton, Silvia e Ana Maria, pelos conhecimentos repassados, atenção, incentivo e por acreditarem em minha dedicação. Em especial, agradeço ao grande professor Maurício por sua amizade, ajuda e conselhos que me ajudaram bastante a escolher meu caminho após a graduação. Agradeço também, as funcionárias Márgeri, Luísa e Mariluse que sempre nos atenderam com muita atenção e simpatia.

À todos os amigos e colegas do IMECC-UNICAMP, pela ótima convivência, ajuda, churrascos e muitos outros momentos agradáveis. Em especial agradeço às amigas Ângela, Carol, Désirée, Larissa e Mariana com as quais compartilhei, também, a TPQ (tensão pré-qualificação) e aos amigos Carlos ("Peruano"), Carlos ("Baiano"), Nívea e Sheila.

Ao Sr. Newton e Dona Raquel, pelo ótimo convívio em vossa pensão, pelo cuidado e atenção para comigo e pela vossa gentil hospitalidade sempre no intuito de torna mais agradável minha estadia em Campinas-SP.

Ao meu orientador Caio, pelo apoio, respeito e confiança.

Aos membros da banca examinadora Professores Hélio Migon e Heleno Bolfarine, pelas sugestões, críticas e contribuições que deram ao presente trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo suporte financeiro.

# Resumo

Uma das suposições dominantes nos modelos de resposta ao item (MRI) é a suposição de normalidade simétrica para modelar o comportamento dos traços latentes. No entanto, tal suposição tem sido questionada em vários trabalhos como, por exemplo, nos trabalhos de Micceri (1989) e Bazán et.al (2006). Recentemente Azevedo et.al (2011) propuseram um MRI com distribuição normal assimétrica centralizada para os traços latentes, considerando a estrutura de um único grupo de indivíduos. No presente trabalho fazemos uma extensão desse modelo para o caso de grupos múltiplos. Desenvolvemos dois algoritmos MCMC para estimação dos parâmetros utilizando a estrutura de dados aumentados para representar a função de resposta ao item (FRI), veja Albert (1992). O primeiro é um amostrador de Gibbs com passos de Metropolis-Hastings. No segundo utilizamos representações estocásticas (gerando uma estrutura hierárquica) das densidades a priori dos traços latentes e parâmetros populacionais conseguindo, assim, formas conhecidas para todas as distribuições condicionais completas, o que nos possibilitou desenvolver o amostrador de Gibbs completo. Comparamos esses algoritmos utilizando como critério o tamanho efetivo de amostra, veja Sahu (2002). O amostrador de Gibbs completo obteve o melhor desempenho. Também avaliamos o impacto do número de respondentes por grupo, número de itens por grupo, número de itens comuns, assimetria da distribuição do grupo de referência e priori, na recuperação dos parâmetros. Os resultados indicaram que nosso modelo recuperou bem todos os parâmetros, principalmente, quando utilizamos a priori de Jeffreys. Além disso, o número de itens por grupo e o número de examinados por grupo, mostraram ter um alto impacto na recuperação dos traços latentes e parâmetros dos itens, respectivamente. Analisamos um conjunto de dados reais que apresenta indícios de assimetria na distribuição dos traços latentes de alguns grupos. Os resultados obtidos com o nosso modelo confirmam a presença de assimetria na maioria dos grupos. Estudamos algumas medidas de diagnóstico baseadas na distribuição preditiva de medidas de discrepância adequadas. Por último, comparamos os modelos simétrico e assimétrico utilizando os critérios sugeridos por Spiegelhalter et al. (2002). O modelo assimétrico se ajustou melhor aos dados segundo todos os critérios.

**Palavras-chave:** Teoria da resposta ao item, distribuição normal assimétrica, parametrização

centrada, métodos MCMC, datos aumentados, priori de Jeffreys.

# Abstract

An usual assumption for parameter estimation in the Item Response Models (IRM) is to assume that the latent traits are random variables which follow a normal distribution. However, many works suggest that this assumption does not apply in many cases. For example, the works of Micceri (1989) and Bazán (2006). Recently Azevedo et.al (2011) proposed an IRM with skew-normal distribution under the centred parametrization for the latent traits, considering one single group of examinees. In the present work, we developed an extension of this model to account for multiple groups. We developed two MCMC algorithms to parameter estimation using the augmented data structure to represent the Item response function (IRF), see Albert (1992). The First is a Metropolis-Hastings within Gibbs sampling. In the second, we use stochastic representations (creating a hierarchical structure) in the prior distribution of the latent traits and population parameters. Therefore, we obtained known full conditional distributions, which enabled us to develop the full Gibbs sampler. We compared these algorithms using the effective sample size criteria, see Sahu (2002). The full Gibbs sampling presented the best performance. We also evaluated the impact of the number of examinees per group, number of items per group, number of common items, priors and asymmetry of the reference group, on the parameter recovery. The results indicated that our approach recovers properly all parameters, mainly, when we consider the Jeffreys prior. Furthermore, the number of items per group and the number of examinees per group, showed to have a high impact on the recovery of the true of latent traits and item parameters, respectively. We analyze a real data set in which we found an evidence of asymmetry in the distribution of latent traits of some groups. The results obtained with our model confirmed the presence of asymmetry in most groups. We studied some diagnostic measures based on predictive distribution of appropriate discrepancy measures. Finally, we compared the symmetric and asymmetric models using the criteria suggested by Spiegelhalter et al. (2002). The asymmetrical model fits better according to all criteria.

**Keywords:** Item response theory, skew-normal distribution, centred parametrization, MCMC methods, augmented data, Jeffreys prior.



# Sumário

|                                                                                          |           |
|------------------------------------------------------------------------------------------|-----------|
| Agradecimentos                                                                           | xiii      |
| Resumo                                                                                   | xv        |
| Abstract                                                                                 | xvii      |
| Lista de Figuras                                                                         | xxv       |
| Lista de Tabelas                                                                         | xxviii    |
| <b>1 Introdução</b>                                                                      | <b>1</b>  |
| 1.1 Aspectos gerais . . . . .                                                            | 1         |
| 1.2 Motivação e revisão da literatura . . . . .                                          | 2         |
| 1.3 Objetivos e organização da dissertação . . . . .                                     | 5         |
| 1.3.1 Objetivos . . . . .                                                                | 5         |
| 1.3.2 Organização da dissertação . . . . .                                               | 6         |
| <b>2 Distribuição Normal Assimétrica sob a Parametrização Centrada</b>                   | <b>7</b>  |
| 2.1 Introdução . . . . .                                                                 | 7         |
| 2.2 A parametrização centrada . . . . .                                                  | 7         |
| 2.3 Aspectos Estatísticos . . . . .                                                      | 10        |
| 2.3.1 Estimação por máxima verossimilhança . . . . .                                     | 10        |
| 2.3.2 Estimação bayesiana . . . . .                                                      | 17        |
| <b>3 Modelo de Grupos Múltiplos com Distribuições Normais Assimétricas Centralizadas</b> | <b>27</b> |
| 3.1 Introdução . . . . .                                                                 | 27        |
| 3.2 Definição do Modelo . . . . .                                                        | 27        |
| 3.3 Estimação bayesiana e algoritmos MCMC . . . . .                                      | 31        |

|          |                                                                       |            |
|----------|-----------------------------------------------------------------------|------------|
| 3.3.1    | MCMC usando a densidade original da Normal assimétrica centralizada . | 32         |
| 3.3.2    | MCMC sob uma estrutura hierárquica . . . . .                          | 36         |
| <b>4</b> | <b>Estudo de Simulação</b>                                            | <b>41</b>  |
| 4.1      | Introdução . . . . .                                                  | 41         |
| 4.2      | Avaliação da convergência . . . . .                                   | 41         |
| 4.2.1    | Escolha de Prioris . . . . .                                          | 42         |
| 4.3      | Estudo de replicação . . . . .                                        | 55         |
| 4.4      | Comentários e conclusões . . . . .                                    | 63         |
| <b>5</b> | <b>Análise de Dados Reais</b>                                         | <b>65</b>  |
| 5.1      | Introdução . . . . .                                                  | 65         |
| 5.2      | Medidas de diagnóstico . . . . .                                      | 65         |
| 5.3      | Estudo longitudinal em Matemática . . . . .                           | 68         |
| 5.3.1    | Estimação dos parâmetros e avaliação do ajuste . . . . .              | 68         |
| 5.4      | Conclusões e comentários . . . . .                                    | 77         |
| <b>6</b> | <b>Comentários finais e sugestões para futuros trabalhos</b>          | <b>79</b>  |
| 6.1      | Comentários finais . . . . .                                          | 79         |
| 6.2      | Sugestões para futuros trabalhos . . . . .                            | 80         |
| 6.2.1    | Uso de outras funções de resposta ao item . . . . .                   | 80         |
| 6.2.2    | Outras classes de modelos de resposta ao item . . . . .               | 80         |
| 6.2.3    | Outros métodos de estimação . . . . .                                 | 82         |
| 6.2.4    | Medidas de diagnóstico . . . . .                                      | 82         |
| <b>A</b> | <b>Detalhes dos algoritmos MCMC</b>                                   | <b>85</b>  |
| A.1      | Algoritmo ADMHWGS . . . . .                                           | 85         |
| A.2      | Algoritmo ADGS . . . . .                                              | 89         |
| <b>B</b> | <b>Gráficos: Estudo de simulação</b>                                  | <b>95</b>  |
| B.1      | Gráficos das estatísticas REQM . . . . .                              | 95         |
| B.2      | Gráficos das estatísticas VRA . . . . .                               | 105        |
| <b>C</b> | <b>Tabelas: estudo de simulação e análise de dados reais</b>          | <b>115</b> |
|          | <b>Referências Bibliográficas</b>                                     | <b>123</b> |

# Lista de Figuras

|      |                                                                                                                                                   |    |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1  | Densidade da normal assimétrica para diferentes valores de $\gamma$ e $\lambda$ . . . . .                                                         | 9  |
| 2.2  | Duas vezes a LVPR de $\lambda$ (painel esquerdo) e $(\psi, \lambda)$ (painel direito) . . . . .                                                   | 11 |
| 2.3  | Distribuições amostrais de $\hat{\alpha}$ (painel esquerdo) e $(\hat{\alpha}, \hat{\psi})$ (painel direito) para $n = 50$                         | 12 |
| 2.4  | Distribuições amostrais de $\hat{\alpha}$ (painel esquerdo) e $(\hat{\alpha}, \hat{\psi})$ (painel direito) para $n = 100$ . . . . .              | 12 |
| 2.5  | Distribuições amostrais de $\hat{\alpha}$ (painel esquerdo) e $(\hat{\alpha}, \hat{\psi})$ (painel direito) para $n = 150$ . . . . .              | 13 |
| 2.6  | Distribuições amostrais de $\hat{\alpha}$ (painel esquerdo) e $(\hat{\alpha}, \hat{\psi})$ (painel direito) para $n = 200$ . . . . .              | 13 |
| 2.7  | Duas vezes a LVPR de $\gamma$ (painel esquerdo) e $(\sigma, \gamma)$ (painel direito) . . . . .                                                   | 14 |
| 2.8  | Distribuições amostrais de $\hat{\mu}$ (painel esquerdo) e $(\hat{\mu}, \hat{\sigma})$ (painel direito) para $n = 50$                             | 15 |
| 2.9  | Distribuições amostrais de $\hat{\mu}$ (painel esquerdo) e $(\hat{\mu}, \hat{\sigma})$ (painel direito) para $n = 100$                            | 15 |
| 2.10 | Distribuições amostrais de $\hat{\mu}$ (painel esquerdo) e $(\hat{\mu}, \hat{\sigma})$ (painel direito) para $n = 150$                            | 16 |
| 2.11 | Distribuições amostrais de $\hat{\mu}$ (painel esquerdo) e $(\hat{\mu}, \hat{\sigma})$ (painel direito) para $n = 200$                            | 16 |
| 2.12 | Correlogramas dos valores simulados para os parâmetros diretos e centrados . .                                                                    | 21 |
| 2.13 | Gráfico de valores simulados para os parâmetros diretos . . . . .                                                                                 | 21 |
| 2.14 | Gráfico de valores simulados para os parâmetros centrados . . . . .                                                                               | 22 |
| 2.15 | Intervalos de credibilidade e HPD's das estimativas dos parâmetros diretos e centrados . . . . .                                                  | 23 |
| 2.16 | Distribuições amostrais da média a posteriori sob a priori Uniforme (as linhas verticais indicam os verdadeiros valores dos parâmetros) . . . . . | 24 |
| 2.17 | Distribuições amostrais da média a posteriori sob a priori Jeffreys (as linhas verticais indicam os verdadeiros valores dos parâmetros) . . . . . | 24 |
| 3.1  | Esquema de Grupo de referência e itens comuns . . . . .                                                                                           | 29 |
| 4.1  | Distribuições a priori de $\lambda_{\theta_k}$ e $\gamma_{\theta_k}$ considerando as prioris de Jeffreys e Uniforme                               | 44 |
| 4.2  | Distribuições de $\lambda_{\theta_k}$ e $\gamma_{\theta_k}$ para diferentes variâncias considerando a priori normal                               | 45 |

|      |                                                                                                                                                                                    |    |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.3  | Valores simulados para os parâmetros populacionais: Algoritmo ADMHGS. Legenda: $\circ$ conjunto 1, $\triangle$ conjunto 2, $\diamond$ conjunto 3, — valor verdadeiro . . . . .     | 47 |
| 4.4  | Valores simulados para os parâmetros populacionais: Algoritmo ADMHGS (continuação). Legenda: $\circ$ conjunto 1, $\triangle$ conjunto 2, $\diamond$ conjunto 3, — valor verdadeiro | 47 |
| 4.5  | Valores simulados para os parâmetros populacionais: Algoritmo ADGS. Legenda: $\circ$ conjunto 1, $\triangle$ conjunto 2, $\diamond$ conjunto 3, — valor verdadeiro . . . . .       | 48 |
| 4.6  | Valores simulados para os parâmetros populacionais: Algoritmo ADGS (continuação). Legenda: $\circ$ conjunto 1, $\triangle$ conjunto 2, $\diamond$ conjunto 3, — valor verdadeiro   | 48 |
| 4.7  | Valores simulados para os parâmetros dos itens 11 e 60: Algoritmo ADMHGS. Legenda: $\circ$ conjunto 1, $\triangle$ conjunto 2, $\diamond$ conjunto 3, — valor verdadeiro . . . . . | 49 |
| 4.8  | Valores simulados para os parâmetros dos itens 11 e 60: Algoritmo ADGS. Legenda: $\circ$ conjunto 1, $\triangle$ conjunto 2, $\diamond$ conjunto 3, — valor verdadeiro . . . . .   | 49 |
| 4.9  | Correlogramas sem espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADMHGS . . . . .                                                                    | 50 |
| 4.10 | Correlogramas sem espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADGS . . . . .                                                                      | 50 |
| 4.11 | Correlogramas com espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADMHGS . . . . .                                                                    | 51 |
| 4.12 | Correlogramas com espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADGS . . . . .                                                                      | 51 |
| 4.13 | Correlogramas sem espaçamento das amostras geradas para os parâmetros dos itens 11 e 60 Algoritmo ADMHGS . . . . .                                                                 | 52 |
| 4.14 | Correlogramas sem espaçamento das amostras geradas para os parâmetros dos itens 11 e 60: Algoritmo ADGS . . . . .                                                                  | 52 |
| 4.15 | Correlogramas com espaçamento das amostras geradas para os parâmetros dos itens 11 e 60: Algoritmo ADMHGS . . . . .                                                                | 53 |
| 4.16 | Correlogramas com espaçamento das amostras geradas para os parâmetros dos itens 11 e 60: Algoritmo ADGS . . . . .                                                                  | 53 |
| 4.17 | Mediana do REQM das estimativas dos traços latentes. Legenda: ( $-\circ-$ Priori 1, $-\triangle-$ Priori 2, $-*-$ Priori 3 e $-\oplus-$ Priori 4) . . . . .                        | 58 |
| 4.18 | Mediana do REQM das estimativas dos parâmetros de discriminação. Legenda: ( $-\circ-$ Priori 1, $-\triangle-$ Priori 2, $-*-$ Priori 3 e $-\oplus-$ Priori 4) . . . . .            | 58 |
| 4.19 | Mediana do REQM das estimativas dos parâmetros de dificuldade. Legenda: ( $-\circ-$ Priori 1, $-\triangle-$ Priori 2, $-*-$ Priori 3 e $-\oplus-$ Priori 4) . . . . .              | 59 |

|      |                                                                                                                                                                                                       |    |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.20 | Mediana do REQM das estimativas das médias populacionais. Legenda: (—○— Priori 1, —△— Priori 2, —*— Priori 3 e —⊕— Priori 4) . . . . .                                                                | 59 |
| 4.21 | Mediana do REQM das estimativas dos desvios-padrão populacionais. Legenda: (—○— Priori 1, —△— Priori 2, —*— Priori 3 e —⊕— Priori 4) . . . . .                                                        | 60 |
| 4.22 | Mediana do REQM das estimativas dos coeficientes de assimetria populacionais. Legenda: (—○— Priori 1, —△— Priori 2, —*— Priori 3 e —⊕— Priori 4) . . . . .                                            | 60 |
| 5.1  | Distribuição dos escores segundo o MGMA. Legenda: linha cheia: escores preditos, —○— escores observados, — — — intervalo de credibilidade de 95% . . . . .                                            | 70 |
| 5.2  | Distribuição dos escores segundo o MGM. Legenda: linha cheia: escores preditos, —○— escores observados, — — — intervalo de credibilidade de 95% . . . . .                                             | 70 |
| 5.3  | Estimativas dos parâmetros populacionais com intervalos de credibilidade de 95% para o MGM e MGMA . . . . .                                                                                           | 71 |
| 5.4  | Valores-p Bayesianos correspondentes as medidas de discrepância baseadas nas estatísticas de Pearson e <i>deviance</i> . Legenda: ● Estatística qui-quadrado, △ Estatística <i>Deviance</i> . . . . . | 72 |
| 5.5  | Estimativas dos parâmetros de discriminação com intervalos de credibilidade de 95% . . . . .                                                                                                          | 73 |
| 5.6  | Estimativas dos parâmetros de dificuldade com intervalos de credibilidade de 95% . . . . .                                                                                                            | 73 |
| 5.7  | Gráfico de dispersão das estimativas dos traços latentes segundo o MGMA e MGM . . . . .                                                                                                               | 74 |
| 5.8  | Distribuições dos traços latentes por série em escala comum . . . . .                                                                                                                                 | 74 |
| 5.9  | Distribuições dos traços latentes com curvas teóricas segundo o modelo simétrico. Legenda: (— curva estimada, — — — curva teórica) . . . . .                                                          | 75 |
| 5.10 | Distribuições dos traços latentes com curvas teóricas segundo o modelo assimétrico. Legenda: (— curva estimada, — — — curva teórica) . . . . .                                                        | 75 |
| B.1  | Valores de REQM das estimativas das médias populacionais considerando AGR=1 . . . . .                                                                                                                 | 95 |
| B.2  | Valores de REQM das estimativas das médias populacionais considerando AGR=2 . . . . .                                                                                                                 | 96 |
| B.3  | Valores de REQM das estimativas das médias populacionais considerando AGR=3 . . . . .                                                                                                                 | 96 |
| B.4  | Valores de REQM das estimativas dos desvios-padrão populacionais considerando AGR=1 . . . . .                                                                                                         | 97 |
| B.5  | Valores de REQM das estimativas dos desvios-padrão populacionais considerando AGR=2 . . . . .                                                                                                         | 97 |
| B.6  | Valores de REQM das estimativas dos desvios-padrão populacionais considerando AGR=3 . . . . .                                                                                                         | 98 |

|      |                                                                                                           |     |
|------|-----------------------------------------------------------------------------------------------------------|-----|
| B.7  | Valores de REQM das estimativas dos coeficientes de assimetria populacionais considerando AGR=1 . . . . . | 98  |
| B.8  | Valores de REQM das estimativas dos coeficientes de assimetria populacionais considerando AGR=2 . . . . . | 99  |
| B.9  | Valores de REQM das estimativas dos coeficientes de assimetria populacionais considerando AGR=3 . . . . . | 99  |
| B.10 | Valores de REQM das estimativas dos parâmetros de discriminação considerando AGR=1 . . . . .              | 100 |
| B.11 | Valores de REQM das estimativas dos parâmetros de discriminação considerando AGR=2 . . . . .              | 100 |
| B.12 | Valores de REQM das estimativas dos parâmetros de discriminação considerando AGR=3 . . . . .              | 101 |
| B.13 | Valores de REQM das estimativas dos parâmetros de dificuldade considerando AGR=1 . . . . .                | 101 |
| B.14 | Valores de REQM das estimativas dos parâmetros de dificuldade considerando AGR=2 . . . . .                | 102 |
| B.15 | Valores de REQM das estimativas dos parâmetros de dificuldade considerando AGR=3 . . . . .                | 102 |
| B.16 | Valores de REQM das estimativas dos traços latentes considerando AGR=1 . . .                              | 103 |
| B.17 | Valores de REQM das estimativas dos traços latentes considerando AGR=2 . . .                              | 103 |
| B.18 | Valores de REQM das estimativas dos traços latentes considerando AGR=3 . . .                              | 104 |
| B.19 | Valores de VRA das estimativas das médias populacionais considerando AGR=1                                | 105 |
| B.20 | Valores de VRA das estimativas das médias populacionais considerando AGR=2                                | 106 |
| B.21 | Valores de VRA das estimativas das médias populacionais considerando AGR=3                                | 106 |
| B.22 | Valores de VRA das estimativas dos desvios-padrão populacionais considerando AGR=1 . . . . .              | 107 |
| B.23 | Valores de VRA das estimativas dos desvios-padrão populacionais considerando AGR=2 . . . . .              | 107 |
| B.24 | Valores de VRA das estimativas dos desvios-padrão populacionais considerando AGR=3 . . . . .              | 108 |
| B.25 | Valores de VRA das estimativas dos coeficientes de assimetria populacionais considerando AGR=1 . . . . .  | 108 |
| B.26 | Valores de VRA das estimativas dos coeficientes de assimetria populacionais considerando AGR=2 . . . . .  | 109 |

|                                                                                                               |     |
|---------------------------------------------------------------------------------------------------------------|-----|
| B.27 Valores de VRA das estimativas dos coeficientes de assimetria populacionais considerando AGR=3 . . . . . | 109 |
| B.28 Valores de VRA das estimativas dos parâmetros de discriminação considerando AGR=1 . . . . .              | 110 |
| B.29 Valores de VRA das estimativas dos parâmetros de discriminação considerando AGR=2 . . . . .              | 110 |
| B.30 Valores de VRA das estimativas dos parâmetros de discriminação considerando AGR=3 . . . . .              | 111 |
| B.31 Valores de VRA das estimativas dos parâmetros de dificuldade considerando AGR=1 . . . . .                | 111 |
| B.32 Valores de VRA das estimativas dos parâmetros de dificuldade considerando AGR=2 . . . . .                | 112 |
| B.33 Valores de VRA das estimativas dos parâmetros de dificuldade considerando AGR=3 . . . . .                | 112 |
| B.34 Valores de VRA das estimativas dos traços latentes considerando AGR=1 . . . .                            | 113 |
| B.35 Valores de VRA das estimativas dos traços latentes considerando AGR=2 . . . .                            | 113 |
| B.36 Valores de VRA das estimativas dos traços latentes considerando AGR=3 . . . .                            | 114 |



# Lista de Tabelas

|     |                                                                                                                                                            |     |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 2.1 | Resumo a posteriori das estimativas dos parâmetros diretos e centrados . . . . .                                                                           | 22  |
| 2.2 | Viés das estimativas no estudo de simulação considerando as duas prioris . . . . .                                                                         | 25  |
| 2.3 | EQM das estimativas no estudo de simulação considerando as duas prioris . . . . .                                                                          | 25  |
| 4.1 | Hiperparâmetros escolhidos para as prioris 3.14, 3.15, 3.19, 3.20 e 3.21 relacionadas ao algoritmo ADMHGS . . . . .                                        | 43  |
| 4.2 | Hiperparâmetros escolhidos para as prioris 3.14, 3.15, 3.35 e 3.36 relacionadas ao algoritmo ADGS . . . . .                                                | 44  |
| 4.3 | Performance dos algoritmos ADMHGS e ADGS no conjunto de dados referente à avaliação do progresso matemático de estudantes de escolas primárias de Londrina | 54  |
| 4.4 | Parâmetros populacionais escolhidos para cada grupo para cada nível do fator AGR . . . . .                                                                 | 55  |
| 4.5 | ANOVA para os traços latentes . . . . .                                                                                                                    | 62  |
| 4.6 | ANOVA para os parâmetros de discriminação . . . . .                                                                                                        | 62  |
| 4.7 | ANOVA para os parâmetros de dificuldade . . . . .                                                                                                          | 62  |
| 4.8 | Tempo de computação do algoritmo ADGS para cada 1000 iterações . . . . .                                                                                   | 63  |
| 5.1 | Resumo a posteriori dos parâmetros populacionais de cada série segundo os modelos simétrico e assimétrico . . . . .                                        | 71  |
| 5.2 | Critérios para comparação de modelos . . . . .                                                                                                             | 77  |
| C.1 | Parâmetros dos itens escolhidos quando NI=20 e NIC=50%: Estudo de simulação                                                                                | 115 |
| C.2 | Parâmetros dos itens escolhidos quando NI=20 e NIC=25%: Estudo de simulação                                                                                | 116 |
| C.3 | Parâmetros dos itens escolhidos quando NI=40 e NIC=25%: Estudo de simulação                                                                                | 117 |
| C.4 | Parâmetros dos itens escolhidos quando NI=40 e NIC=50%: Estudo de simulação                                                                                | 118 |
| C.5 | ANOVA para os traços latentes considerando todas interações: Estudo de simulação . . . . .                                                                 | 119 |

|     |                                                                                                           |     |
|-----|-----------------------------------------------------------------------------------------------------------|-----|
| C.6 | ANOVA para os parâmetros de discriminação considerando todas interações:<br>Estudo de simulação . . . . . | 120 |
| C.7 | ANOVA para os parâmetros de dificuldade considerando todas interações: Es-<br>tudo de simulação . . . . . | 121 |
| C.8 | Estimativas dos parâmetros de discriminação e dificuldade: análise de dados reais                         | 122 |

# Capítulo 1

## Introdução

### 1.1 Aspectos gerais

O estudo de variáveis latentes é de grande importância em muitas áreas do conhecimento. Tais variáveis possuem como peculiaridade o fato de não poderem ser medidas de forma direta. Exemplos de variáveis latentes são: o conhecimento em Matemática, veja Bazán et.al (2004); nível de maturidade de uma organização, veja Alexandre et.al (2002) e o grau de satisfação de clientes, veja Bayley (2001). As inferências a respeito dessas variáveis são feitas com base nas respostas de indivíduos a instrumentos de medida adequados. Por exemplo, uma prova de matemática, um questionário sobre práticas de gestão de uma organização e uma pesquisa de opinião a respeito das características de um produto. Desta forma faz-se necessário considerar modelos que representem de modo satisfatório o relacionamento entre as variáveis observáveis (respostas aos testes) e as variáveis não observáveis.

A Teoria da Resposta ao Item (TRI) é um conjunto de modelos os quais, em geral, são denominados de modelos de resposta ao item (MRI). De uma forma geral, esses modelos representam o relacionamento entre características não observáveis de indivíduos (ou alguma outra unidade amostral de interesse), também chamadas de traços latentes, algumas características dos instrumentos de medida (teste, questionários etc) e as respostas dadas a esses testes. Mais especificamente, os MRI's modelam a probabilidade de um indivíduo obter um certo escore (pontuação) em um determinado item, condicionada aos traços latentes e as características desse item. Discussões a respeito da TRI e de seus diversos aspectos podem ser encontradas em Hambleton e Swaminathan (1985), Hambleton et al. (1991), Lord (1980), Lord e Novick (1968) e Andrade et al. (2000), por exemplo. Aplicações em vários campos do conhecimento podem ser encontradas em Azevedo (2008), Tavares et al. (2004), Douglas (1999), Tavares e Andrade (2006) e Glass et al. (2009), dentre outros.

## 1.2 Motivação e revisão da literatura

Devido a ampla aplicação da TRI em diversos campos de pesquisa e também com o surgimento de planejamentos, principalmente na área educacional, mais complexos, verificou-se a necessidade de desenvolvimento de modelos de resposta ao item que pudessem levar em consideração as diversas características desses estudos. Dessa forma, é de grande relevância o estudo de MRI's mais flexíveis, de métodos de estimação mais adequados e de medidas de diagnóstico que permitam avaliar a qualidade do ajuste desses modelos.

Particularmente, uma das suposições dominantes na utilização da TRI na análise de dados, tem sido a atribuição da distribuição normal para modelar o comportamento dos traços latentes. Entretanto, tal suposição tem sido contestada em diversos trabalhos. Por exemplo, Micceri (1989) mostra vários conjuntos de dados psicométricos onde a suposição de normalidade da distribuição dos traços latentes não é verificada e é usada apenas por conveniência. Os trabalhos de Azevedo et al. (2009a,b) também apresentam conjuntos de dados onde o pressuposto de normalidade e até mesmo de simetria são questionáveis. Tais pressupostos são também questionados por Samejima (1997). Outros exemplos podem ser encontrados em Bazán et al. (2006).

Por outro lado, muitos trabalhos na literatura apresentam evidências de que a falta de normalidade da distribuição dos traços latentes pode conduzir a estimativas viesadas. Um dos primeiros é devido à Seong (1990), onde é mostrado, para modelos dicotômicos, que a má especificação da distribuição das habilidades pode ter um impacto significativo tanto nas estimativas dos parâmetros dos itens quanto na dos traços latentes. Nesse trabalho foi utilizado o método de máxima verossimilhança marginal (MVM), veja Bock and Aitkin (1981), para estimar os parâmetros dos itens. Kirisci et al. (2001) também obtiveram conclusões similares para o mesmo tipo de modelo, não somente no que diz respeito a falta de normalidade da distribuição dos traços latentes, mas também a respeito da falta de unidimensionalidade destes. Eles mostraram fortes evidências de que a falta de normalidade da distribuição dos traços latentes pode conduzir a estimativas dos parâmetros dos itens viesadas.

Resultados obtidos por De Ayala and Sava-Bolesta (1999), DeMars (2003) e Wollack et al. (2002) para o modelo de resposta nominal indicaram que a distribuição dos traços latentes é responsável por um razoável percentual (em torno de 40%) da variabilidade e da acurácia das estimativas dos parâmetros dos itens. Além disso, como descrito em Azevedo and Andrade (2010), quando a forma da distribuição dos traços latentes (distribuição normal, por exemplo) é apropriada mas os hiperparâmetros não estão devidamente especificados, resultados viesados podem ocorrer.

A escolha de distribuições a priori adequadas para os parâmetros dos itens pode vir a amenizar o viés induzido pela falta de algum dos supracitados pressupostos. No entanto, os impactos

da não-normalidade da distribuição dos traços latentes é ainda significativa; veja Swaminathan et al. (2003), Harwell e Janosky (1991) e Gilfford e Swaminathan (1990), por exemplo.

Muitos trabalhos têm proposto distribuições alternativas para os traços latentes. Por exemplo, Mislevy (1986) considera uma estimação de máxima verossimilhança não-paramétrica da distribuição dos traços latentes baseada em métodos de integração por quadratura gaussiana; veja Baker e Kim (2004). Tais métodos foram desenvolvidos sob MVM e estimação bayesiana marginal, para estimação dos parâmetros dos itens. Por outro lado, Mislevy (1984) considera uma mistura finita de distribuições normais. Há também trabalhos que consideram estimação não paramétrica baseadas em: Histogramas empíricos, veja Mislevy (1984) e Woods (2007), curva Davidian, ver Woods (2009), e também curva de Ramsay, ver Woods (2006, 2008).

Recentemente, Bazán et al. (2006) propuseram um MRI com distribuição normal assimétrica para os traços latentes baseada na parametrização direta (PD); veja Pewsey (2000). Eles não consideram a estimação do parâmetro de assimetria juntamente com o ajuste do modelo, e nem discutiram aspectos relativos à identificabilidade deste. Dessa forma, a apropriada modelagem da distribuição dos traços latentes é de extrema importância para a obtenção de resultados confiáveis.

Por outro lado, é bastante comum, principalmente em dados educacionais, a presença de indivíduos pertencentes à diferentes grupos (populações). Estes grupos podem ser caracterizados por diferentes graus de escolaridade, região, sexo, tipo de escola, etc. Neste caso, levar em consideração as possíveis diferenças existentes entre os grupos é mais apropriado para a obtenção de resultados confiáveis. Nesse tipo de estudo, o primeiro passo para que os resultados relativos aos vários grupos possam ser comparáveis é a exigência de itens comuns nos testes aplicados a estes grupos, criando uma estrutura de ligação entre os mesmos. Um procedimento usual neste tipo de situação, é fazer a estimação em cada um dos grupos e após isso utilizar uma técnica de equalização, veja Kolen e Brennan (1995), por exemplo.

Uma Abordagem alternativa é o Modelo para grupos múltiplos introduzido por Bock e Zimowski (1997). Este modelo supõe que as diferenças entre os grupos podem ser traduzidas por diferentes distribuições de probabilidade. Tais distribuições estão associadas ao processo de amostragem dos indivíduos, mas também, podem ser vistas como distribuições a priori para os traços latentes (no contexto bayesiano).

Usualmente, são utilizadas distribuições normais simétricas (com diferentes parâmetros). A grande vantagem dessa abordagem reside no fato de que a equalização é feita automaticamente no próprio processo de estimação. Desta forma, as estimativas dos parâmetros não estarão mais sujeitas a diferenças devidas ao método de equalização escolhido. Além disso, essa metodologia requer um menor número de itens comuns, em comparação com outros métodos Andrade

(1999).

Modelos mais elaborados necessitam de métodos de estimação que levem em consideração as complexidades inerentes. Dessa forma, o processo de estimação se constitui em uma das etapas mais importantes da utilização da TRI (Baker and Kim (2004) e Azevedo (2003)). Métodos de estimação adequados tornam possível a obtenção de resultados confiáveis. Uma metodologia que tem mostrado potencial utilidade na estimação de modelos da TRI são os então chamados algoritmos de Monte Carlo via Cadeias de Markov (MCMC) Gilks et al. (1996). Tais algoritmos permitem considerar facilmente diferentes distribuições para os traços latentes, diversas estruturas de covariância, modelos de regressão para os parâmetros populacionais (curvas de crescimento), ausência de respostas pelo planejamento, ausência de repostas devido a mecanismos de não-resposta informativos dentre outras. Portanto, a implementação computacional destes métodos, permite a aplicação da TRI em diversas situações de interesse.

Vários trabalhos têm desenvolvido algoritmos MCMC para estimação de MRI's. O primeiro deles é devido a Albert (1992), que desenvolve um amostrador de *Gibbs*, veja Gelfand and Smith (1999), sob a metodologia de dados aumentados Tanner and Wong (1987), para o modelo proibido de 2 parâmetros considerando um único grupo de respondentes. Mais tarde Patz and Junker (1999) desenvolveram um amostrador de *Gibbs* com passos de *Metropolis-Hastings* para os modelos logísticos de 1, 2 e 3 parâmetros e para o modelo de crédito parcial generalizado. Outros exemplos podem ser vistos em Ghosh et al. (2000) e Wollack et al. (2002). Este último propõem um amostrador de *Gibbs* com rejeição adaptativa para o ajuste do modelo de resposta nominal sob normalidade dos traços latentes, utilizando o pacote WinBugs, veja Lunn et al. (2000). Um amostrador de *Gibbs* com a estrutura de dados aumentados para o modelo proibido de 3 parâmetros foi desenvolvido por Béguin and Glas (2001) que também o estendeu para o caso multidimensional. Fox and Glas (2001) também usam o método de *Gibbs* para estimação do modelo proibido de 2 parâmetros com estrutura multinível. Outros trabalhos que utilizam MCMC na TRI são os de Bazán (2006), Siharay (2006), Azevedo (2008), Azevedo and Andrade (2010), Azevedo et al. (2011).

Um outro aspecto relevante diz respeito à avaliação da qualidade do ajuste dos modelos. Nos últimos anos, têm-se explorado algumas técnicas, nessa direção, aplicáveis em MRI's mais simples. Um importante trabalho nessa linha, é o de Sinharay (1996), que utiliza ferramentas baseadas na distribuição preditiva, veja Gelman et al. (2004), dos escores e de algumas medidas de discrepância para avaliar o ajuste dos modelos logísticos de 2 e 3 parâmetros. Entretanto, para modelos com estruturas mais complexas, existem poucas ferramentas disponíveis. Particularmente, com relação aos modelos de grupos múltiplos e assimétricos existem diversas lacunas a serem preenchidas. Azevedo et al. (2011) e Azevedo et al. (2012), adaptaram a metodologia

considerada por Siharay (1996), para o caso assimétrico e de grupos múltiplos, respectivamente. No entanto, muito ainda precisa ser explorado, principalmente no que diz respeito à verificação do afastamento de suposições específicas, tais como, ausência de normalidade para os traços latentes e existência de unidimensionalidade. Além disso, métodos de diagnóstico adequados podem ser úteis para avaliar a qualidade do instrumento de medida (prova, questionário, etc). Por isso, no contexto da TRI, é fundamental a verificação da qualidade do ajuste não só do modelo como um todo, como também da avaliação por grupos ou por item, por exemplo.

## 1.3 Objetivos e organização da dissertação

### 1.3.1 Objetivos

O principal objetivo do presente trabalho é propor um modelo de grupos múltiplos com distribuições normais assimétricas centralizadas para os traços latentes no contexto da Teoria de Resposta ao Item, possibilitando, não só, contemplar possíveis diferenças entre os grupos envolvidos na análise, mas também permitir uma maior flexibilidade para o comportamento dos traços latentes. Tal modelo constituir-se-á em uma generalização dos trabalhos de: Azevedo et al. (2011), Albert (1992), Bock and Zimowski (1997) e Bazán (2005), deste último, com relação ao modelo PNA.

Para tal, seguindo Azevedo et al. (2011), utilizaremos o modelo probito unidimensional de 2 parâmetros. Além disso intencionamos, especificamente:

- Propor um modelo probito unidimensional de 2 parâmetros para grupos múltiplos com distribuições normais assimétricas centralizadas (Pewsey (2000)) para modelar os traços latentes dos diferentes grupos;
- Estudar aspectos de identificabilidade do modelo proposto;
- Propor algoritmos MCMC para estimação do modelo proposto baseados, tanto na densidade original da normal assimétrica, quanto em sua representação estocástica proposta por Henze (1986);
- Estudar os mecanismos de validação de modelos considerados por Azevedo et al. (2012);
- Implementar a metodologia desenvolvida na linguagem de programação Ox<sup>TM</sup>  
<http://www.doornik.com/products.html>;

- Realizar um estudo de simulação para avaliar a acurácia do modelo e do método de estimação proposto, e também, avaliar a influência de fatores específicos na precisão das estimativas;
- Analisar o conjunto de dados considerado em Azevedo et al. (2012).

### 1.3.2 Organização da dissertação

O presente trabalho está organizado da seguinte forma. No Capítulo 1 apresentamos a motivação do nosso estudo com uma revisão da literatura e indicamos os nossos objetivos. No Capítulo 2, é feita uma revisão da distribuição normal assimétrica centralizada discutindo alguns aspectos relacionados a estimação sob os pontos de vista frequentista e bayesiano. Já no Capítulo 3 apresentamos o modelo proposto e também os algoritmos MCMC desenvolvidos para a estimação do mesmo. O Capítulo 4 possui a seguinte divisão: na seção 4.2 avaliamos a convergência dos algoritmos propostos e os comparamos em termos do tamanho efetivo de amostra (Sahu, 2002). Na seção 4.3 mostramos os resultados de um estudo de replicação realizado utilizando o algoritmo que apresentou melhor performance no estudo feito na seção 4.2. E finalmente na seção 4.4 apresentamos algumas conclusões e comentários adicionais. No Capítulo 5 analisamos um conjunto de dados referente a evolução do conhecimento matemático de crianças do ensino fundamental de escolas públicas da cidade de Londrina - PR. Por último, no Capítulo 6, tecemos mais alguns comentários e apresentamos sugestões para futuros trabalhos

## Capítulo 2

# Distribuição Normal Assimétrica sob a Parametrização Centrada

### 2.1 Introdução

Neste capítulo apresentamos uma revisão da distribuição normal assimétrica sob a parametrização centrada. Mostraremos sua relação com a distribuição normal assimétrica usual, bem como algumas de suas propriedades. Questões inferenciais também serão abordadas sob ambas metodologias, clássica e bayesiana.

### 2.2 A parametrização centrada

A distribuição normal assimétrica (NA) pertence a uma subclasse de distribuições formada a partir da classe das distribuições elípticas, veja Branco e Valle (2004), para mais detalhes. Esta tem sido utilizada na modelagem de dados assimétricos em várias áreas, incluindo a educacional, de acordo com Bazán et.al (2004) e Azevedo et.al (2009a).

Visando facilitar aspectos inferenciais, algumas representações estocásticas foram propostas na literatura. Duas representações bem conhecidas são as de Henze (1986) e Sahu et.al (2003). Azevedo et.al (2009b) mostram que essas densidades são exatamente as mesmas sob a parametrização centrada. Dessa forma, nossos desenvolvimentos se aplicam em qualquer uma dessas representações. Seguiremos utilizando a de Henze.

Uma variável aleatória  $Y$  tem distribuição normal assimétrica com parâmetro de localização  $\alpha \in \mathbb{R}$ , parâmetro de escala  $\psi \in \mathbb{R}^+$  e parâmetro de assimetria  $\lambda \in \mathbb{R}$  (notação:  $NA(\alpha, \psi, \lambda)$ ) se sua f.d.p é dada por

$$f(y; \alpha, \psi, \lambda) = 2\psi^{-1}\phi\left(\frac{y-\alpha}{\psi}\right)\Phi\left(\lambda\frac{y-\alpha}{\psi}\right)I(y)_{(-\infty, +\infty)}, \quad (2.1)$$

onde  $\phi$  e  $\Phi$  denotam a função de densidade e função de distribuição da distribuição normal padrão, respectivamente. A esperança e variância de  $Y$  são dadas por

$$\mathbb{E}(Y) = \alpha + \psi\delta r \quad \text{e} \quad \text{Var}(Y) = \psi^2(1 - r^2\delta^2),$$

em que,

$$r = \sqrt{\frac{2}{\pi}}$$

Já  $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$  que assume valores no intervalo  $(-1, 1)$ , pode ser usado como uma parametrização alternativa com interpretação similar a de  $\lambda$ .

Azzalini (1985) introduziu a parametrização centrada definida a partir da seguinte identidade

$$Y_c = \alpha + \psi Z = \mu + \sigma Z_0, \quad (2.2)$$

em que  $Z \sim \text{NA}(0, 1, \lambda)$ ,  $Z_0 = \sigma_z^{-1}(Z - \mu_z)$  e

$$\mu_z = \mathbb{E}(Z) = r\delta, \quad \sigma_z^2 = \text{Var}(Z) = 1 - \mu_z^2.$$

Os parâmetros  $\mu$  e  $\sigma$  são definidos implicitamente de modo que (2.2) seja satisfeita. A parametrização centrada (PC) é formada então por  $(\mu, \sigma, \gamma)$  cujas expressões explícitas são dadas por

$$\begin{aligned} \mu = \mathbb{E}(Y_c) &= \alpha + \psi\mu_z, \\ \sigma = \sqrt{\text{Var}(Y_c)} &= \psi\sqrt{1 - \mu_z^2}, \\ \gamma = \frac{\mathbb{E}[(Y_c - \mathbb{E}(Y_c))^3]}{\text{Var}(Y_c)^{3/2}} &= r\delta^3 \left[ \frac{4}{\pi} - 1 \right] [1 - \mu_z^2]^{-3/2}. \end{aligned} \quad (2.3)$$

Neste caso,  $\mu$  e  $\sigma$  representam, literalmente, o valor esperado e o desvio padrão da variável aleatória  $Y_c$ . O parâmetro  $\gamma$  denota o coeficiente de assimetria de Pearson, este assume valores no intervalo  $(-0,99527; 0,99527)$ . Para a normal assimétrica sob a parametrização centralizada usaremos a seguinte notação:  $Y_c \sim \text{NA}_c(\mu, \sigma, \gamma)$ . A partir da definição dada pela equação (2.2),

podemos chegar a densidade de  $Y_c$ . Pelo método do jacobiano temos

$$\begin{aligned} f_{Y_c}(y) &= 2 \frac{\sqrt{\sigma_z^2}}{\sigma} \phi \left[ \frac{\sqrt{\sigma_z^2}}{\sigma} \left( y - \mu + \frac{\sigma}{\sqrt{\sigma_z^2}} \mu_z \right) \right] \Phi \left( \lambda \left[ \frac{\sqrt{\sigma_z^2}}{\sigma} \left( y - \mu + \frac{\sigma}{\sqrt{\sigma_z^2}} \mu_z \right) \right] \right) \\ &= 2\omega^{-1} \phi(\omega^{-1}(y - \xi)) \Phi[\lambda(\omega^{-1}(y - \xi))], \end{aligned} \quad (2.4)$$

que possui a forma de uma normal assimétrica de parâmetros  $\xi$ ,  $\omega$  e  $\lambda$  em que,

$$\xi = \mu - \frac{\sigma \mu_z}{\sigma_z} \quad \text{e} \quad \omega = \frac{\sigma}{\sigma_z}.$$

Em termo dos parâmetros centrados  $\mu$ ,  $\sigma$  e  $\gamma$  temos

$$\begin{aligned} \xi &= \mu - \sigma \gamma^{1/3} s, \\ \omega &= \sigma \sqrt{1 + \gamma^{2/3} s^2}, \end{aligned} \quad (2.5)$$

$$\begin{aligned} \lambda &= \frac{\gamma^{1/3} s}{\sqrt{r^2 + s^2 \gamma^{2/3} (r^2 - 1)}} \text{ em que} \\ s &= \left( \frac{2}{4 - \pi} \right)^{1/3}. \end{aligned} \quad (2.6)$$

Na Figura 2.1 vemos a densidade dada em (2.4) com  $\mu = 10$ ,  $\sigma = 5$  e diferentes valores do coeficiente de assimetria  $\gamma$  juntamente com o correspondente valor do parâmetro de assimetria  $\lambda$ . Podemos ver que a reparametrização preserva o comportamento assimétrico da densidade dada em (2.1).

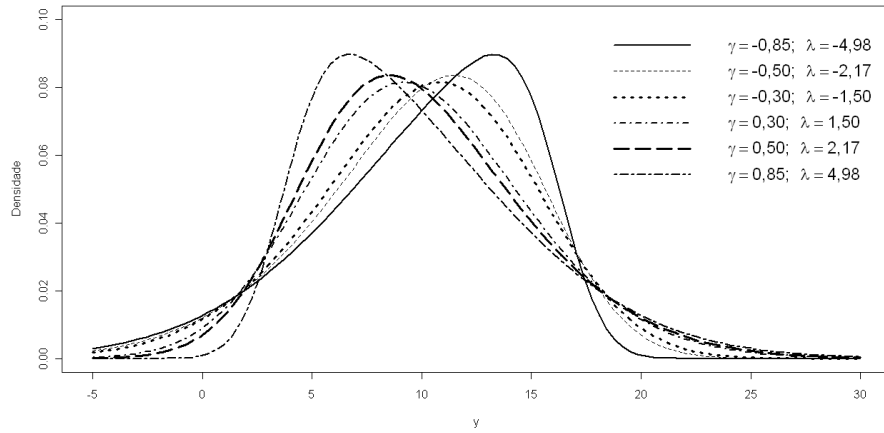


Figura 2.1: Densidade da normal assimétrica para diferentes valores de  $\gamma$  e  $\lambda$

## 2.3 Aspectos Estatísticos

Vamos agora discutir alguns métodos de estimação dos parâmetros do modelo normal assimétrico. Na subseção 2.2.1 discutiremos o método de máxima verossimilhança. Este apresenta alguns problemas, como por exemplo, a não convergência dos estimadores quando se considera a parametrização direta. Na subseção 2.2.2 abordaremos a metodologia bayesiana, discutindo a escolha de prioris, representação hierárquica e a construção de um algoritmo MCMC para estimação dos parâmetros.

### 2.3.1 Estimação por máxima verossimilhança

Desde o artigo pioneiro de Azzalini (1985), tem-se conhecimento da grande dificuldade na estimação por máxima verossimilhança no modelo NA. Tais dificuldades estão relacionadas com a não convergência dos estimadores para particulares resultados amostrais. Por exemplo, no caso uniparamétrico ( $NA(0, 1, \lambda)$ ), dado uma amostra aleatória dessa distribuição pode-se mostrar (Branco et.al, 2004) que: se todas as observações  $x'_i$ s são positivas (negativas) então a verossimilhança será monotonamente crescente (decrecente) e o estimador de máxima verossimilhança de  $\lambda$  será infinito. No caso geral (três parâmetros) observamos que, há sempre um ponto de inflexão em  $\lambda = 0$  na verossimilhança perfilada. Correspondentemente, a matriz de informação de Fisher esperada é singular, quando  $\lambda \rightarrow 0$ . Vamos ilustrar tais problemas. Dora-vante utilizaremos as siglas PD, para nos referir à parametrização direta (ou parâmetros diretos) e PC para fazer referência à parametrização centrada (ou parâmetros centrados). Temos que, para uma amostra aleatória  $Y_1, \dots, Y_n$  de uma  $NA(\alpha, \psi, \lambda)$  a função de log-verossimilhança é dada por

$$l_{PD}(\alpha, \psi, \lambda) = \frac{n}{2} \ln\left(\frac{2}{\pi}\right) - n \ln \psi - \frac{1}{2} \sum_{i=1}^n \left(\frac{y_i - \alpha}{\psi}\right)^2 + \sum_{i=1}^n \ln \left\{ \Phi \left[ \lambda \left(\frac{y_i - \alpha}{\psi}\right) \right] \right\}. \quad (2.7)$$

Uma vez que a log-verossimilhança depende dos três parâmetros, consideraremos a correspondente log-verossimilhança perfilada. Por exemplo, a log-verossimilhança perfilada de  $\lambda$  é dada por  $l_{PD}(\hat{\alpha}(\lambda), \hat{\psi}(\lambda), \lambda)$ , onde  $\hat{\alpha}(\lambda)$  e  $\hat{\psi}(\lambda)$  denotam os valores de  $\alpha$  e  $\psi$  que maximizam  $l_{PD}$  para algum valor de  $\lambda$ .

Seguindo Arellano-Valle and Azzalini (2008), para facilitar a interpretação, iremos considerar a log-verossimilhança perfilada relativa (LVPR), obtida pela subtração do valor máximo  $l_{PD}(\hat{\alpha}(\lambda), \hat{\psi}(\lambda), \hat{\lambda})$ . Então o máximo da nova função será igual a zero. A LVPR para  $(\psi, \lambda)$  é definida de maneira similar.

A Figura 2.2 mostra os gráficos de duas vezes a LVPR de  $\lambda$  e  $(\psi, \lambda)$  este último apresentado num gráfico de contornos, considerando uma amostra aleatória de tamanho 200 da  $NA(0, 1, 1)$ . Utilizamos o pacote `sn` do software R para gerar os resultados desta subseção.

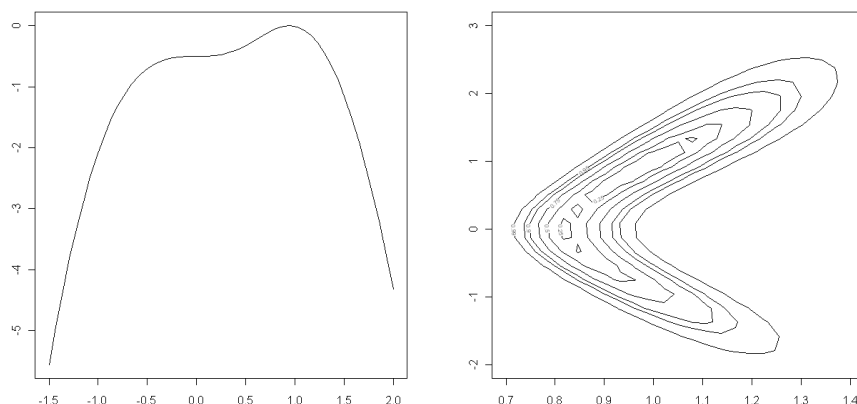


Figura 2.2: Duas vezes a LVPR de  $\lambda$  (painel esquerdo) e  $(\psi, \lambda)$  (painel direito)

O gráfico à esquerda exibe uma forma não quadrática da log-verossimilhança gerada pela existência de um ponto de inflexão em  $\lambda = 0$ . Já no gráfico à direita vemos uma forma bastante peculiar dos contornos. Isso indica que pode existir um ponto de sela na superfície de log-verossimilhança de  $(\psi, \lambda)$ .

Este ponto de inflexão não é um traço peculiar destes dados. Arnold et al. (1993) e Arellano-Valle and Azzalini (2008) utilizam conjuntos de dados reais que exibem o mesmo comportamento da LVPR. O mesmo pode ser visto em Chiogna (2005) que utiliza dados simulados.

O outro problema que está relacionado à existência deste ponto de inflexão é a singularidade da matriz de informação de Fisher esperada quando  $\lambda$  tende a zero. Este fato afeta não somente as propriedades assintóticas dos estimadores de máxima verossimilhança como também impede que avaliemos a existência de assimetria usando testes baseados nesta matriz, como por exemplo, o teste de Wald. Para ter uma ideia direta deste problema, conduzimos um pequeno estudo de simulação. Nesse estudo foram geradas 1000 amostras de tamanho 50, 100, 150 e 200 da  $NA(0, 1, 1)$ , e para cada uma delas foram calculados os estimadores de máxima verossimilhança  $(\hat{\alpha}, \hat{\psi}, \hat{\lambda})$ .

As Figuras 2.3 a 2.6 mostram as correspondentes distribuições amostrais de  $\hat{\alpha}$  e de  $(\hat{\alpha}, \hat{\psi})$  para os diferentes tamanhos amostrais, em forma de histograma e gráfico de dispersão, respectivamente.

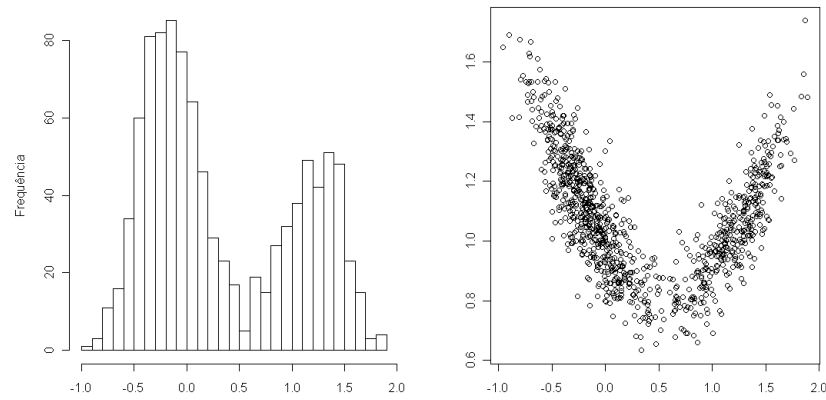


Figura 2.3: Distribuições amostrais de  $\hat{\alpha}$  (painel esquerdo) e  $(\hat{\alpha}, \hat{\psi})$  (painel direito) para  $n = 50$

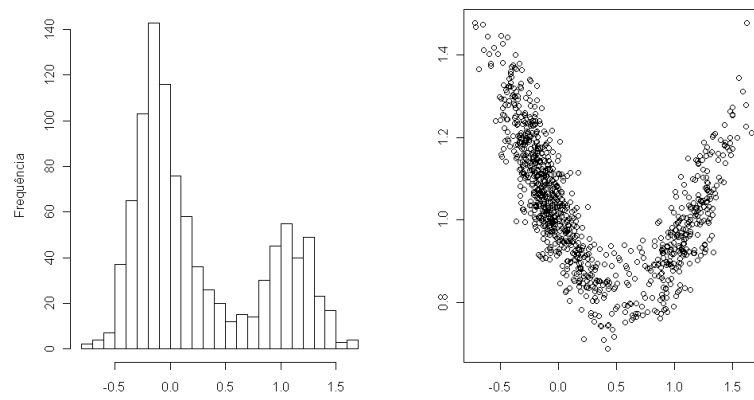


Figura 2.4: Distribuições amostrais de  $\hat{\alpha}$  (painel esquerdo) e  $(\hat{\alpha}, \hat{\psi})$  (painel direito) para  $n = 100$

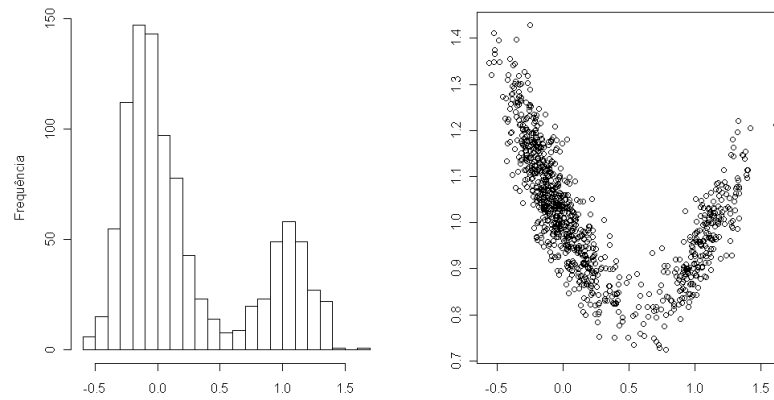


Figura 2.5: Distribuições amostrais de  $\hat{\alpha}$  (painel esquerdo) e  $(\hat{\alpha}, \hat{\psi})$  (painel direito) para  $n = 150$

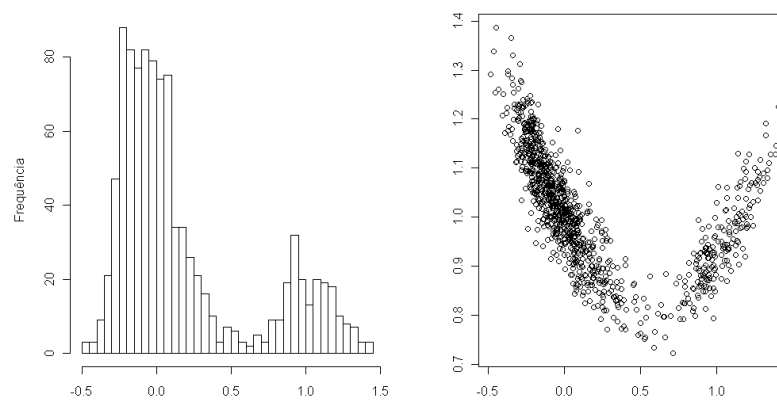


Figura 2.6: Distribuições amostrais de  $\hat{\alpha}$  (painel esquerdo) e  $(\hat{\alpha}, \hat{\psi})$  (painel direito) para  $n = 200$

A forma dessas distribuições amostrais sugerem, basicamente, que existe um viés nas estimativas que permanece ainda quando  $n \rightarrow \infty$ . Segundo Arellano-Valle and Azzalini (2008) este problema pode persistir até mesmo a alguma distância do ponto  $\lambda = 0$ .

Passando agora à parametrização centrada temos que a log-verossimilhança é dada por

$$l_{PC}(\mu, \sigma, \gamma) = \frac{n}{2} \ln\left(\frac{2}{\pi}\right) - n \ln(\omega) - \frac{1}{2} \sum_{i=1}^n \left(\frac{y_i - \xi}{\omega}\right)^2 + \sum_{i=1}^n \ln \left\{ \Phi \left[ \lambda \left( \frac{y_i - \xi}{\omega} \right) \right] \right\}, \quad (2.8)$$

em que  $\xi, \omega$  e  $\lambda$  são definidos como nas equações 2.6. A Figura 2.7 mostra o gráfico de duas vezes a LVPR agora utilizando os parâmetros  $\sigma$  e  $\gamma$ .

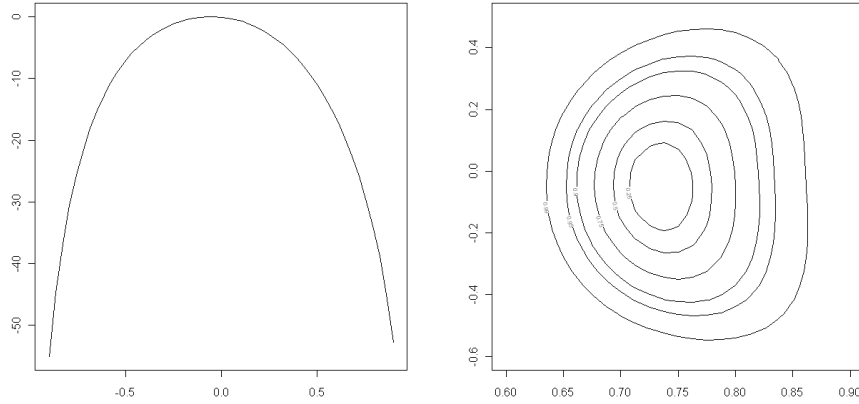


Figura 2.7: Duas vezes a LVPR de  $\gamma$  (painel esquerdo) e  $(\sigma, \gamma)$  (painel direito)

Diferentemente do que foi visto na Figura 2.2 os novos gráficos exibem um comportamento mais regular e mais próximo de uma função convexa. Além disso, não há ponto de inflexão em  $\gamma = 0$ .

Também o gráfico de contornos visto à direita não fornece indícios de que haja um ponto de sela na superfície de verossimilhança de  $(\sigma, \gamma)$ .

Agora vamos repetir o experimento descrito no início desta subseção. Note que PD= (0, 1, 1) corresponde à PC= (0, 564; 0, 826; 0, 137). As Figuras 2.8 a 2.11 mostram os gráficos das distribuições amostrais de  $\hat{\mu}$  o novo parâmetro de localização e  $(\hat{\mu}, \hat{\sigma})$ .

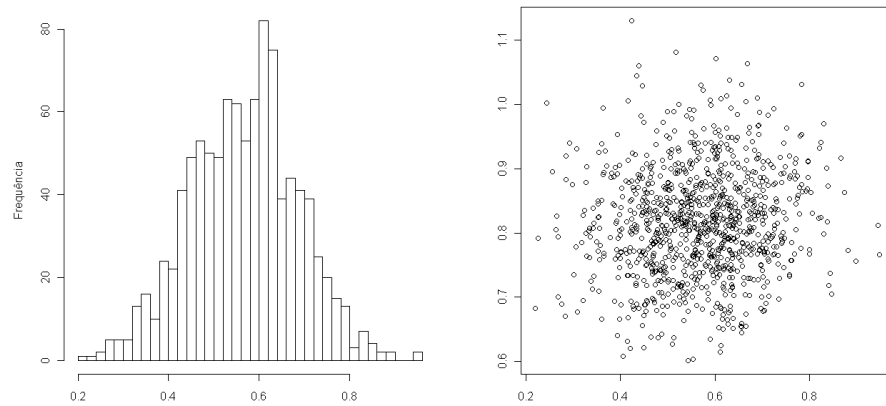


Figura 2.8: Distribuições amostrais de  $\hat{\mu}$  (painel esquerdo) e  $(\hat{\mu}, \hat{\sigma})$  (painel direito) para  $n = 50$

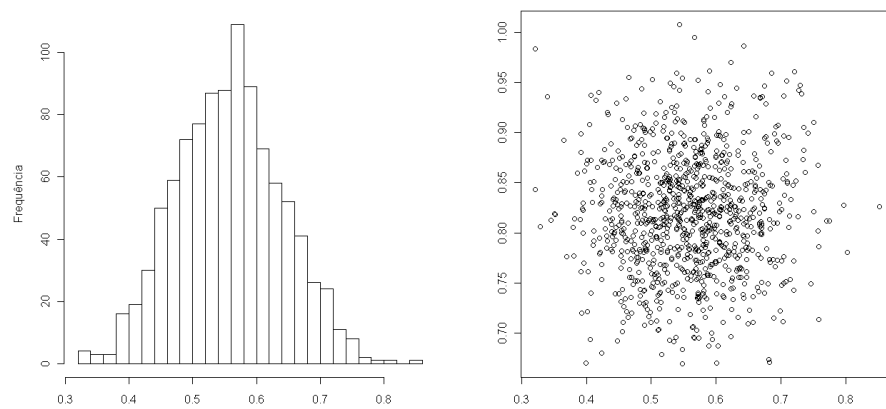


Figura 2.9: Distribuições amostrais de  $\hat{\mu}$  (painel esquerdo) e  $(\hat{\mu}, \hat{\sigma})$  (painel direito) para  $n = 100$

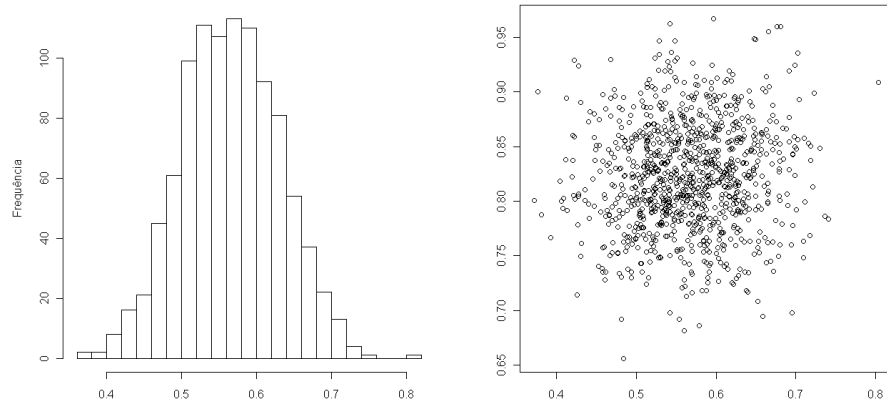


Figura 2.10: Distribuições amostrais de  $\hat{\mu}$  (painel esquerdo) e  $(\hat{\mu}, \hat{\sigma})$  (painel direito) para  $n = 150$

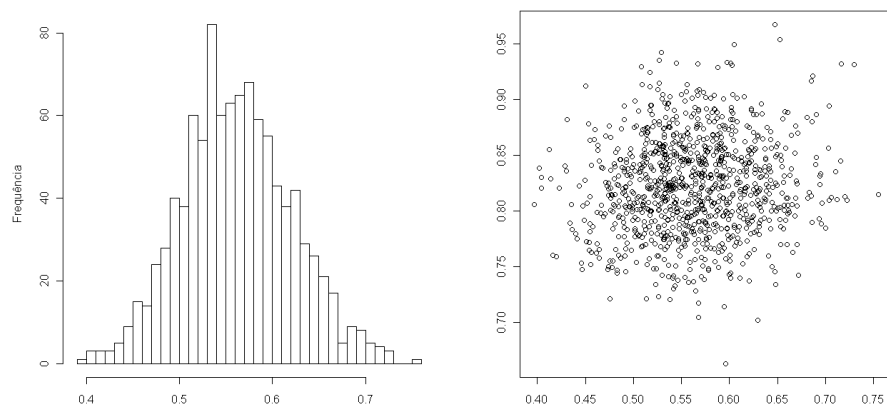


Figura 2.11: Distribuições amostrais de  $\hat{\mu}$  (painel esquerdo) e  $(\hat{\mu}, \hat{\sigma})$  (painel direito) para  $n = 200$

Os gráficos sugerem que as distribuições amostrais de  $\hat{\mu}$  e  $(\hat{\mu}, \hat{\sigma})$  são mais próximas da distribuição normal do que as de  $\hat{\alpha}$  e  $(\hat{\alpha}, \hat{\psi})$ . De fato, a reparametrização remove a singularidade da matriz de informação de Fisher, e a distribuição limite das novas estimativas é uma distribuição normal multivariada com variância assintótica  $\text{diag}(1/\sigma^2, \sigma^2/2, 1/6)$ ; veja Azzalini (1985) e Chiogna (2005).

### 2.3.2 Estimação bayesiana

Vamos agora abordar o problema de estimação no modelo normal assimétrico utilizando metodologia bayesiana. Essa metodologia, é mais adequada nos casos em que o estimador de máxima verossimilhança é infinito, como ilustrado em Branco et al. (2004) para o caso uniparamétrico. De modo geral, segundo Arellano-Valle and Azzalini (2008) problemas na estimação podem ocorrer a menos que se tenha uma priori fortemente informativa.

Liseo e Loperfido (2004) propõem uma aproximação para a priori de Jeffreys e um algoritmo MCMC para o caso uniparamétrico chegando a estimativas com boas propriedades frequentistas.

Nosso principal objetivo nessa subseção é verificar a existência desses possíveis problemas comparando as duas parametrizações, como feito na subseção anterior.

#### Especificação a priori

Iremos discutir primeiro a escolha da distribuição a priori para o parâmetro de assimetria  $\lambda$  no caso uniparamétrico. Uma possibilidade é utilizar a priori de Jeffreys, que é dada por

$$\begin{aligned} f^J(\lambda) &\propto \sqrt{I(\lambda)}, \\ &\propto \sqrt{\int_{-\infty}^{\infty} 2x^2 \phi(x) \frac{\phi^2(\lambda x)}{\Phi(\lambda x)} dx}. \end{aligned} \quad (2.9)$$

Esta possui as seguintes propriedades, verificadas por Liseo e Loperfido (2004):

- $f^J(\lambda)$  é simétrica em torno de  $\lambda = 0$  e decrescente em  $|\lambda|$ ;
- a cauda de  $f^J(\lambda)$  é da ordem  $O(\lambda^{-3/2})$ .

Bayes and Branco (2007) propõem uma aproximação para a essa priori baseada no método de Chaibud e Branco (2003). Tal aproximação é dada por:

$$f^J(\lambda) \approx t\left(0, \frac{\pi^2}{4}, \frac{1}{2}\right). \quad (2.10)$$

Está aproximação possui a mesma cauda da priori de Jeffreys, isto é, de ordem  $O(\lambda^{-3/2})$ . Uma outra possibilidade é considerar uma priori vaga para o parâmetro  $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$  que é limitado,  $|\delta| < 1$  e portanto, iremos obter uma priori própria. Neste caso, uma escolha natural seria considerar uma distribuição uniforme em  $[-1, 1]$ , que induz no espaço paramétrico de  $\lambda$  a distribuição t-Student especificada a seguir

$$\lambda \sim t(0, \frac{1}{2}, 2). \quad (2.11)$$

Considerando o caso de três parâmetros, Liseo e Loperfido (2004) utilizando o método de Berger e Bernardo (1992), obtiveram a seguinte distribuição a priori conjunta

$$f(\alpha, \psi, \lambda) \propto \frac{1}{\psi} g(\lambda), \quad (2.12)$$

em que,  $g(\lambda)$  é alguma função integrável do parâmetro  $\lambda$ . Os autores sugerem utilizar, por simplicidade, a priori de Jeffreys obtida para o caso uniparamétrico no lugar de  $g(\lambda)$ . Lembramos que a priori de Jeffreys para os parâmetros de localização e escala é proporcional a  $\frac{1}{\psi}$ . Considerando independência a priori entre os parâmetros usaremos a seguinte especificação a priori

$$p(\alpha, \psi, \lambda) \propto \frac{1}{\psi} p(\lambda), \quad (2.13)$$

em que  $p(\lambda)$  é uma densidade  $t(0, \varphi^2, d)$ . Quando  $d = \frac{1}{2}$  e  $\varphi^2 = \frac{\pi^2}{4}$  obtemos a aproximação da priori de Jeffreys e para  $d = 2$  e  $\varphi^2 = \frac{1}{2}$  obtemos a priori induzida por  $\delta \sim U(-1, 1)$ . Portanto,

$$p(\alpha, \psi, \lambda) \propto \frac{1}{\psi} \left( 1 + \frac{\lambda^2}{d\varphi^2} \right)^{-\frac{d+1}{2}}. \quad (2.14)$$

### Estrutura hierárquica e um algoritmo MCMC

Consideraremos a representação estocástica de Henze (1986) dada por

$$Y_i = \alpha + \psi \left( \delta H_i + \sqrt{1 - \delta^2} Z_i \right), \quad (2.15)$$

em que  $H_i \sim HN(0, 1)$  e  $Z_i \sim N(0, 1)$  são variáveis aleatórias independentes. A notação  $HN$  denota a distribuição normal positiva, ou seja,  $N(0, 1) I_{[0, \infty)}(h_i)$  mais conhecida como Half-Normal.

Condicionando em  $H_i$  obtemos

$$\begin{aligned} Y_i|H_i &\sim N(\alpha + \psi\delta H_i, \psi^2(1 - \delta^2)), \\ H_i &\sim HN(0, 1). \end{aligned} \quad (2.16)$$

Visando facilitar o trabalho computacional, trabalharemos com a reparametrização  $\eta = \psi\delta$  e  $\tau = \psi\sqrt{1 - \delta^2}$ , então:

$$\begin{aligned} Y_i|H_i &\sim N(\alpha + \eta H_i, \tau^2), \\ H_i &\sim HN(0, 1). \end{aligned} \quad (2.17)$$

A verossimilhança na nova parametrização é dada por

$$L(\alpha, \tau, \eta) = \frac{1}{\tau^n} \exp\left(-\frac{1}{2\tau^2} \sum_{i=1}^n (y_i - \alpha - \eta h_i)^2\right) \exp\left(-\frac{1}{2} \sum_{i=1}^n h_i^2\right). \quad (2.18)$$

A distribuição a priori  $\lambda \sim t(0, \varphi^2, d)$ , também pode ser reescrita em forma hierárquica da seguinte maneira

$$\begin{aligned} \lambda|w &\sim N(0, \frac{\varphi^2}{w}), \\ w &\sim \text{Gamma}(\frac{d}{2}, \frac{d}{2}). \end{aligned} \quad (2.19)$$

Dessa forma a priori dada em (2.14) pode ser reescrita como:

$$p(\lambda, \alpha, \psi, w) \propto \frac{1}{\psi^2} \exp\left(-\frac{1}{2} \frac{\lambda^2 w}{\varphi^2}\right) w^{\frac{d+1}{2}-1} \exp\left(-\frac{d}{2} w\right), \quad (2.20)$$

a qual induz a seguinte priori na nova reparametrização,

$$p(\alpha, \tau, \eta, w) \propto \frac{1}{\tau^2} \exp\left(-\frac{1}{2} \frac{\eta^2 w}{\tau^2 \varphi^2}\right) w^{\frac{d+1}{2}-1} \exp\left(-\frac{d}{2} w\right). \quad (2.21)$$

Finalmente, a distribuição a posteriori é dada por

$$f(\alpha, \tau, \eta, w|H) \propto p(\alpha, \tau, \eta, w) L(\alpha, \tau, \eta). \quad (2.22)$$

Essa reparametrização, também permite-nos obter formas conhecidas para todas as condicionais completas. Estas são representadas pelas expressões a seguir

$$\begin{aligned}
W|\alpha, \tau, \eta, H, Y &\sim \text{Gamma}\left(\frac{d+1}{2}, \frac{1}{2}\left(\frac{\eta^2}{\varphi^2\tau^2} + d\right)\right), \\
H_i|\alpha, \tau, \eta, w, Y &\sim N\left(\frac{(y_i - \alpha)\eta}{\tau^2 + \eta^2}, \frac{\tau^2}{\tau^2 + \eta^2}\right), I(h_i)_{[0, \infty)} \quad \forall i, \quad i = 1, \dots, n \\
\alpha|\tau, \eta, w, H, Y &\sim N\left(\frac{\sum_{i=1}^n (y_i - \eta h_i)}{n}, \frac{\tau^2}{n}\right), \\
\eta|\alpha, \tau, w, H, Y &\sim N\left(\frac{\sum_{i=1}^n (y_i - \alpha)h_i}{\frac{w}{\varphi^2} + \sum_{i=1}^n h_i^2}, \frac{\tau^2}{\frac{w}{\psi_y^2} + \sum_{i=1}^n h_i^2}\right), \\
\frac{1}{\tau^2}|\alpha, \eta, w, H, Y &\sim \text{Gamma}\left(\frac{n+1}{2}, \frac{1}{2}\left(\frac{w\eta^2}{\varphi^2} + \sum_{i=1}^n (y_i - \alpha - \eta h_i)^2\right)\right).
\end{aligned} \tag{2.23}$$

Os parâmetros originais podem ser recuperados da seguinte forma:

$$\psi = \sqrt{\tau^2 + \eta^2} \quad \text{e} \quad \lambda = \frac{\eta}{\tau}. \tag{2.24}$$

Já os parâmetros centrados são obtidos como nas equações (2.6). A partir das expressões (2.23) implementamos o amostrador de Gibbs em linguagem Ox<sup>TM1</sup>. Utilizando este algoritmo realizamos a estimação dos parâmetros diretos e centrados a partir de uma amostra de tamanho 50 da  $NA(0, 1, 1) \equiv NA_c(0, 564; 0, 826; 0, 137)$ .

Foram feitas 30000 simulações considerando a priori de Jeffreys. As 3000 primeiras observações foram descartadas e foram tomados valores de 27 em 27, obtendo amostras a posteriori de tamanho 1000. Este procedimento foi repetido mais 2 vezes com outros dois diferentes conjuntos de valores iniciais, para que pudéssemos aplicar o diagnóstico de Gelman-Rubin, veja Gelman et al. (2004). As estatísticas de Gelman-Rubin, calculadas a partir das amostras finais, foram todas iguais a 1. Isto indica que a convergência foi alcançada após 3000 iterações.

Nas Figuras podemos ver os correlogramas e os gráficos de valores simulados, respectivamente, para os parâmetros diretos e centrados.

A Tabela 2.1 apresenta os verdadeiros valores dos parâmetros, média, mediana, erros-padrão (EP), intervalos de credibilidade de 95% (IC (95%)) e HPD<sup>2</sup>'s também de 95% dos valores simulados. Os HPD's foram calculados utilizando o método descrito em Chen and Shao (1999). Já na Figura 2.15 são apresentadas as estimativas dos parâmetros com os respectivos intervalos

<sup>1</sup>Veja: <http://www.doornik.com/products.html>

<sup>2</sup>Sigla do inglês: *Highest posterior density*

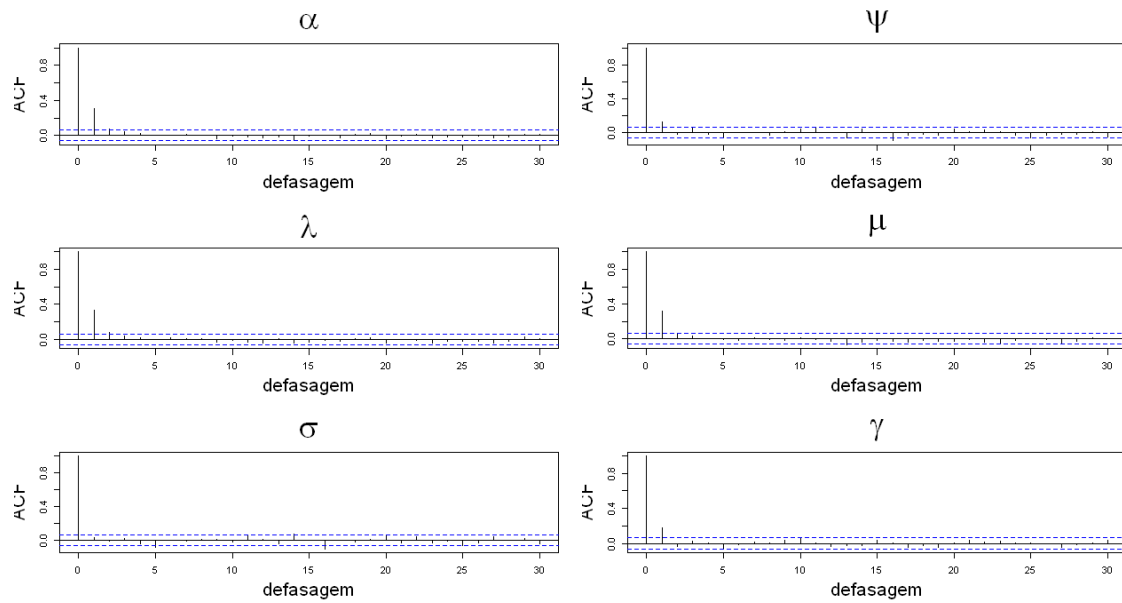


Figura 2.12: Correlogramas dos valores simulados para os parâmetros diretos e centrados

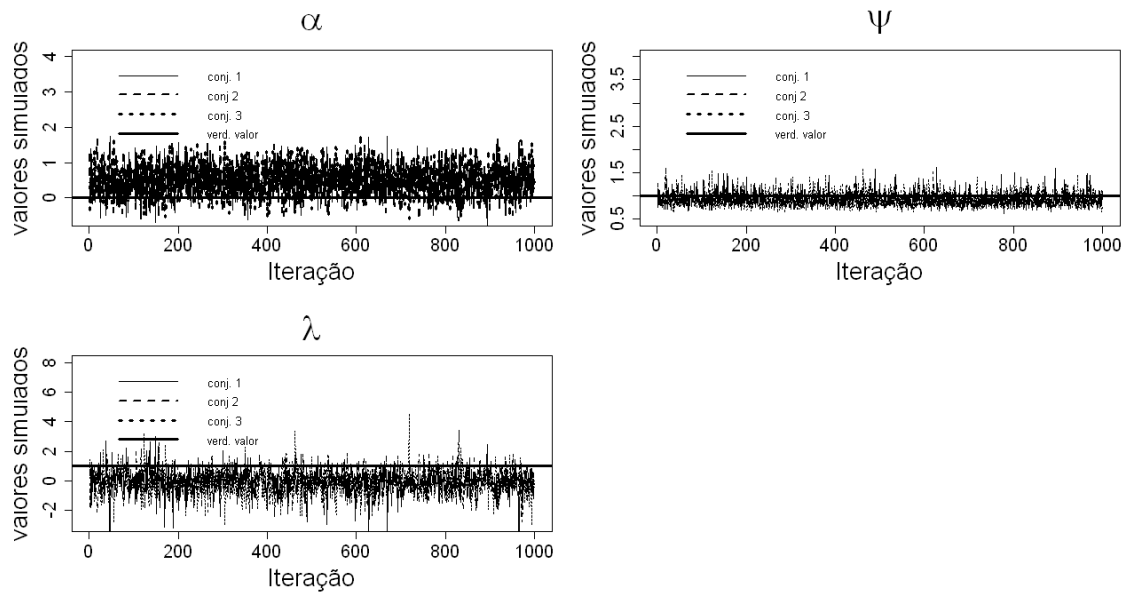


Figura 2.13: Gráfico de valores simulados para os parâmetros diretos

de credibilidade e HPD's. Podemos ver que, os erros-padrão dos parâmetros centrados são menores e também seus IC's e HPD's têm os menores comprimentos, principalmente, para os parâmetros de escala e forma.

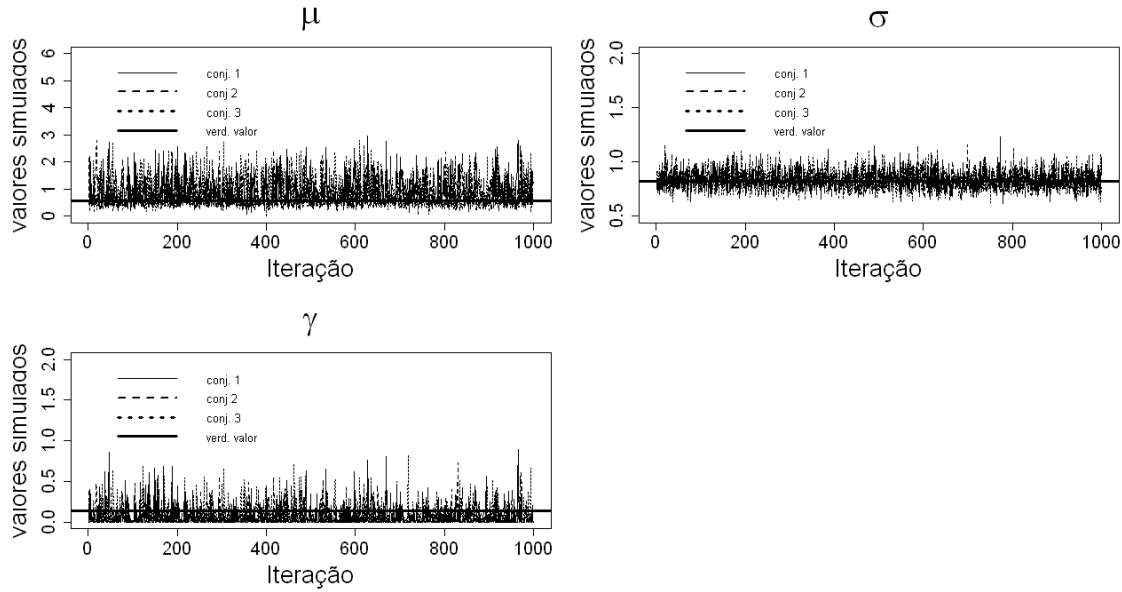


Figura 2.14: Gráfico de valores simulados para os parâmetros centrados

Tabela 2.1: Resumo a posteriori das estimativas dos parâmetros diretos e centrados

|    |           | Verd.valor | Média | Mediana | EP    | IC (95%) |       | HPD (95%) |       |
|----|-----------|------------|-------|---------|-------|----------|-------|-----------|-------|
| PD | $\alpha$  | 0,00       | 0,491 | 0,457   | 0,582 | -0,506   | 1,628 | -0,581    | 1,519 |
|    | $\psi$    | 1,00       | 1,057 | 1,008   | 0,205 | 0,778    | 1,552 | 0,771     | 1,502 |
|    | $\lambda$ | 1,00       | 0,275 | 0,229   | 1,170 | -1,906   | 2,665 | -1,912    | 2,648 |
| PC | $\mu$     | 0,564      | 0,983 | 0,722   | 0,595 | 0,405    | 2,541 | 0,347     | 2,332 |
|    | $\sigma$  | 0,826      | 0,897 | 0,888   | 0,096 | 0,734    | 1,115 | 0,715     | 1,092 |
|    | $\gamma$  | 0,137      | 0,148 | 0,064   | 0,187 | 0,000    | 0,657 | 0,000     | 0,572 |

## Estudo de simulação

Bayes (2005) considera a estimação do parâmetro de assimetria  $\lambda$  no modelo NA uniparamétrico e observa, através de estudos de simulação, que as estimativas obtidas considerando a média a posteriori sob priori de Jeffreys podem ser infinitas. Ele conclui que, em geral, a inferência bayesiana sob a priori uniforme gera resultados melhores do que a baseada na priori de Jeffreys.

Considerando o caso geral, conduzimos um pequeno estudo de simulação. Neste foram geradas 1000 amostras (sob as mesmas condições discutidas anteriormente) de tamanho  $n$  da  $NA(0, 1, 1) \equiv NA_c(0, 564; 0, 826; 0, 137)$ , com  $n = 50, 100, 150, 200$ . Para cada uma delas calculamos a média a posteriori, considerando as duas especificações a priori discutidas anteriormente, na PD e PC. Também avaliamos o viés e o erro quadrático médio empíricos definidos genericamente como

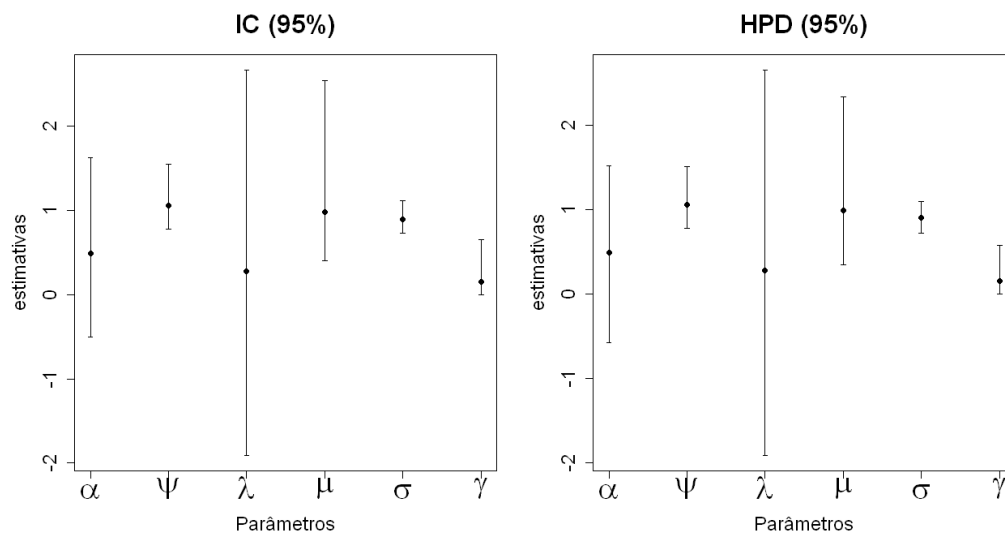


Figura 2.15: Intervalos de credibilidade e HPD's das estimativas dos parâmetros diretos e centrados

$$\text{Vies}(\theta) = \frac{\sum_{i=1}^m (\theta - \tilde{\theta})}{m} \quad \text{e} \quad \text{EQM}(\theta) = \frac{\sum_{i=1}^m (\theta - \tilde{\theta})^2}{m} \quad (2.25)$$

em que  $\theta$  é o parâmetro de interesse,  $\tilde{\theta}$  é a estimativa e  $m$  representa o número de amostras simuladas. As Figuras 2.16 e 2.17 apresentam os histogramas das médias a posteriori de cada um dos parâmetros diretos e centrados, para cada uma das especificações a priori, utilizando amostras de tamanho 50. Observando estas, podemos notar que a falta de precisão dos estimadores de Bayes dos parâmetros diretos não está relacionada com uma particular amostra, independentemente, da priori escolhida. Observando a Figura 2.17 (c), podemos notar que, assim como no caso uniparamétrico, as estimativas do parâmetro de assimetria  $\lambda$  apresentaram alguns valores (em torno de 5%) bastante discrepantes. Já as estimativas dos parâmetros centrados, estão mais concentradas em torno dos verdadeiros valores, sob ambas especificações a priori.

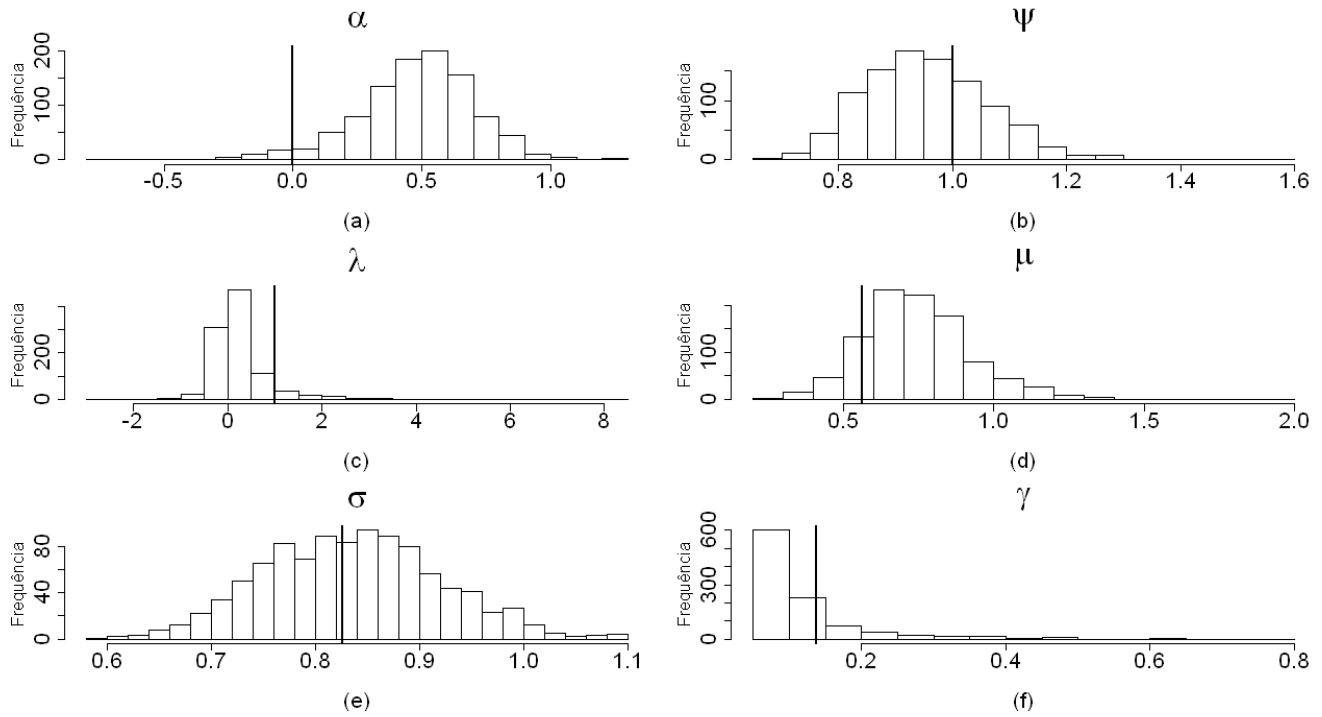


Figura 2.16: Distribuições amostrais da média a posteriori sob a priori Uniforme (as linhas verticais indicam os verdadeiros valores dos parâmetros)

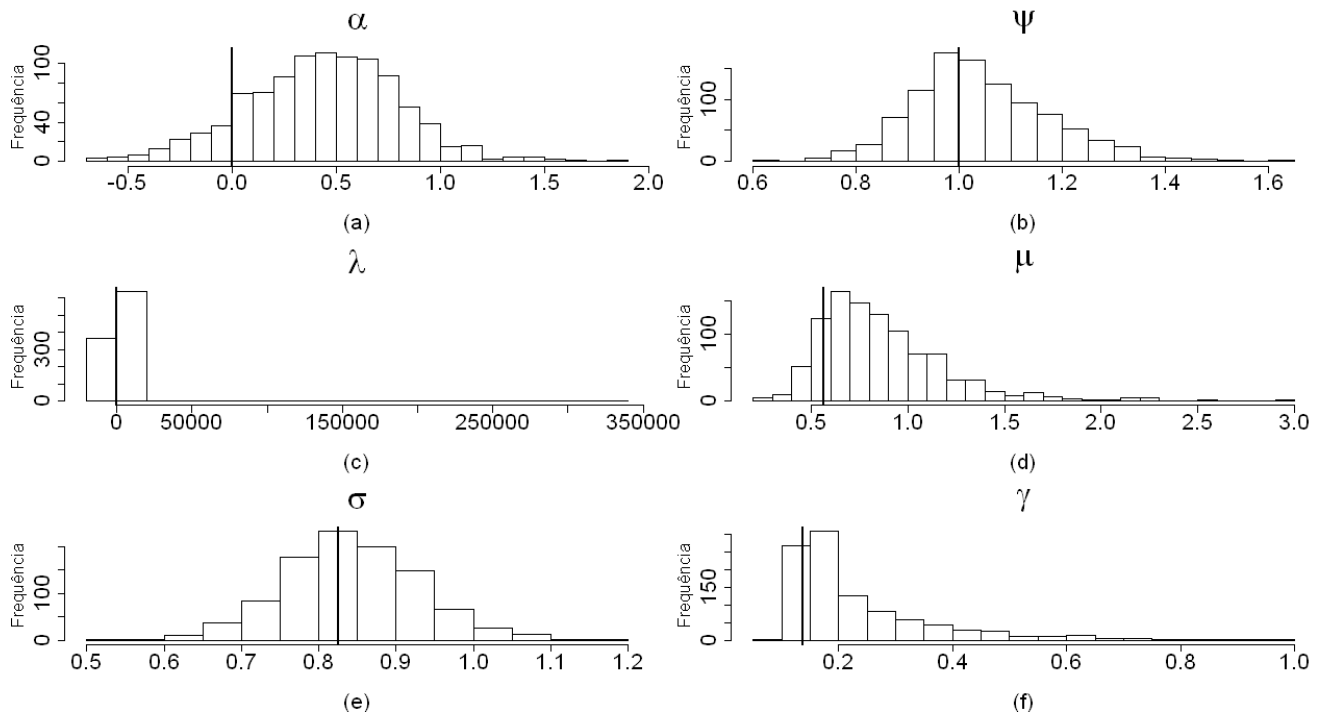


Figura 2.17: Distribuições amostrais da média a posteriori sob a priori Jeffreys (as linhas verticais indicam os verdadeiros valores dos parâmetros)

O viés das estimativas assim como o EQM podem ser vistos nas Tabelas 2.2 e 2.3, respectivamente. Observamos que as estimativas dos parâmetros centrados possuem menor viés e menor EQM, principalmente o coeficiente de assimetria  $\gamma$  em comparação ao parâmetro de assimetria  $\lambda$ . Isso ocorre com ambas as prioris e independentemente do tamanho da amostra.

Notamos também, um valor muito discrepante no viés e no EQM das estimativas de  $\lambda$  considerando a priori de Jeffreys e  $n = 50$ . Isso ocorreu, como já havíamos citado, por conta dos altos valores que foram gerados com o tamanho de amostra 50. O mesmo não ocorreu com amostras de tamanhos maiores.

Quando comparamos somente os parâmetros centrados com relação as prioris, notamos que a priori uniforme apresenta o menor viés e EQM quando o tamanho da amostra é pequeno ( $n = 50$ ). Nas amostras maiores parece não haver diferenças significativas.

Tabela 2.2: Viés das estimativas no estudo de simulação considerando as duas prioris

|    | $n$       | Priori Uniforme |        |        |        | Priori Jeffreys |        |        |        |
|----|-----------|-----------------|--------|--------|--------|-----------------|--------|--------|--------|
|    |           | 50              | 100    | 150    | 200    | 50              | 100    | 150    | 200    |
| PD | $\alpha$  | -0,480          | -0,454 | -0,443 | -0,418 | -0,433          | -0,399 | -0,383 | -0,364 |
|    | $\psi$    | 0,044           | 0,053  | 0,067  | 0,069  | -0,045          | 0,001  | 0,020  | 0,037  |
|    | $\lambda$ | 0,757           | 0,739  | 0,733  | 0,705  | -517,308        | 0,375  | 0,592  | 0,577  |
| PC | $\mu$     | -0,185          | -0,141 | -0,117 | -0,085 | -0,304          | -0,195 | -0,145 | -0,104 |
|    | $\sigma$  | -0,010          | -0,009 | -0,002 | -0,003 | -0,018          | -0,001 | -0,007 | -0,001 |
|    | $\gamma$  | 0,018           | 0,032  | 0,041  | 0,048  | -0,099          | -0,037 | -0,011 | 0,006  |

Tabela 2.3: EQM das estimativas no estudo de simulação considerando as duas prioris

|    | $n$       | Priori Uniforme |       |       |       | Priori Jeffreys |        |       |       |
|----|-----------|-----------------|-------|-------|-------|-----------------|--------|-------|-------|
|    |           | 50              | 100   | 150   | 200   | 50              | 100    | 150   | 200   |
| PD | $\alpha$  | 0,283           | 0,257 | 0,246 | 0,219 | 0,324           | 0,262  | 0,233 | 0,206 |
|    | $\psi$    | 0,014           | 0,011 | 0,011 | 0,009 | 0,021           | 0,011  | 0,009 | 0,009 |
|    | $\lambda$ | 1,030           | 0,887 | 0,791 | 0,687 | 139183740       | 18,261 | 0,876 | 0,704 |
| PC | $\mu$     | 0,070           | 0,047 | 0,038 | 0,027 | 0,202           | 0,102  | 0,071 | 0,043 |
|    | $\sigma$  | 0,007           | 0,004 | 0,002 | 0,002 | 0,008           | 0,004  | 0,002 | 0,002 |
|    | $\gamma$  | 0,006           | 0,009 | 0,008 | 0,007 | 0,029           | 0,015  | 0,011 | 0,009 |

Em resumo temos:

- As estimativas mais precisas foram obtidas considerando a parametrização centrada.
- As estimativas dos parâmetros centrados parecem ser menos sensíveis a escolha da priori (Uniforme ou Jeffreys)

De modo geral concluímos que, assim como na abordagem frequentista, a utilização da parametrização centrada na estimação bayesiana é bem mais atrativa do que a parametrização direta.

# Capítulo 3

## Modelo de Grupos Múltiplos com Distribuições Normais Assimétricas Centralizadas

### 3.1 Introdução

No presente capítulo desenvolvemos o modelo probito de 2 parâmetros para grupos múltiplos, com distribuições normais assimétricas centralizadas para os traços latentes. Desenvolvemos também algoritmos MCMC para estimação dos parâmetros, utilizando o esquema de dados aumentados, veja Tanner and Wong (1987). Nesses algoritmos, usamos tanto a densidade original da distribuição normal assimétrica, quanto uma representação estocástica da mesma. O presente capítulo é uma extensão dos trabalhos de Azevedo et.al (2011) e Azevedo et.al (2010).

### 3.2 Definição do Modelo

Considere a situação onde temos  $K$  grupos compostos de  $n_k$  indivíduos cada, num total de  $n = \sum_{k=1}^K n_k$ . Estes serão submetidos a testes com  $I_k$  itens cada. Os testes possuem itens comuns e a estrutura pode ser reconhecida como um delineamento em blocos incompletos, veja Montgomery (2004). Estes itens são escolhidos de um conjunto total de  $I$  itens, com  $\sum_{k=1}^K I_k \geq I$ . Assumiremos que cada grupo possui um número razoável de indivíduos. Dora-vante, utilizaremos o termo grupo para nos referir de modo indistinto à uma população ou à uma amostra dela extraída. Definamos, então, as seguintes notações:  $\theta_{jk}$  é o traço latente do indivíduo  $j$  ( $j = 1, \dots, n_k$ ) pertencente ao grupo  $k$  ( $k = 1, \dots, K$ ),  $\boldsymbol{\theta}_{.k} = (\theta_{j1}, \dots, \theta_{jK})^\top$  é o vetor dos

traços latentes dos indivíduos do grupo  $k$ ,  $\boldsymbol{\theta}_{..} = (\boldsymbol{\theta}_{.1}, \dots, \boldsymbol{\theta}_{.K})^\top$  é o vetor de todos os traços latentes;  $Y_{ijk}$  é a resposta do indivíduo  $j$ , do grupo  $k$  no item  $i$  ( $i = 1, \dots, I$ ),  $\mathbf{Y}_{.jk} = (Y_{1jk}, \dots, Y_{Ijk})^\top$  será o vetor de repostas em todos os itens do indivíduo  $j$  do grupo  $k$ ,  $\mathbf{Y}_{...} = (\mathbf{Y}_{.1}^\top, \dots, \mathbf{Y}_{.n_k}^\top)^\top$  é o vetor com os conjuntos de repostas dos indivíduos e  $(y_{ijk}, \mathbf{y}_{.jk}^\top, \mathbf{y}_{..k}^\top, \mathbf{y}_{...}^\top)^\top$  são os respectivos valores observados;  $\boldsymbol{\zeta}_i$  é o vetor de parâmetros do item  $i$ ,  $\boldsymbol{\zeta} = (\boldsymbol{\zeta}_1^\top, \dots, \boldsymbol{\zeta}_I^\top)^\top$  é o vetor com todos os parâmetros de todos os itens,  $\boldsymbol{\eta}_{\theta_k}$  é o vetor de parâmetros populacionais do grupo  $k$  e  $\boldsymbol{\eta}_{\theta} = (\boldsymbol{\eta}_{\theta_1}^\top, \dots, \boldsymbol{\eta}_{\theta_K}^\top)^\top$  é o vetor com os parâmetros populacionais de todos os grupos.

Desta forma, definimos o modelo probito de 2 parâmetros assimétrico para grupos múltiplos (MP2AGM) da seguinte forma

### Modelo de resposta ao item (parte funcional)

$$\begin{aligned} Y_{ijk} | \theta_{jk}, \boldsymbol{\zeta}_i &\sim \text{Bernoulli}(P_{ijk}), \\ P_{ijk} = P(Y_{ijk} | \theta_{jk}, \boldsymbol{\zeta}_i) &= \Phi(a_i \theta_{jk} - b_i). \end{aligned} \quad (3.1)$$

### Distribuição Latente (parte estrutural)

$$\theta_{jk} | \boldsymbol{\eta}_{\theta_k} \sim \text{NA}_c(\mu_{\theta_k}, \sigma_{\theta_k}, \gamma_{\theta_k}), \quad (3.2)$$

em que  $\boldsymbol{\zeta}_i = (a_i, b_i)^\top$  e  $\boldsymbol{\eta}_{\theta_k} = (\mu_{\theta_k}, \sigma_{\theta_k}, \gamma_{\theta_k})^\top$ . Na equação (3.1) temos que,  $a_i$  é o parâmetro de discriminação e  $b_i = a_i b_i^*$ , onde  $b_i^*$  é o parâmetro de dificuldade original, veja Baker and Kim (2004), por exemplo, para mais detalhes e interpretações.

Duas questões muito importantes estão relacionadas com a identificabilidade do modelo e a definição da métrica. Na verdade estas questões estão intrinsecamente ligadas, sendo que, a identificabilidade é uma condição necessária e suficiente para a definição da métrica. A comparação das estimativas dos diferentes grupos é muito importante na prática, veja Andrade et.al (2000). Esta comparação só é possível se estabelecermos uma métrica de modo que, a escala das estimativas de todos os parâmetros seja a mesma. No modelo de grupos múltiplos isso é feito da seguinte forma. Inicialmente, fazemos com que os testes aplicados aos diferentes grupos possuam alguns itens comuns. Os itens comuns são necessários para criar uma estrutura de ligação entre os grupos. Depois, a média e a variância da escala é identificada fixando-se a métrica de um grupo de referência. Para isso, devemos fixar os parâmetros populacionais da distribuição dos traços latentes do grupo de referência, veja Book and Zimowski (1997). Portanto, ao fixarmos a escala do grupo de referência, estaremos fixando também a escala dos demais grupos, uma vez que estes estão ligados através dos itens comuns.

A Figura 3.1 mostra um exemplo desse esquema, considerando 3 grupos respondendo a

três provas parcialmente distintas com 30 itens cada selecionados de um total de 40 itens. O Grupo 1 é o de referência e, aos traços latentes dos indivíduos desse grupo foram atribuídas distribuições  $D_1$  cujos parâmetros foram fixados em  $(0, 1)$ .

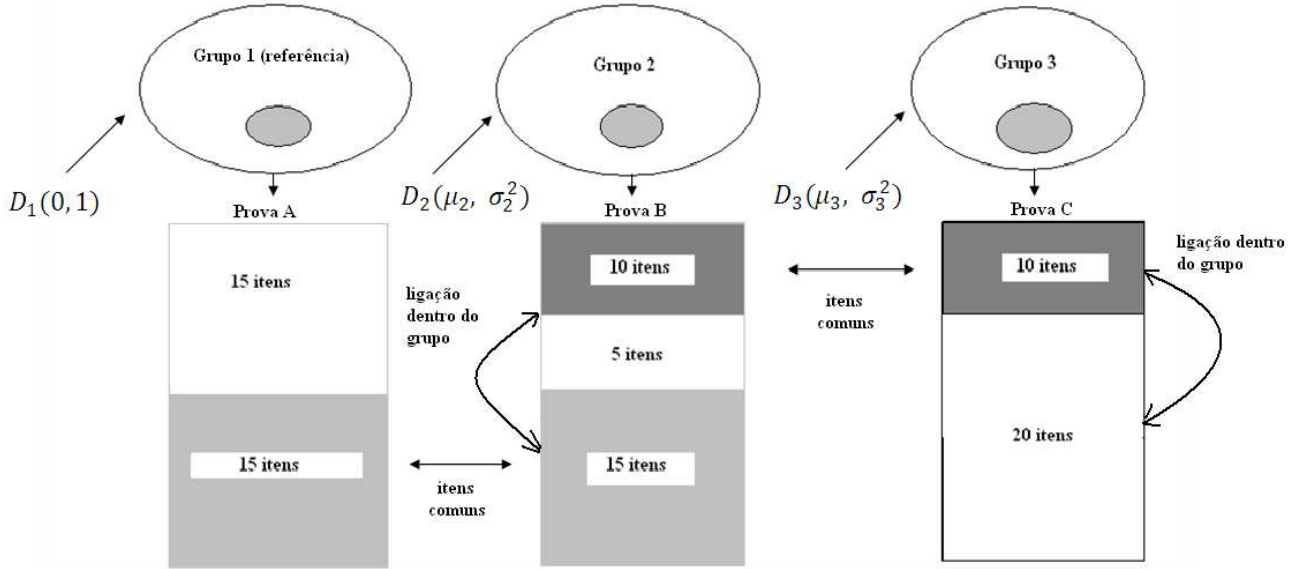


Figura 3.1: Esquema de Grupo de referência e itens comuns

Seguindo o desenvolvimento de Azevedo et.al (2010) vamos provar que o modelo (3.1) é identificável. Vamos considerar, inicialmente, um único grupo. Tomando  $K = 1$  na equação (3.1) temos:

$$P(Y_{ij1}|\theta_{j1}, \zeta_i) = \Phi(a_i\theta_{j1} - b_i). \quad (3.3)$$

Considere agora a seguinte pressuposição:

$$\theta_{j1} \sim \text{NA}(\alpha_{\theta_1}, \psi_{\theta_1}, \lambda_{\theta_1}), \quad (3.4)$$

ou seja, seguindo a notação do capítulo 2,  $\theta_{j1}$  segue distribuição normal assimétrica na parametrização direta. Sob (3.4) e tomando  $\alpha_{\theta_1} = 0$  e  $\psi_{\theta_1} = 1$  temos que  $\mathbb{E}(\theta_{j1}|\lambda_{\theta_1}) = \delta r = h(\lambda_{\theta_1})$  e  $\text{Var}(\theta_{j1}|\lambda_{\theta_1}) = 1 - r^2\delta^2 = g(\lambda_{\theta_1})$ . Ou seja, considerar a distribuição normal assimétrica padrão não garante a identificabilidade do modelo, uma vez que, a média e a variância dos traços latentes não estão fixas. Desta maneira, se não forem feitas restrições aos parâmetros dos itens, o modelo (3.3) não estará identificado. Isto acontece porque:

$$\begin{aligned}
 P(Y_{ij1}|\theta_{j1}, \zeta_i) &= \Phi(a_i(\theta_{j1} - b_i)) = \Phi\left(\frac{a_i}{\alpha}(\alpha\theta_{j1} - \alpha b_i)\right), \\
 &= \Phi\left(\frac{a_i}{\alpha}(\alpha\theta_{j1} + \beta - \alpha b_i - \beta)\right), \\
 &= \Phi(a_i^*(\theta_{j1}^* - b_i^*)) = P(Y_{ij1} = 1|\theta_{j1}^*, \zeta_i^*),
 \end{aligned} \tag{3.5}$$

em que,  $\alpha, \beta \in \mathbb{R}$  são constantes. Desta forma, diferentes conjuntos de valores dos parâmetros podem levar a mesma verossimilhança. Entretanto, assumindo que:

$$\theta_{j1} \sim \text{NA}_c(0, 1, \gamma_{\theta_1}), \tag{3.6}$$

teremos que  $\mathbb{E}(\theta_{j1}|\lambda_{\theta_1}) = 0$  e  $\text{Var}(\theta_{j1}|\lambda_{\theta_1}) = 1$ . Isto garante não somente que o modelo (3.3) está identificado, mas também que todas as estimativas estarão na métrica  $(0, 1)$ . Assim, o uso da normal assimétrica centralizada garante a identificabilidade do modelo (3.3). Note que, não é mais possível fazer transformações como as que foram feitas em (3.5). Portanto, é suficiente assumir (3.6) para identificar o modelo (3.3) com a pressuposição de distribuição normal assimétrica para os traços latentes.

Na presença de  $K$  grupos, além de considerar o esquema de grupo de referência com itens comuns descrito anteriormente, devemos assumir ainda as seguintes restrições:

$$\begin{aligned}
 \theta_{j1} &\sim \text{NA}_c(0, 1, \gamma_{\theta_1}), \\
 \theta_{jk} &\sim \text{NA}_c(\mu_{\theta_k}, \sigma_{\theta_k}, \gamma_{\theta_k}), \quad k = 2, \dots, K,
 \end{aligned} \tag{3.7}$$

em que,  $\theta_{j1}$  representa os traços latentes dos indivíduos do grupo de referência. Analogamente ao modelo de um grupo, note que não é possível considerar transformações do tipo (3.5) nos parâmetros do grupo de referência e, conseqüentemente, também não será possível fazê-lo nos parâmetros dos demais grupos, uma vez que, a escala desses estará ligada através dos itens comuns. Isso garante que o modelo (3.1) está identificado e que as estimativas dos seus parâmetros estarão na métrica  $(0, 1)$ .

### 3.3 Estimação bayesiana e algoritmos MCMC

A inferência bayesiana é baseada na distribuição a posteriori dos parâmetros, que é proporcional ao produto da verossimilhança e distribuição a priori. Considerando os modelos de resposta ao item, geralmente, não é possível obter de forma explícita as distribuições a posteriori de interesse. Entretanto, os algoritmos MCMC fornecem aproximações empíricas dessas distribuições, sob algumas condições. Discussões detalhadas desses métodos podem ser encontradas em Gamerman and Lopes (2006), Gelman et.al (2004) e Robert and Casella (2005).

De modo a facilitar a implementação dos algoritmos MCMC, utilizaremos neste trabalho o esquema de dados aumentados, veja Tanner and Wong (1987). Esta estrutura permite obter distribuições condicionais completas com forma funcional mais simples; vide Albert (1992). Além disso, ao se considerar um esquema de dados aumentados, introduzimos um modelo mais geral. Uma outra vantagem é que, esta abordagem permite construir mecanismos para verificação de qualidade de ajuste do modelo de modo mais direto, veja Fox (2001). Denotando por  $Z_{ijk}$  o conjunto de dados aumentados, temos que o modelo (3.1) pode ser representado da seguinte forma:

$$Y_{ijk} = \mathbb{1}_{(Z_{ijk} \geq 0)} \quad \text{em que,} \quad Z_{ijk} \sim N(a_i \theta_{jk} - b_i, 1).$$

Podemos verificar que,

$$Z_{ijk} | (\theta_{jk}, \zeta_i, y_{ijk}) = \begin{cases} N(a_i \theta_{jk} - b_i, 1) \mathbb{1}_{(Z_{ijk} \geq 0)}, & \text{se } Y_{ijk} = 1, \\ N(a_i \theta_{jk} - b_i, 1) \mathbb{1}_{(Z_{ijk} < 0)}, & \text{se } Y_{ijk} = 0, \end{cases} \quad (3.8)$$

Também não é difícil verificar que  $Y_{ijk} = \mathbb{1}_{(Z_{ijk} \geq 0)}$  é equivalente ao modelo (3.1). Dessa forma, as verossimilhanças dos dados aumentados condicionado aos dados observados e aquela oriunda somente dos dados observados são equivalentes.

Para descrever as possíveis omissões de dados geradas por eventos não controlados, tais como, não-resposta ou erros de codificação, definimos o seguinte conjunto de variáveis indicadoras

$$V_{ijk} = \begin{cases} 1, & \text{se para o item } i \text{ foi observada resposta do indivíduo } j \text{ do grupo } k, \\ 0, & \text{caso contrário.} \end{cases} \quad (3.9)$$

Assumiremos que os dados omissos são gerados por mecanismos não informativos MAR, veja Little and Rubin (2002). De modo similar às variáveis aumentadas, as variáveis indicadoras de resposta são condicionalmente independentes, entre indivíduos, itens e populações, dado um

vetor de parâmetros  $\delta$  que não está relacionado com  $(\theta_{..}, \zeta, \eta_{\theta})$ . Desta forma a verossimilhança aumentada é dada por:

$$\begin{aligned}
 L(\theta_{..}, \zeta, \eta_{\theta}, \delta | \mathbf{z}_{...}, \mathbf{y}_{...}, \mathbf{v}_{...}) &\propto p(\mathbf{z}_{...} | \mathbf{y}_{...}, \mathbf{v}_{...}, \theta_{..}, \zeta, \eta_{\theta}, \delta) p(\mathbf{v}_{...} | \theta_{..}, \zeta, \delta) \\
 &= p(\mathbf{z}_{...} | \mathbf{y}_{...}, \theta_{..}, \zeta) p(\mathbf{v}_{...} | \delta) \\
 &\propto p(\mathbf{z}_{...} | \mathbf{y}_{...}, \theta_{..}, \zeta) \\
 &\propto \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} \exp \{ -0,5(z_{ijk} - a_i \theta_{jk} + b_i)^2 \} \mathbb{1}_{(Z_{ijk}, y_{ijk})}, \quad (3.11)
 \end{aligned}$$

em que,  $I_{jk}$  é o conjunto de itens apresentados ao indivíduo  $j$  do grupo  $k$  e  $\mathbb{1}_{(Z_{ijk}, y_{ijk})}$  representa a função indicadora  $\mathbb{1}_{(Z_{ijk} < 0, y_{ijk} = 0)} + \mathbb{1}_{(Z_{ijk} \geq 0, y_{ijk} = 1)}$ . Note que (3.10) segue do fato de que as variáveis indicadoras são regidas por um processo MAR e (3.11) segue de (3.8).

Nas próximas subseções apresentaremos dois algoritmos MCMC. No primeiro, utilizaremos a densidade (2.4) para quantificar nossa incerteza a priori a respeito dos traços latentes. No segundo, faremos o mesmo, mas dessa vez utilizaremos a representação estocástica da normal assimétrica.

### 3.3.1 MCMC usando a densidade original da Normal assimétrica centralizada

Inicialmente, vamos discutir a escolha de prioris para os parâmetros. Nossas escolhas serão baseadas na obtenção de posteriors próprias e também na pertinência à família de modelos de conjugação condicional, Gelman (2006). Assumiremos a seguinte forma geral da priori conjunta dos parâmetros:

$$p(\theta_{..}, \zeta, \eta_{\theta} | \eta_{\zeta}, \eta_{\eta}) = \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(\theta_{jk} | \eta_{\theta_k}) \right\} \left\{ \prod_{i=1}^I p(\zeta_i | \eta_{\zeta}) \right\} \left\{ \prod_{k=1}^K p(\eta_{\theta_k} | \eta_{\eta}) \right\}, \quad (3.12)$$

em que,  $\eta_{\zeta}$  e  $\eta_{\eta}$  são hiperparâmetros convenientes associados à  $\zeta$  e  $\eta_{\theta}$ , respectivamente. Estamos, também, admitindo independência a priori entre os itens e os parâmetros de diferentes populações, o que faz sentido na prática.

A distribuição a priori dos traços latentes é dada por (2.4). Ou seja,

$$p(\theta_{jk}) \propto \omega_{\theta_k}^{-1} \phi(\omega_{\theta_k}^{-1}(\theta_{jk} - \xi_{\theta_k})) \Phi[\lambda_{\theta_k}(\omega_{\theta_k}^{-1}(\theta_{jk} - \xi_{\theta_k}))], \quad (3.13)$$

em que,

$$\begin{aligned}\xi_{\theta_k} &= \mu_{\theta_k} - \sigma_{\theta_k} \gamma_{\theta_k}^{1/3} s, \\ \omega_{\theta_k} &= \sigma_{\theta_k} \sqrt{1 + \gamma_{\theta_k}^{2/3} s^2}, \\ \lambda_{\theta_k} &= \frac{\gamma_{\theta_k}^{1/3} s}{\sqrt{r^2 + s^2 \gamma_{\theta_k}^{2/3} (r^2 - 1)}} \\ s &= \left(\frac{2}{4 - \pi}\right)^{1/3}.\end{aligned}$$

Para os parâmetros dos itens, uma escolha razoável é a distribuição normal multivariada (com  $a_i > 0$ ), veja Azevedo (2008). Ou seja,

$$p(\boldsymbol{\zeta}_i) \propto \exp \left[ -0,5(\boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta)^\top \boldsymbol{\Psi}_\zeta^{-1} (\boldsymbol{\zeta}_i - \boldsymbol{\mu}_\zeta) \right] \mathbb{1}_{(a_i > 0)}. \quad (3.14)$$

Para os parâmetros populacionais, consideraremos dois conjuntos de prioris. Primeiro seguiremos utilizando a mesma priori considerada no Capítulo 2 que é dada por:

$$p(\xi_{\theta_k}, \omega_{\theta_k}, \lambda_{\theta_k}) = p(\xi_{\theta_k}, \omega_{\theta_k}) p(\lambda_{\theta_k}) \propto \frac{1}{\omega_{\theta_k}} \left( 1 + \frac{\lambda_{\theta_k}^2}{d\varphi^2} \right)^{-\frac{d+1}{2}}, \quad (3.15)$$

em que,

$$p(\xi_{\theta_k}, \omega_{\theta_k}) \propto \frac{1}{\omega_{\theta_k}}, \quad (3.16)$$

$$p(\lambda_{\theta_k}) \propto \left( 1 + \frac{\lambda_{\theta_k}^2}{d\varphi^2} \right)^{-\frac{d+1}{2}}. \quad (3.17)$$

A partir de (3.15) podemos obter a aproximação da priori de Jeffreys (fazendo,  $d = \frac{1}{2}$  e  $\varphi^2 = \frac{\pi^2}{4}$ ) e a priori induzida por  $\delta_{\theta_k} \sim U(-1, 1)$  (fazendo,  $d = 2$  e  $\varphi^2 = \frac{1}{2}$ ), conforme discutido na seção 2.3.2 do Capítulo 2. A segunda escolha é baseada na seguinte priori conjunta:

$$p(\xi_{\theta_k}, \omega_{\theta_k}, \lambda_{\theta_k}) = p(\xi_{\theta_k}) p(\omega_{\theta_k}) p(\lambda_{\theta_k}), \quad (3.18)$$

considerando,

$$\xi_{\theta_k} \sim N(\mu_{\xi_\theta}, \sigma_{\xi_\theta}^2), \quad (3.19)$$

$$\omega_{\theta_k}^2 \sim \text{Gama-Inv.}(\alpha_{\omega_\theta}, \beta_{\omega_\theta}), \quad (3.20)$$

$$\lambda_{\theta_k} \sim N(\mu_{\lambda_\theta}, \sigma_{\lambda_\theta}^2). \quad (3.21)$$

Desta forma, considerando a verossimilhança aumentada (3.11) e as prioris (3.14) e (3.15), temos que a distribuição a posteriori de  $(\mathbf{Z}_{...}, \boldsymbol{\theta}_{..}, \boldsymbol{\zeta}, \boldsymbol{\eta}_\theta)$  é dada por:

$$\begin{aligned} p(\mathbf{Z}_{...}, \boldsymbol{\theta}_{..}, \boldsymbol{\zeta}, \boldsymbol{\eta}_\theta | \mathbf{y}_{...}, \boldsymbol{\eta}_\zeta, \boldsymbol{\eta}_\eta) &\propto p(\mathbf{Z}_{...} | \boldsymbol{\theta}_{..}, \boldsymbol{\zeta}, \mathbf{y}_{...}) p(\boldsymbol{\theta}_{..} | \boldsymbol{\eta}_\theta) p(\boldsymbol{\zeta} | \boldsymbol{\eta}_\zeta) p(\boldsymbol{\eta}_\theta | \boldsymbol{\eta}_\eta) \\ &= \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} p(z_{ijk} | \theta_{jk}, \zeta_i, y_{ijk}) \right\} \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(\theta_{jk} | \boldsymbol{\eta}_{\theta_k}) \right\} \left\{ \prod_{i=1}^I p(\zeta_i) \right\} \\ &\times \left\{ \prod_{k=1}^K p(\xi_{\theta_k}, \omega_{\theta_k}, \lambda_{\theta_k}) \right\} \\ &\propto \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} \exp \{ -0,5(z_{ijk} - a_i \theta_{jk} + b_i)^2 \} \mathbb{1}_{(Z_{ijk}, y_{ijk})} \right\} \\ &\times \left\{ \prod_{k=1}^K \omega_{\theta_k}^{-n_k} \prod_{j=1}^{n_k} \phi \left( \frac{\theta_{jk} - \xi_{\theta_k}}{\omega_{\theta_k}} \right) \Phi \left[ \left( \frac{\theta_{jk} - \xi_{\theta_k}}{\omega_{\theta_k}} \right) \right] \right\} \\ &\times \left\{ \prod_{i=1}^I \exp \left[ -0,5(\zeta_i - \boldsymbol{\mu}_\zeta)^\top \boldsymbol{\Psi}_\zeta^{-1} (\zeta_i - \boldsymbol{\mu}_\zeta) \right] \mathbb{1}_{(a_i > 0)} \right\} \\ &\times \left\{ \prod_{k=1}^K \frac{1}{\omega_{\theta_k}} \left( 1 + \frac{\lambda_{\theta_k}^2}{d\varphi^2} \right)^{-\frac{d+1}{2}} \right\}. \end{aligned} \quad (3.22)$$

Se ao invés da priori (3.15), considerarmos as prioris (3.19), (3.20) e (3.21) obteremos a seguinte posteriori:

$$\begin{aligned}
p(\mathbf{Z}_{\dots}, \boldsymbol{\theta}_{\dots}, \boldsymbol{\zeta}, \boldsymbol{\eta}_{\theta} | \mathbf{y}_{\dots}, \boldsymbol{\eta}_{\zeta}, \boldsymbol{\eta}_{\eta}) &\propto p(\mathbf{Z}_{\dots} | \boldsymbol{\theta}_{\dots}, \boldsymbol{\zeta}, \mathbf{y}_{\dots}) p(\boldsymbol{\theta}_{\dots} | \boldsymbol{\eta}_{\theta}) p(\boldsymbol{\zeta} | \boldsymbol{\eta}_{\zeta}) p(\boldsymbol{\eta}_{\theta} | \boldsymbol{\eta}_{\eta}) \\
&= \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} p(z_{ijk} | \theta_{jk}, \zeta_i, y_{ijk}) \right\} \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(\theta_{jk} | \boldsymbol{\eta}_{\theta_k}) \right\} \left\{ \prod_{i=1}^I p(\zeta_i) \right\} \\
&\times \left\{ \prod_{k=1}^K p(\xi_{\theta_k}, \omega_{\theta_k}, \lambda_{\theta_k}) \right\} \\
&\propto \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} \exp \{ -0,5(z_{ijk} - a_i \theta_{jk} + b_i)^2 \} \mathbb{1}_{(Z_{ijk}, y_{ijk})} \right\} \\
&\times \left\{ \prod_{k=1}^K \omega_{\theta_k}^{-n_k} \prod_{j=1}^{n_k} \phi \left( \frac{\theta_{jk} - \xi_{\theta_k}}{\omega_{\theta_k}} \right) \Phi \left[ \left( \frac{\theta_{jk} - \xi_{\theta_k}}{\omega_{\theta_k}} \right) \right] \right\} \\
&\times \left\{ \prod_{i=1}^I \exp \{ -0,5(\zeta_i - \boldsymbol{\mu}_{\zeta})^{\top} \boldsymbol{\Psi}_{\zeta}^{-1} (\zeta_i - \boldsymbol{\mu}_{\zeta}) \} \mathbb{1}_{(a_i > 0)} \right\} \\
&\times \left\{ \prod_{k=1}^K \exp \left\{ -\frac{1}{2\sigma_{\xi_{\theta}}^2} (\xi_{\theta_k} - \mu_{\xi_{\theta}})^2 \right\} \right\} \left\{ \prod_{k=1}^K (\omega_{\theta_k}^2)^{-\alpha_{\omega_{\theta}} - 1} \exp(-\beta_{\omega_{\theta}} \frac{1}{\omega_{\theta_k}^2}) \right\} \\
&\times \left\{ \prod_{k=1}^K \exp \left\{ -\frac{1}{2\sigma_{\lambda_{\theta}}^2} (\lambda_{\theta_k} - \mu_{\lambda_{\theta}})^2 \right\} \right\}. \tag{3.23}
\end{aligned}$$

Podemos ver que, as distribuições a posteriori (3.22) e (3.23) têm formas intratáveis, o que justifica a necessidade de algoritmos MCMC para obter amostras dessas. Devido ao esquema de dados aumentados, as distribuições condicionais completas dos parâmetros dos itens e dos dados aumentados têm formas conhecidas e de fácil amostragem. Entretanto, para os traços latentes e parâmetros populacionais essas distribuições não possuem formas conhecidas. Por conta disso, não poderemos utilizar o amostrador de *Gibbs* completo. Para contornar esse problema, podemos utilizar passos de *Metropolis-Hastings* dentro do amostrador de *Gibbs* (MHGS)<sup>1</sup>, veja Patz and Junker (1999a,b). Entretanto, diferentemente dessa abordagem, estamos considerando aqui a estrutura de dados aumentados ao invés da verossimilhança original. Portanto, utilizaremos o que Azevedo et.al (2010) chamaram de algoritmo (ADMHGS)<sup>2</sup>.

Para realizar os passos de *Metropolis-Hastings* precisamos escolher densidades proposta, veja Gelman (2004), tanto para os traços latentes, quanto para os parâmetros populacionais.

<sup>1</sup>Sigla do inglês: *Metropolis-Hastings within Gibbs sampling*

<sup>2</sup>Sigla do inglês: *Augmented data Metropolis-Hastings within Gibbs sampling*

Para os traços latentes, utilizaremos

$$J_t(\theta_{jk}^{(*)}|\theta_{jk}^{(t-1)}) = N(\theta_{jk}^{(t-1)}, \sigma_\theta^2), \quad (3.24)$$

em que,  $J_t(\cdot)$  denota a densidade proposta (ou densidade de transição) no passo  $t$ . Para os parâmetros populacionais, utilizaremos as seguintes densidades propostas:

$$J_t(\xi_k^{(*)}|\xi_k^{(t-1)}) = N(\xi_k^{(t-1)}, \sigma_0^2), \quad (3.25)$$

$$J_t(\omega_k^{(*)}|\omega_k^{(t-1)}) = \text{Log-Normal}(\omega_k^{(t-1)}, \sigma_0^2), \quad (3.26)$$

$$J_t(\lambda_k^{(*)}|\lambda_k^{(t-1)}) = N(\lambda_k^{(t-1)}, \sigma_0^2). \quad (3.27)$$

Denotando por  $(\cdot)$  o conjunto de todos os outros parâmetros, temos que o algoritmo (ADMHGS) proposto, para  $t = 1, 2, \dots, B, \dots, M$ , onde  $B$  é o *burn-in* e  $M$  é o tamanho de amostra gerado; simula iterativamente todas as quantidades desconhecidas na seguinte ordem:

- Inicie o algoritmo escolhendo valores iniciais convenientes;
- Simule as variáveis aumentadas  $Z_{ijk}$  de  $Z_{ijk}|\cdot$ ,  $k = 1, \dots, K$ ,  $j = 1, \dots, n_k$  e  $i \in I_{jk}$ ;
- Simule  $\theta_{jk}$  de  $\theta_{jk}|\cdot$ ,  $k = 1, \dots, K$  e  $j = 1, \dots, n_k$ ;
- Simule  $\zeta_i$  de  $\zeta_i|\cdot$ ,  $i \in I_{jk}$ ;
- Simule  $\xi_{\theta_k}$  de  $\xi_{\theta_k}|\cdot$ ,  $k = 1, \dots, K$ ;
- Simule  $\omega_{\theta_k}$  de  $\omega_{\theta_k}|\cdot$ ,  $k = 1, \dots, K$ ;
- Simule  $\lambda_{\theta_k}$  de  $\lambda_{\theta_k}|\cdot$ ,  $k = 1, \dots, K$ .

Uma descrição mais detalhada desses algoritmos pode ser vista no Apêndice A.1.

### 3.3.2 MCMC sob uma estrutura hierárquica

Uma outra opção para implementação do algoritmo MCMC, seria fazê-lo utilizando representações estocástica das densidades a priori visando reduzir, ou até mesmo, eliminar os passos de *Metropolis-Hastings* dentro do amostrador de *Gibbs*. O número de densidades proposta usadas nos passos de *Metropolis-Hastings* pode ser uma problema na implementação do algoritmo, vide Patz and Junker (1999), portanto, nosso objetivo é obter todas as distribuições condicionais completas com formas conhecidas. Isso possibilitará a utilização do amostrador de Gibbs

completo, como veremos a seguir.

Como vimos, estamos assumindo como distribuição a priori para os traços latentes uma  $\text{NA}_c(\mu_{\theta_k}, \sigma_{\theta_k}^2, \gamma_{\theta_k})$ , vide equação (3.2). Utilizando a representação estocástica de Henze (1986), temos que

$$\theta_{jk} = \xi_{\theta_k} + \omega_{\theta_k} \left( \delta_{\theta_k} H_{jk} + \sqrt{1 - \delta_{\theta_k}^2} Q_{jk} \right), \quad (3.28)$$

em que,  $H_{jk} \sim \text{HN}(0, 1)$ ,  $Q_{jk} \sim N(0, 1)$ ,  $H_{jk} \perp Z_{jk}$ ,  $\forall j, k$ . Podemos ainda reescrever (3.28) de forma hierárquica da seguinte maneira,

$$\theta_{jk} | (h_{jk}, \boldsymbol{\eta}_{\theta_k}) \sim N(\xi_{\theta_k} + \tau_{\theta_k} h_{jk}, \varsigma_{\theta_k}^2), \quad (3.29)$$

$$H_{jk} \sim \text{HN}(0, 1), \quad (3.30)$$

em que,

$$\begin{aligned} \tau_{\theta_k} &= \omega_{\theta_k} \delta_{\theta_k}, \\ \varsigma_{\theta_k}^2 &= \omega_{\theta_k}^2 (1 - \delta_{\theta_k}^2). \end{aligned}$$

Note que,  $\tau_{\theta_k}$  e  $\varsigma_{\theta_k}$  são uma reparametrização feita de modo a facilitar o trabalho computacional, conforme discutido no Capítulo 2, vide equação (2.17).

Considerando agora os parâmetros populacionais, podemos notar que a priori (3.17) pode ser reescrita de maneira hierárquica da seguinte forma,

$$\lambda_{\theta_k} | t_k \sim N(0, \frac{\varphi^2}{t_k}), \quad (3.31)$$

$$T_k \sim \text{Gama}(\frac{d}{2}, \frac{d}{2}). \quad (3.32)$$

Logo, a priori (3.15) pode ser reescrita da seguinte como,

$$p(\lambda_{\theta_k}, \xi_{\theta_k}, \omega_{\theta_k}, t_k) \propto \frac{1}{\omega_{\theta_k}} \exp\left(-\frac{1}{2} \frac{\lambda_{\theta_k}^2 t_k}{\varphi^2}\right) t_k^{\frac{d+1}{2}-1} \exp\left(-\frac{d}{2} t_k\right). \quad (3.33)$$

Considerando a reparametrização teremos,

$$p(\xi_{\theta_k}, \tau_{\theta_k}, \varsigma_{\theta_k}, t_k) \propto \frac{1}{\omega_{\theta_k}} \exp\left(-\frac{1}{2} \frac{\tau_{\theta_k}^2 t_k}{\varsigma_{\theta_k}^2 \varphi^2}\right) t_k^{\frac{d+1}{2}-1} \exp\left(-\frac{d}{2} t_k\right). \quad (3.34)$$

A equação (3.34) é uma priori conjunta para o conjunto de parâmetros  $(\xi_{\theta_k}, \tau_{\theta_k}, \varsigma_{\theta_k})$  e variáveis latentes  $T_k$ .

Por outro lado, podemos notar que, a equação (3.29) pode ser vista como um modelo de regressão com variável resposta  $\theta_{jk}$ , intercepto  $\xi_{\theta_k}$  e inclinação  $\tau_{\theta_k}$ . Portanto, analogamente ao que foi feito para os parâmetros dos itens, poderíamos utilizar, também, uma distribuição normal bivariada como priori para o vetor  $(\xi_{\theta_k}, \tau_{\theta_k})^\top$ , o qual denotaremos por  $\beta_{\theta_k}$ , e uma priori Gama-Inversa para o parâmetro de variância  $\varsigma_{\theta_k}^2$ . Ou seja,

$$\beta_{\theta_k} \sim N(\mu_\theta, \Sigma_\theta), \quad (3.35)$$

$$\varsigma_{\theta_k}^2 \sim \text{Gama-Inv.}(\alpha_{\varsigma_\theta}, \beta_{\varsigma_\theta}), \quad (3.36)$$

em que,

$$\beta_{\theta_k} = \begin{pmatrix} \xi_{\theta_k} \\ \tau_{\theta_k} \end{pmatrix}, \quad \mu_\theta = \begin{pmatrix} \mu_{\xi_\theta} \\ \mu_{\tau_\theta} \end{pmatrix} \quad \text{e} \quad \Sigma_\theta = \begin{pmatrix} \sigma_{\xi_\theta}^2 & \rho \sigma_{\xi_\theta}^2 \sigma_{\tau_\theta}^2 \\ \rho \sigma_{\xi_\theta}^2 \sigma_{\tau_\theta}^2 & \sigma_{\tau_\theta}^2 \end{pmatrix}.$$

Assumiremos, agora, a seguinte forma geral para priori conjunta dos parâmetros,

$$\begin{aligned} p(\theta_{..}, \mathbf{h}_{..}, \zeta, \boldsymbol{\eta}_\theta, \mathbf{t}_{..} | \boldsymbol{\eta}_\zeta, \boldsymbol{\eta}_\eta) &= \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(\theta_{jk} | \boldsymbol{\eta}_{\theta_k}) \right\} \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(h_{jk}) \right\} \left\{ \prod_{i=1}^I p(\zeta_i | \boldsymbol{\eta}_\zeta) \right\} \\ &\times \left\{ \prod_{k=1}^K p(\boldsymbol{\eta}_{\theta_k} | \boldsymbol{\eta}_\eta) \right\} \left\{ \prod_{k=1}^K p(t_k) \right\}, \end{aligned} \quad (3.37)$$

em que,  $\mathbf{h}_{..} = (\mathbf{h}_{1.}, \dots, \mathbf{h}_{n_k.})^\top$  e  $\mathbf{t}_{..} = (t_1, \dots, t_K)^\top$ . As distribuições a priori para os traços latentes e variáveis  $\mathbf{h}$  são dadas por (3.29) e (3.30), respectivamente. Para os parâmetros dos itens, continuaremos utilizando a priori (3.14). Já para os parâmetros populacionais, podemos utilizar a priori conjunta (3.34) ou as prioris (3.35) e (3.36). Considerando, então, a verossimilhança aumentada (3.11) e a priori (3.37), temos que a distribuição a posteriori de  $(\mathbf{Z}_{..}, \boldsymbol{\theta}_{..}, \mathbf{h}_{..}, \zeta, \boldsymbol{\eta}_\theta, \mathbf{t}_{..})$

é dada por:

$$\begin{aligned}
p(\boldsymbol{\theta}_{..}, \mathbf{h}_{..}, \boldsymbol{\zeta}, \boldsymbol{\eta}_{\theta}, \mathbf{t}_{.} | \mathbf{y}_{...}, \boldsymbol{\eta}_{\zeta}, \boldsymbol{\eta}_{\eta}) &\propto p(\mathbf{Z}_{...} | \boldsymbol{\theta}_{..}, \boldsymbol{\zeta}, \mathbf{y}_{...}) p(\boldsymbol{\theta}_{..} | \mathbf{h}_{..}, \boldsymbol{\eta}_{\theta}) p(\boldsymbol{\zeta} | \boldsymbol{\eta}_{\zeta}) p(\boldsymbol{\eta}_{\theta} | \mathbf{t}_{.}, \boldsymbol{\eta}_{\eta}) p(\mathbf{t}_{.}) \\
&= \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} p(z_{ijk} | \theta_{jk}, \zeta_i, y_{ijk}) \right\} \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(\theta_{jk} | h_{jk}, \boldsymbol{\eta}_{\theta_k}) \right\} \\
&\times \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(h_{jk}) \right\} \left\{ \prod_{i=1}^I p(\zeta_i) \right\} \left\{ \prod_{k=1}^K p(\boldsymbol{\eta}_{\theta_k} | \boldsymbol{\eta}_{\eta_k}) \right\} \left\{ \prod_{k=1}^K p(t_k) \right\} \\
&\propto \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} \exp \left\{ -0,5(z_{ijk} - a_i \theta_{jk} + b_i)^2 \right\} \mathbb{1}_{(Z_{ijk}, y_{ijk})} \right\} \\
&\times \left\{ \prod_{k=1}^K \varsigma_{\theta_k}^{-\frac{n_k}{2}} \prod_{j=1}^{n_k} \exp \left\{ -\frac{1}{2\varsigma_{\theta_k}^2} [\theta_{jk} - \xi_{\theta_k} - \tau_{\theta_k} h_{jk}]^2 \right\} \right\} \\
&\times \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \exp \left( -\frac{1}{2} h_{jk}^2 \right) \mathbb{1}_{(h_{jk} > 0)} \right\} \\
&\times \left\{ \prod_{i=1}^I \exp \left[ -0,5(\zeta_i - \boldsymbol{\mu}_{\zeta})^{\top} \boldsymbol{\Psi}_{\zeta}^{-1} (\zeta_i - \boldsymbol{\mu}_{\zeta}) \right] \mathbb{1}_{(a_i > 0)} \right\} \\
&\times \left\{ \prod_{k=1}^K \frac{1}{\varsigma_{\theta_k}^2} \exp \left( -\frac{1}{2} \frac{\tau_{\theta_k}^2 t_k}{\varsigma_{\theta_k}^2 \varphi^2} \right) t_k^{\frac{d+1}{2}-1} \exp \left( -\frac{d}{2} t_k \right) \right\}, \tag{3.38}
\end{aligned}$$

se utilizarmos a priori (3.34), ou

$$\begin{aligned}
p(\boldsymbol{\theta}_{..}, \mathbf{h}_{..}, \boldsymbol{\zeta}, \boldsymbol{\eta}_{\theta} | \mathbf{y}_{...}, \boldsymbol{\eta}_{\zeta}, \boldsymbol{\eta}_{\eta}) &\propto p(\mathbf{Z}_{...} | \boldsymbol{\theta}_{..}, \boldsymbol{\zeta}, \mathbf{y}_{...}) p(\boldsymbol{\theta}_{..} | \mathbf{h}_{..}, \boldsymbol{\eta}_{\theta}) p(\boldsymbol{\zeta} | \boldsymbol{\eta}_{\zeta}) p(\boldsymbol{\eta}_{\theta} | \boldsymbol{\eta}_{\eta}) \\
&= \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} p(z_{ijk} | \theta_{jk}, \zeta_i, y_{ijk}) \right\} \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(\theta_{jk} | h_{jk}, \boldsymbol{\eta}_{\theta_k}) \right\} \\
&\times \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} p(h_{jk}) \right\} \left\{ \prod_{i=1}^I p(\zeta_i) \right\} \left\{ \prod_{k=1}^K p(\boldsymbol{\eta}_{\theta_k} | \boldsymbol{\eta}_{\eta_k}) \right\} \left\{ \prod_{k=1}^K p(t_k) \right\} \\
&\propto \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \prod_{i \in I_{jk}} \exp \left\{ -0,5(z_{ijk} - a_i \theta_{jk} + b_i)^2 \right\} \mathbb{1}_{(Z_{ijk}, y_{ijk})} \right\} \\
&\times \left\{ \prod_{k=1}^K \varsigma_{\theta_k}^{-\frac{n_k}{2}} \prod_{j=1}^{n_k} \exp \left\{ -\frac{1}{2\varsigma_{\theta_k}^2} [\theta_{jk} - \xi_{\theta_k} - \tau_{\theta_k} h_{jk}]^2 \right\} \right\} \\
&\times \left\{ \prod_{k=1}^K \prod_{j=1}^{n_k} \exp \left( -\frac{1}{2} h_{jk}^2 \right) \mathbb{1}_{(h_{jk} > 0)} \right\} \\
&\times \left\{ \prod_{i=1}^I \exp \left[ -0,5(\zeta_i - \boldsymbol{\mu}_{\zeta})^{\top} \boldsymbol{\Psi}_{\zeta}^{-1} (\zeta_i - \boldsymbol{\mu}_{\zeta}) \right] \mathbb{1}_{(a_i > 0)} \right\} \\
&\times \left\{ \exp \left\{ -\frac{1}{2} (\boldsymbol{\beta}_{\theta} - \boldsymbol{\mu}_{\theta})^{\top} \boldsymbol{\Sigma}_{\theta}^{-1} (\boldsymbol{\beta}_{\theta} - \boldsymbol{\mu}_{\theta}) \right\} \right\} \\
&\times \left\{ (\varsigma_{\theta_k}^2)^{-(\alpha_{\varsigma_{\theta}}+1)} \exp \left( -\frac{\beta_{\varsigma_{\theta}}}{\varsigma_{\theta_k}^2} \right) \right\}, \tag{3.39}
\end{aligned}$$

se utilizarmos as prioris (3.35) e (3.36). A partir das posteriores (3.38) e (3.39) obtemos todas as distribuições condicionais completas com formas conhecidas, veja Apêndice A.2. Portanto, podemos agora utilizar o amostrador de *Gibbs* completo, sem a necessidade dos passos de Metropolis-Hastings. Denotaremos esse segundo algoritmo por ADGS<sup>3</sup>. O algoritmo ADGS pode ser executado na seguinte ordem:

- Inicie o algoritmo escolhendo valores iniciais convenientes;
- Simule as variáveis aumentadas  $Z_{ijk}$  de  $Z_{ijk}|\cdot$ ,  $k = 1, \dots, K$ ,  $j = 1, \dots, n_k$  e  $i \in I_{jk}$ ;
- Simule  $h_{jk}$  de  $H_{jk}|\cdot$ ,  $k = 1, \dots, K$  e  $j = 1, \dots, n_k$ ;
- Simule  $\theta_{jk}$  de  $\theta_{jk}|\cdot$ ,  $k = 1, \dots, K$  e  $j = 1, \dots, n_k$ ;
- Simule  $\zeta_i$  de  $\zeta_i|\cdot$ ,  $i \in I_{jk}$ ;
- Simule  $t_k$  de  $T_k|\cdot$ ,  $k = 1, \dots, K$ ;
- Simule  $\xi_{\theta_k}$  de  $\xi_{\theta_k}|\cdot$ ,  $k = 1, \dots, K$ ;
- Simule  $\tau_{\theta_k}$  de  $\tau_{\theta_k}|\cdot$ ,  $k = 1, \dots, K$ ;
- Simule  $\varsigma_{\theta_k}^2$  de  $\varsigma_{\theta_k}^2|\cdot$ ,  $k = 1, \dots, K$ ;

se utilizarmos a distribuição a posteriori (3.38). Já se utilizarmos a distribuição a posteriori (3.39), devemos seguir os seguintes passos:

- Inicie o algoritmo escolhendo valores iniciais convenientes;
- Simule as variáveis aumentadas  $Z_{ijk}$  de  $Z_{ijk}|\cdot$ ,  $k = 1, \dots, K$ ,  $j = 1, \dots, n_k$  e  $i \in I_{jk}$ ;
- Simule  $h_{jk}$  de  $H_{jk}|\cdot$ ,  $k = 1, \dots, K$  e  $j = 1, \dots, n_k$ ;
- Simule  $\theta_{jk}$  de  $\theta_{jk}|\cdot$ ,  $k = 1, \dots, K$  e  $j = 1, \dots, n_k$ ;
- Simule  $\zeta_i$  de  $\zeta_i|\cdot$ ,  $i \in I_{jk}$ ;
- Simule  $\beta_{\theta_k}$  de  $\beta_{\theta_k}|\cdot$ ,  $k = 1, \dots, K$ ;
- Simule  $\varsigma_{\theta_k}^2$  de  $\varsigma_{\theta_k}^2|\cdot$ ,  $k = 1, \dots, K$ .

Para mais detalhes, veja Apêndice A.2.

---

<sup>3</sup>Sigla do inglês: *Augmented data Gibbs sampling*

# Capítulo 4

## Estudo de Simulação

### 4.1 Introdução

Ao longo deste capítulo, avaliaremos o desempenho do modelo de grupos múltiplos assimétrico e também a eficácia do método de estimação proposto. Primeiramente, iremos avaliar a convergência dos nossos algoritmos e também compará-los em termos do tamanho efetivo de amostra, como feito em Sahu (2002). Em seguida, um estudo de recuperação de parâmetros será conduzido. Neste, avaliamos o impacto de alguns fatores, tais como, número total de respondentes, número de itens por grupo, na precisão das estimativas. Por fim, apresentamos algumas conclusões e comentários.

### 4.2 Avaliação da convergência

Nesta seção avaliaremos os dois algoritmos MCMC propostos no Capítulo 3 em termos da convergência: avaliando a magnitude das autocorrelações e, também, avaliamos sensibilidade com relação à escolha de valores iniciais. As análises serão feitas com base em um único conjunto de respostas, que foram simuladas considerando-se a seguinte situação: 4 grupos (nomeadamente 1, 2, 3 e 4) respondendo a 4 testes (nomeadamente 1, 2, 3 e 4) parcialmente distintos, compostos da seguinte forma:

- teste 1: 20 itens;
- teste 2: teste 1 + 20 itens inéditos;
- teste 3: os últimos 20 itens do teste 2 + 20 itens inéditos;
- teste 4: os últimos 20 itens do teste 3 + 20 itens inéditos.

Isso nos dá um conjunto total de 80 itens. Essa situação é similar à situação real que será analisada no Capítulo 5. Estamos considerando também o mesmo padrão de não resposta dos dados reais. Assim poderemos avaliar a convergência de nossos algoritmos sob circunstâncias muito comuns na prática, como a presença de dados faltantes e desbalanceamento entre grupos, com a vantagem de conhecer os verdadeiros valores dos parâmetros.

Os parâmetros dos itens foram fixados nos seguintes intervalos:  $a_i \in [0, 7; 1, 4]$  e  $b_i^* \in [-2; 4]$ . Os valores dos parâmetros de dificuldade foram fixados de acordo com a média e desvio padrão dos traços latentes simulados e, a partir daí, foram fixados os valores dos parâmetros de discriminação de modo a contemplar várias situações (itens difíceis com discriminação alta, difíceis com discriminação baixa, etc). Já os traços latentes foram simulados, de forma independente, a partir de distribuições normais assimétricas centralizadas com médias  $\mu_\theta = (0; 1; 1, 4; 2)^\top$ , desvios-padrão  $\sigma_\theta = (1; 0, 88; 0, 62; 0, 77)^\top$  e coeficientes de assimetria  $\gamma_\theta = (0; 0, 14; 0, 5; -0, 5)^\top$  para os grupos 1, 2, 3 e 4 respectivamente. Os valores escolhidos para as médias e desvios-padrão são próximos das estimativas obtidas no Capítulo 6 na análise de dados reais. Como já dissemos, é nosso interesse simular uma situação bem parecida com uma que realmente ocorreu. Já os valores dos coeficientes de assimetria foram escolhidos, de modo a contemplar vários níveis de assimetria (no caso, assimetria nula, fraca, forte à direita e forte à esquerda).

Os tamanhos amostrais foram fixados em 556, 556, 401 e 294 para os grupos 1, 2, 3 e 4 respectivamente. Seguindo DeMars (2003), para assegurar que as estimativas fiquem na mesma métrica dos valores simulados aplicou-se uma transformação aos traços latentes simulados afim de deixá-los com as médias e variâncias desejadas. A transformação foi a seguinte:

$$\theta_{jk}^{**} = \sigma_{\theta_k} s_{\theta_k}^{-1} (\theta_{jk}^* - \bar{\theta}_k) + \mu_{\theta_k},$$

em que  $\theta_{jk}^*$  são os traços latentes simulados e  $\bar{\theta}_k$  e  $s_{\theta_k}^{-1}$  são a média e o desvio-padrão amostrais dos traços latentes simulados, respectivamente.

### 4.2.1 Escolha de Priors

Discutiremos agora a escolha de hiperparâmetros para as especificações a priori que foram adotadas, vide subseções 3.3.1 e 3.3.2. Para os traços latentes, como dito anteriormente, assumiremos a priori (3.2). Para os parâmetros dos itens, utilizamos o vetor de médias (1; 0) e a matriz diagonal (0, 5; 9), ou seja, assumimos independência a priori. Com relação às médias das priors dos parâmetros dos itens, temos que, para a discriminação o valor 1 indica que, em média, esperamos que os itens apresentem uma boa discriminação. Para o parâmetros de difi-

culdade, o valor 0 indica que, em média, espera-se que os itens tenham dificuldade em torno do valor médio da escala (0, 1). Com relação às variâncias, o valor 9 para as dificuldades, permite que estas variem razoavelmente, uma vez que os testes podem diferir bastante entre si. Por outro lado a discriminação tem de ser positiva (pelo menos, no que diz respeito a conhecimento), não muito baixa e nem muito elevada. Dessa forma, o valor 0,5 parece ser razoável. Para os parâmetros populacionais (reparametrizados) estamos considerando, as prioris de Jeffreys e a priori induzida pela uniforme, vide subseção 2.3.2 e, também, as prioris conjugadas definidas nas subseções 3.3.1 e 3.3.2. No último caso, fixamos hiperparâmetros que nos levaram a prioris com médias 0 para as médias e coeficientes de assimetria e média 1 para os desvios-padrão. Ou seja, estamos assumindo, a priori, que as médias são iguais, os desvios-padrão são em média iguais ao da escala (0,1) e que não existe assimetria. As Tabelas 4.1 e 4.2 resumem os valores que adotamos para todos os hiperparâmetros. Doravante, denotaremos por priori 1, 2, 3 e 4 os quatro conjuntos de prioris definidos. Note que a diferença entre esses conjuntos está apenas nas prioris dos parâmetros populacionais. Ou seja, as prioris 1 e 2, são as prioris de Jeffreys e a induzida pela uniforme, respectivamente. As prioris 3 e 4, levam as médias 0, 1 e 0 nas prioris dos parâmetros originais, como dito anteriormente. Sendo que, a priori 3 leva a variâncias aproximadamente 1 e a priori 4, conduz a variâncias aproximadamente 10, exceto para o coeficiente de assimetria, cujas variâncias foram mantidas próximas de zero. Notamos que, a distribuição de  $\gamma_{\theta_k}$  tende a se concentrar nos valores -0,9 e 0,9 com o aumento da variância. Tal comportamento também é observado quando utilizamos a priori de Jeffreys. A Figura 4.1 apresenta as prioris de  $\gamma_{\theta_k}$  quando utilizamos as prioris de Jeffreys e a priori induzida pela uniforme para  $\lambda_{\theta_k}$ . Já a Figura (4.2) compara as distribuições de  $\lambda_{\theta_k}$  e  $\gamma_{\theta_k}$  para diferentes variâncias, quando utilizamos uma priori normal para  $\lambda_{\theta_k}$ . As prioris do coeficiente de assimetria são representadas pelos histogramas de seus valores simulados.

Tabela 4.1: Hiperparâmetros escolhidos para as prioris 3.14, 3.15, 3.19, 3.20 e 3.21 relacionadas ao algoritmo ADMHGS

| Priori | Hiperparâmetros |              |                   |                                             |                                                   |                                                     |
|--------|-----------------|--------------|-------------------|---------------------------------------------|---------------------------------------------------|-----------------------------------------------------|
|        | $\mu_\zeta$     | $\Psi_\zeta$ | $(d; \varphi^2)$  | $(\mu_{\xi_\theta}, \sigma_{\xi_\theta}^2)$ | $(\alpha_{\omega_\theta}, \beta_{\omega_\theta})$ | $(\mu_{\lambda_\theta}, \sigma_{\lambda_\theta}^2)$ |
| 1      | (1; 0)          | (0,5; 9)     | $(0, 5; \pi^2/4)$ | -                                           | -                                                 | -                                                   |
| 2      | (1; 0)          | (0,5; 9)     | (2; 0, 5)         | -                                           | -                                                 | -                                                   |
| 3      | (1; 0)          | (0,5; 9)     | -                 | $(0; 0, 85^2)$                              | (3,1; 2,8)                                        | $(0; 1^2)$                                          |
| 4      | (1; 0)          | (0,5; 9)     | -                 | $(0; 0, 31^2)$                              | (2,1; 2,1)                                        | $(0; 2, 2^2)$                                       |

Tabela 4.2: Hiperparâmetros escolhidos para as prioris 3.14, 3.15, 3.35 e 3.36 relacionadas ao algoritmo ADGS

| Priori | Hiperparâmetros |              |                   |                                             |                                                     |                                               |
|--------|-----------------|--------------|-------------------|---------------------------------------------|-----------------------------------------------------|-----------------------------------------------|
|        | $\mu_\zeta$     | $\Psi_\zeta$ | $(d; \varphi^2)$  | $(\mu_{\xi_\theta}, \sigma_{\xi_\theta}^2)$ | $(\alpha_{\zeta_\theta^2}, \beta_{\zeta_\theta^2})$ | $(\mu_{\tau_\theta}, \sigma_{\tau_\theta}^2)$ |
| 1      | (1; 0)          | (0,5; 9)     | $(0, 5; \pi^2/4)$ | -                                           | -                                                   | -                                             |
| 2      | (1; 0)          | (0,5; 9)     | $(2; 0, 5)$       | -                                           | -                                                   | -                                             |
| 3      | (1; 0)          | (0,5; 9)     | -                 | $(0; 0, 7^2)$                               | $(3,4; 3)$                                          | $(0; 0, 8^2)$                                 |
| 4      | (1; 0)          | (0,5; 9)     | -                 | $(0; 3, 1^2)$                               | $(2,1; 2,32)$                                       | $(0; 0, 23^2)$                                |

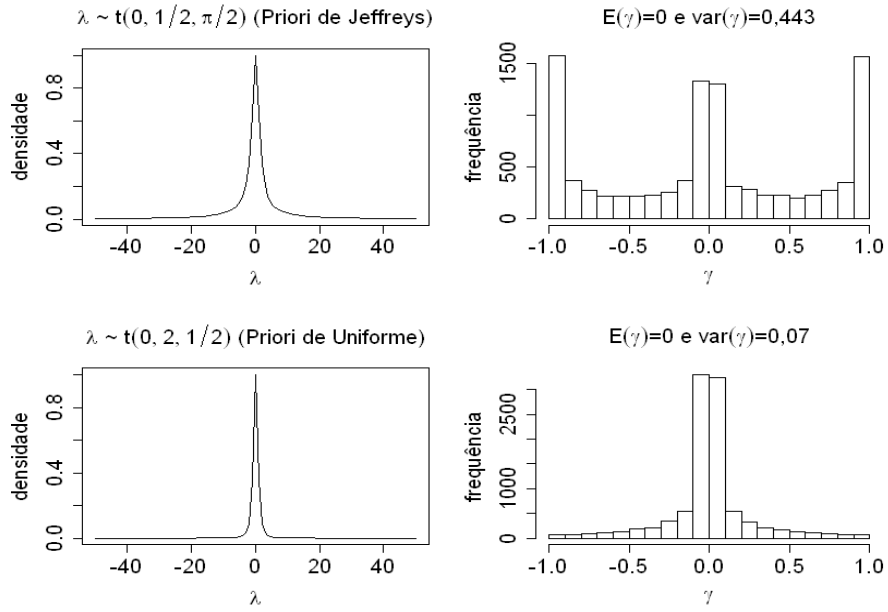


Figura 4.1: Distribuições a priori de  $\lambda_{\theta_k}$  e  $\gamma_{\theta_k}$  considerando as prioris de Jeffreys e Uniforme

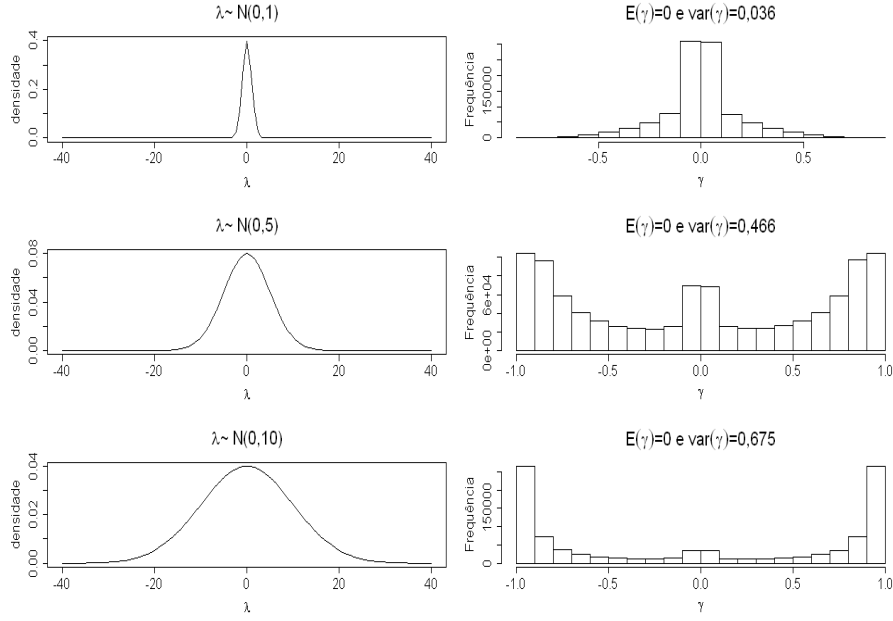


Figura 4.2: Distribuições de  $\lambda_{\theta_k}$  e  $\gamma_{\theta_k}$  para diferentes variâncias considerando a priori normal

Seguindo a recomendação de Gelman (2004), geramos 3 cadeias para cada algoritmo, iniciadas com três diferentes conjuntos de valores. No primeiro conjunto (conjunto 1) utilizamos valores de referência da TRI. O segundo conjunto (conjunto 2) foi todo gerado aleatoriamente a partir de distribuições de probabilidade apropriadas. Já o terceiro conjunto (conjunto 3) foi obtido a partir da amostra. A seguir apresentamos uma descrição mais detalhada desses valores.

#### Conjunto 1:

- Parâmetros populacionais:  $\mu_{\theta} = (0; 0; 0; 0)^{\top}$ ,  $\sigma_{\theta} = (1; 1; 1; 1)^{\top}$ ,  $\gamma_{\theta} = (0; 0; 0; 0)^{\top}$ ;
- Parâmetros dos itens: todos os itens de discriminação foram fixados em um e todos os itens de dificuldade foram fixados em zero;
- Proficiências:  $\theta_{jk} \sim N(0, 1)$ .

#### Conjunto 2:

- Parâmetros populacionais:  $\mu_{\theta_k} \sim N(0; 10)$ ,  $\sigma_{\theta_k} \sim \text{GI}(2, 1; 1, 1)$ ,  $\gamma_{\theta_k} \sim U(-1; 1)$ ;
- Parâmetros dos itens:  $a_i \sim \text{Log-normal}(0, 1)$  e  $b_i \sim N(0, 1)$ ;
- Proficiências:  $\theta_{jk} \sim N(0, 1)$ .

**Conjunto 3:**

- Parâmetros populacionais: Estimativas pelo método dos momentos;
- Parâmetros dos itens: Estatísticas da teoria clássica de medidas, veja Andrade et.al (2000) e Lord (1980);
- Proficiências:  $\theta_{jk}$  = Escores padronizados.

Foram gerados 100000 valores utilizando os algoritmos ADMHGS e ADGS com a priori 1. Simulações preliminares sugeriram que todas as cadeias convergiram após 1000 iterações, exceto as cadeias relacionadas aos coeficientes de assimetria, cuja convergência ocorreu após 20000 iterações. Portanto, a convergência de todas as cadeias ocorreu após 20000 iterações, e consideramos estas como período de aquecimento (*burn in*). As estatísticas de Gelman-Rubin calculadas após o aquecimento variaram entre 1,00 e 1,02 quando utilizamos o algoritmo ADMHGS e entre 1,00 e 1,01 quando foi utilizado o ADGS. Isso indica que a convergência realmente foi alcançada. Para obter amostras independentes, das posteriores de interesse, consideramos um espaçamento entre os valores gerados, uma vez que, ambos algoritmos geraram observações correlacionadas. Nas amostras geradas pelo algoritmo ADMHGS, pegamos valores de 80 em 80. Nas amostras geradas pelo algoritmo ADGS, foi suficiente pegarmos valores de 30 em 30 para reduzir a autocorrelação. As Figuras 4.3 a 4.8 apresentam os gráficos de valores simulados dos parâmetros populacionais e dos parâmetros dos itens 11 e 60 que foram escolhidos aleatoriamente. Os gráficos mostram os valores após o período de aquecimento, considerando os três conjuntos de valores iniciais e com espaçamento da ordem de 80 observações, para facilitar a visualização. Podemos ver que as cadeias estão bem misturadas, indicando que os diferentes valores iniciais não alteram, significativamente, os resultados.

As Figuras 4.9 a 4.16 mostram os correlogramas, com e sem espaçamento, das amostras geradas para os parâmetros populacionais, exceto para a média e desvio-padrão do grupo 1 que foram fixados em zero e um para garantir a identificabilidade, e dos itens 11 e 60, utilizando a priori 1 e o conjunto de valores iniciais 1. Os correlogramas das amostras obtidas utilizando os demais conjuntos de valores iniciais são muito semelhantes à estes e não serão mostrados.

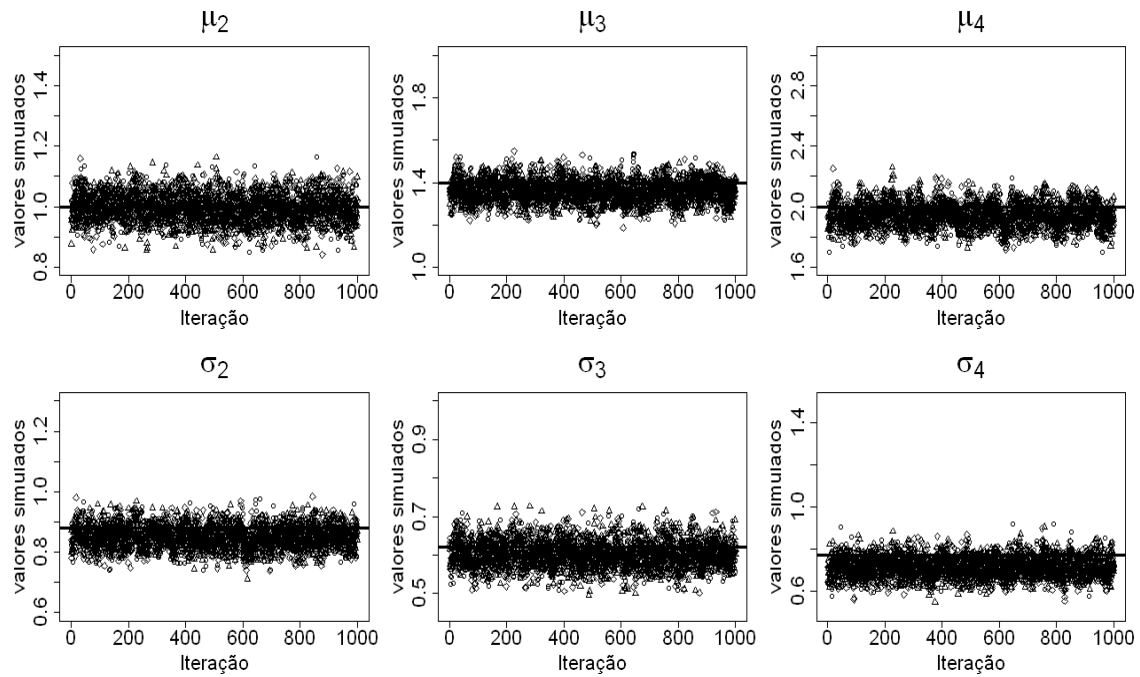


Figura 4.3: Valores simulados para os parâmetros populacionais: Algoritmo ADMHGS. Legenda:  $\circ$  conjunto 1,  $\triangle$  conjunto 2,  $\diamond$  conjunto 3, — valor verdadeiro

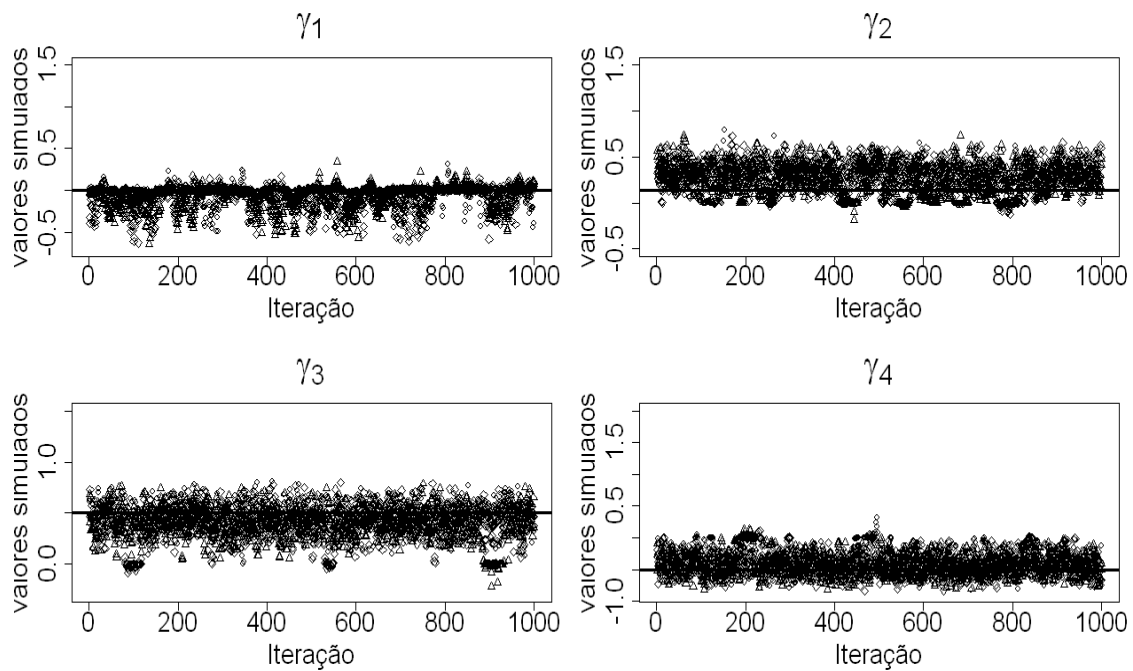


Figura 4.4: Valores simulados para os parâmetros populacionais: Algoritmo ADMHGS (continuação). Legenda:  $\circ$  conjunto 1,  $\triangle$  conjunto 2,  $\diamond$  conjunto 3, — valor verdadeiro

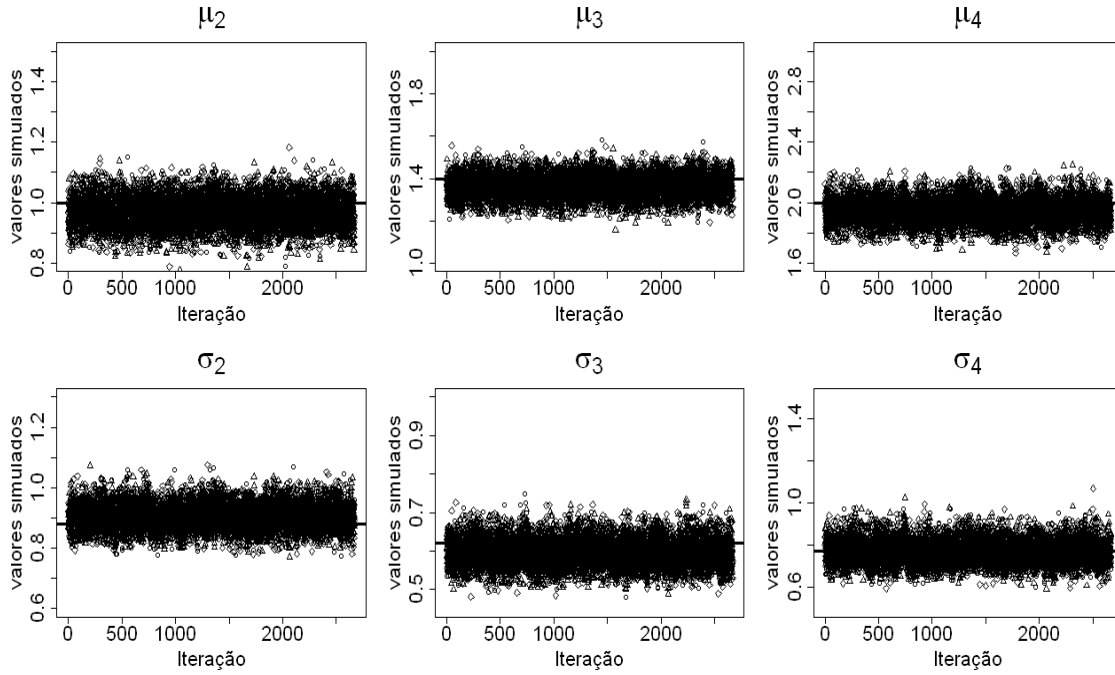


Figura 4.5: Valores simulados para os parâmetros populacionais: Algoritmo ADGS. Legenda:  $\circ$  conjunto 1,  $\triangle$  conjunto 2,  $\diamond$  conjunto 3, — valor verdadeiro

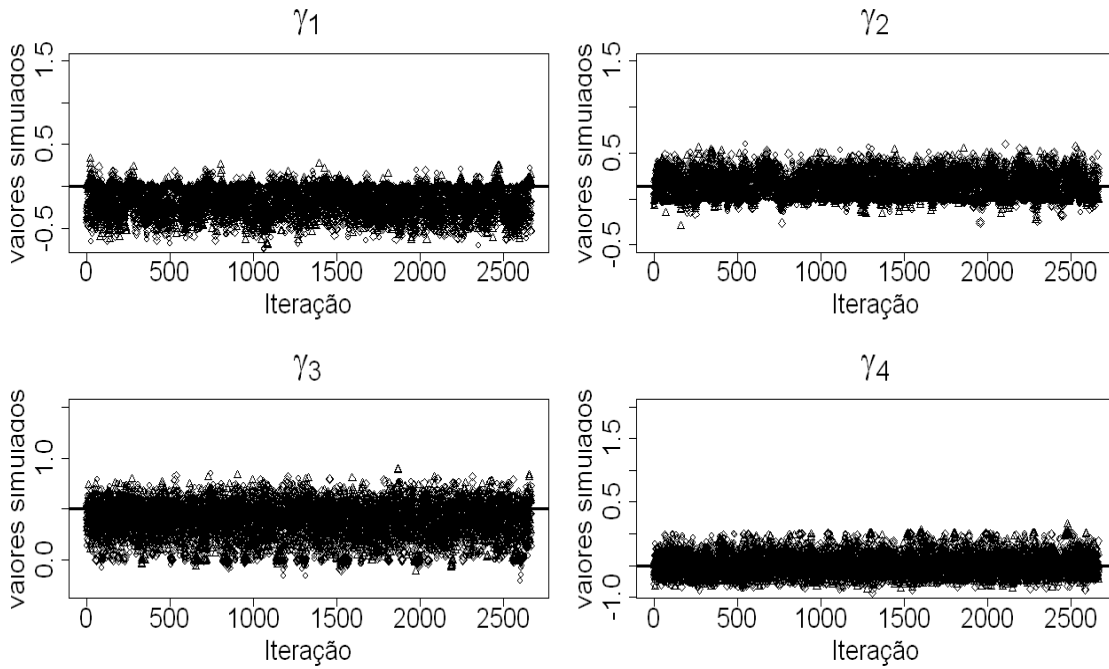


Figura 4.6: Valores simulados para os parâmetros populacionais: Algoritmo ADGS (continuação). Legenda:  $\circ$  conjunto 1,  $\triangle$  conjunto 2,  $\diamond$  conjunto 3, — valor verdadeiro

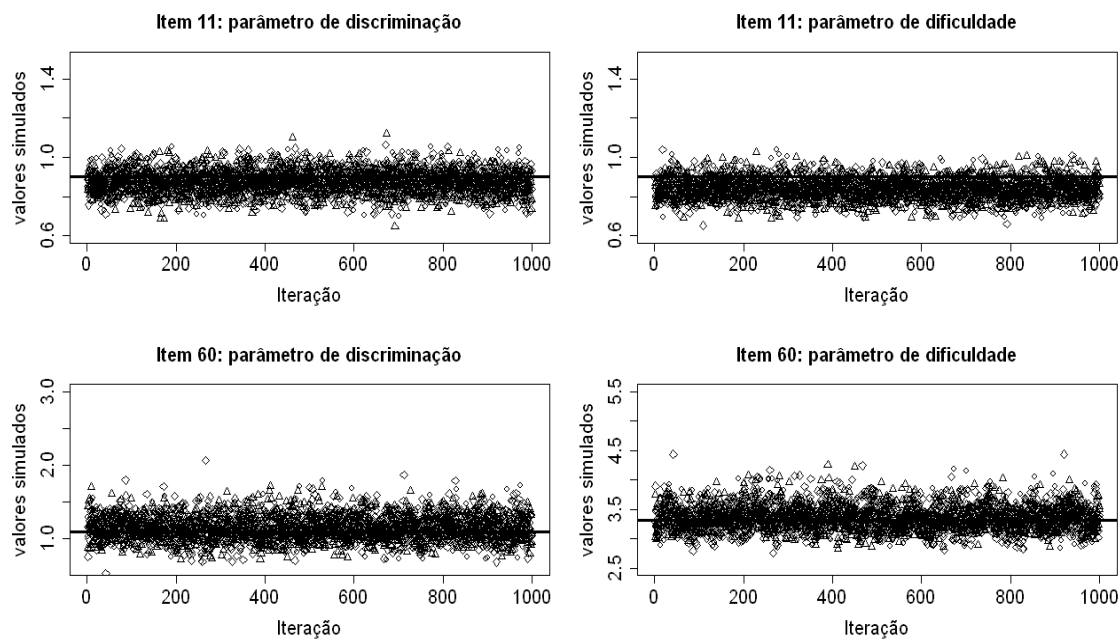


Figura 4.7: Valores simulados para os parâmetros dos itens 11 e 60: Algoritmo ADMHGS. Legenda:  $\circ$  conjunto 1,  $\triangle$  conjunto 2,  $\diamond$  conjunto 3, — valor verdadeiro

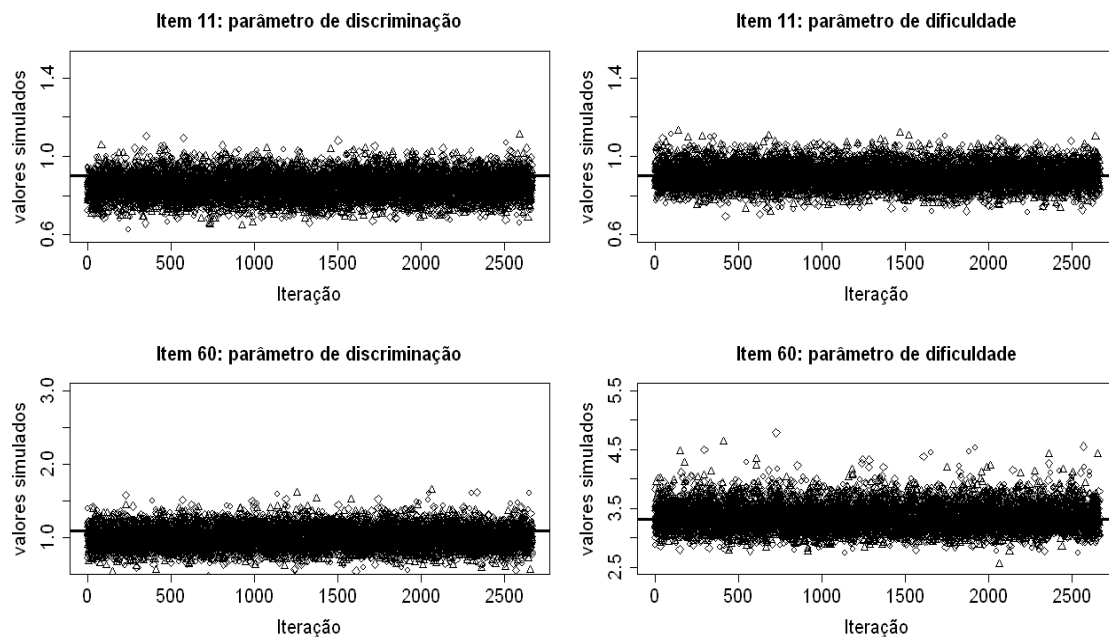


Figura 4.8: Valores simulados para os parâmetros dos itens 11 e 60: Algoritmo ADGS. Legenda:  $\circ$  conjunto 1,  $\triangle$  conjunto 2,  $\diamond$  conjunto 3, — valor verdadeiro

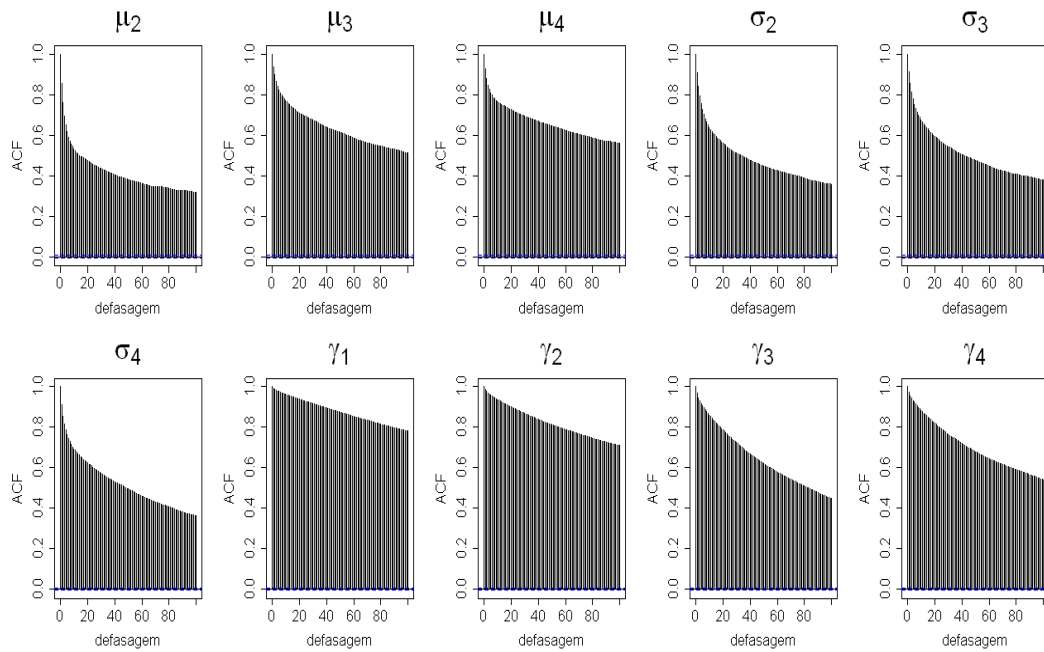


Figura 4.9: Correlogramas sem espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADMHGS

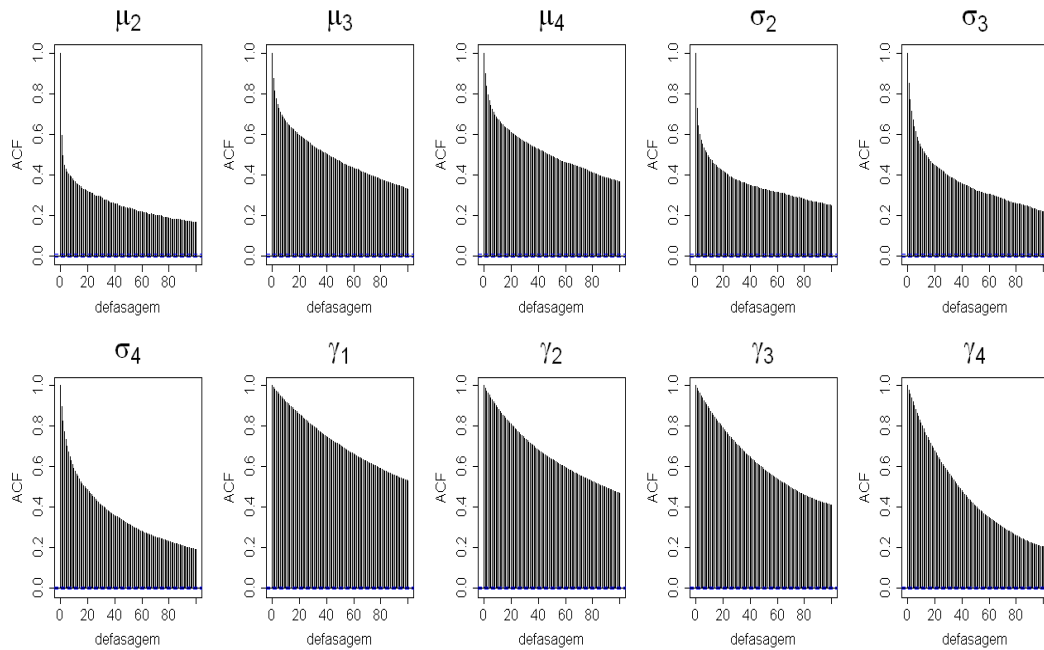


Figura 4.10: Correlogramas sem espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADGS

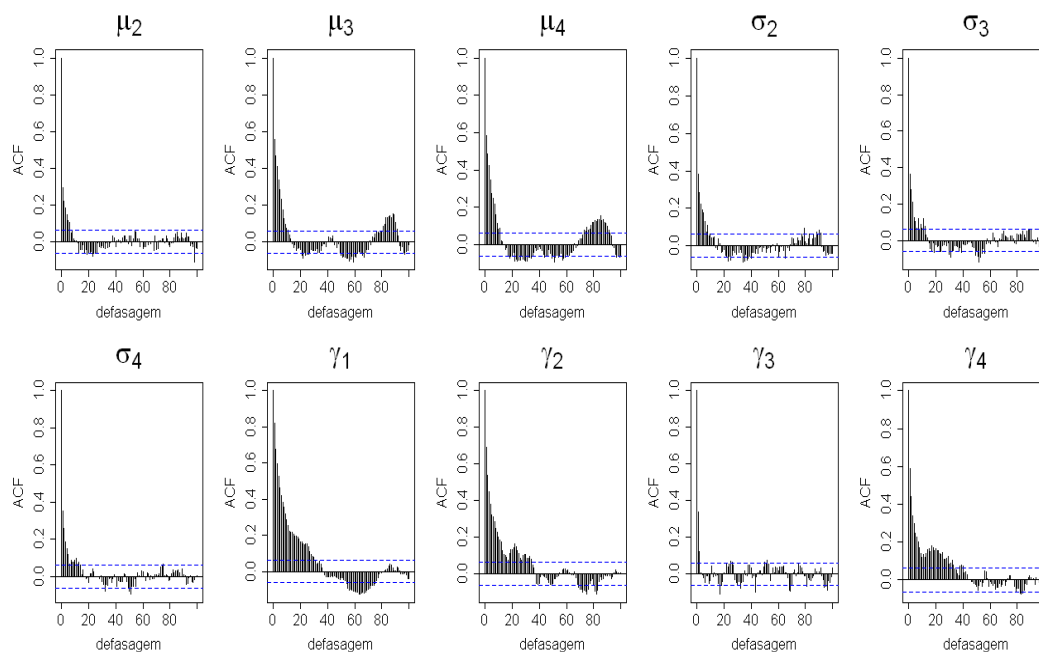


Figura 4.11: Correlogramas com espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADMHGS

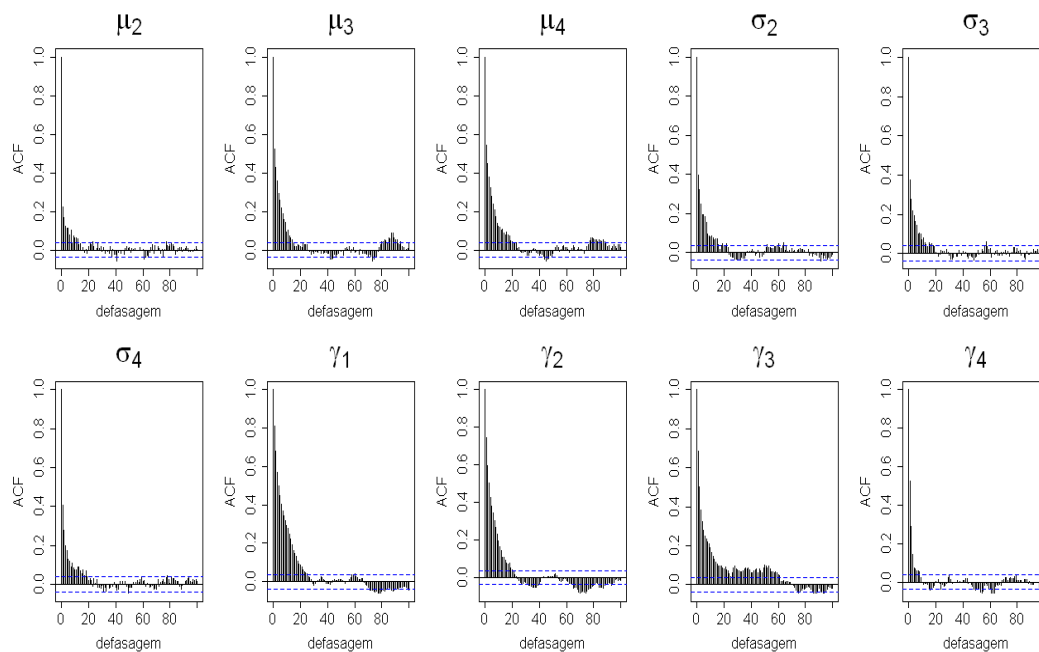


Figura 4.12: Correlogramas com espaçamento das amostras geradas para os parâmetros populacionais: Algoritmo ADGS

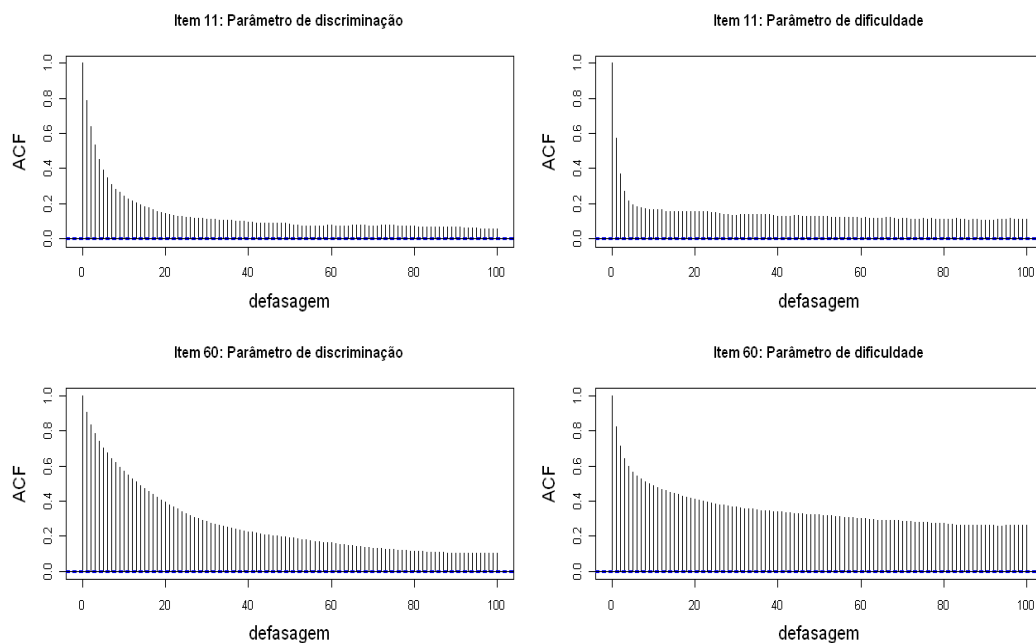


Figura 4.13: Correlogramas sem espaçamento das amostras geradas para os parâmetros dos itens 11 e 60 Algoritmo ADMHGS

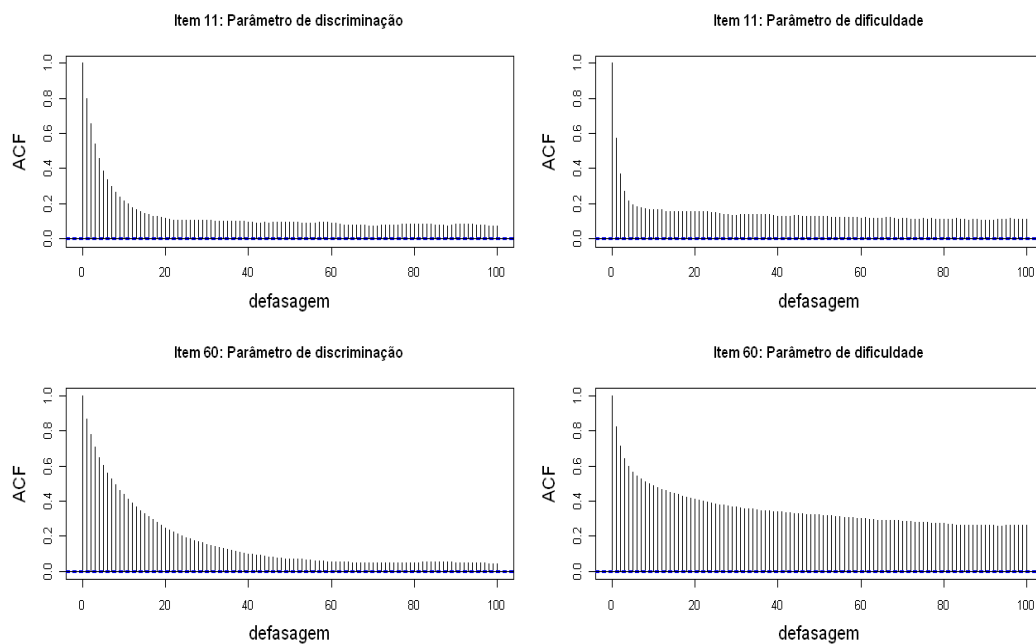


Figura 4.14: Correlogramas sem espaçamento das amostras geradas para os parâmetros dos itens 11 e 60: Algoritmo ADGS

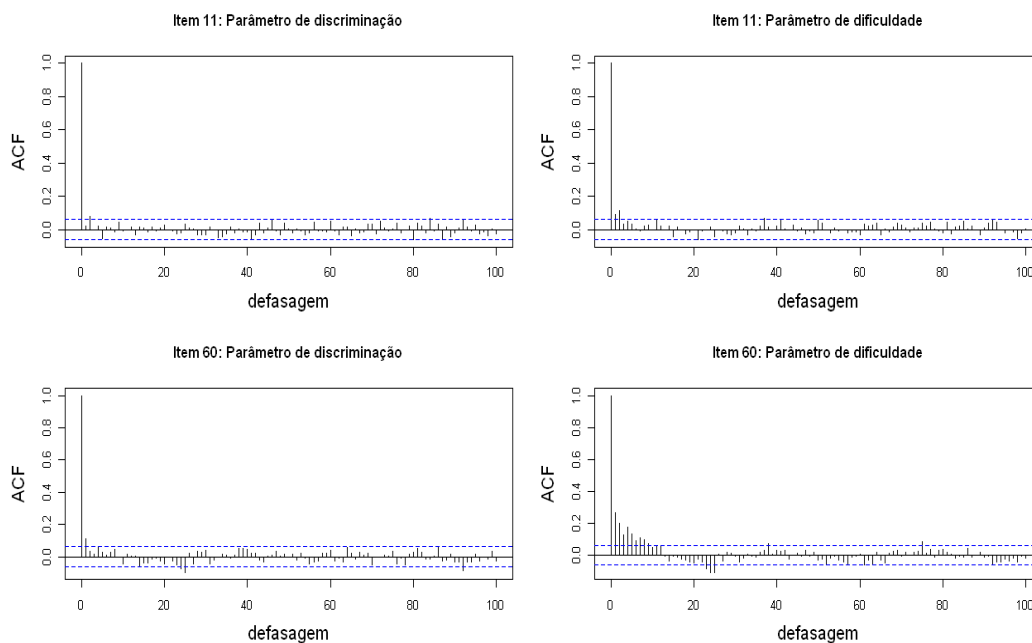


Figura 4.15: Correlogramas com espaçamento das amostras geradas para os parâmetros dos itens 11 e 60: Algoritmo ADMHGS

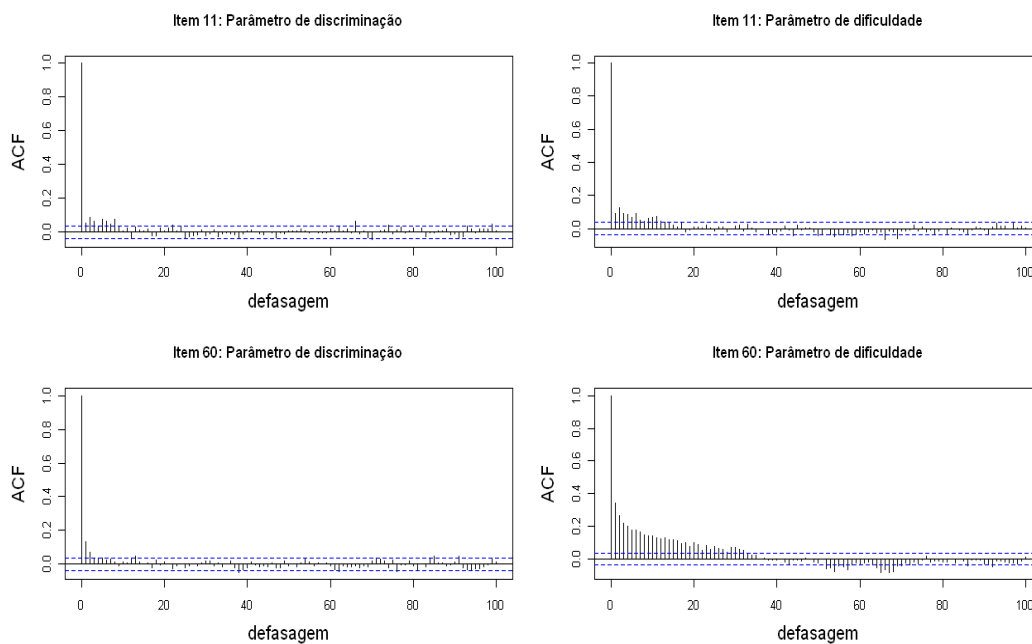


Figura 4.16: Correlogramas com espaçamento das amostras geradas para os parâmetros dos itens 11 e 60: Algoritmo ADGS

Os resultados anteriores sugerem que ambos algoritmos convergem após um certo número de iterações. Porém, tendo em vista os longos períodos de computação demandados pelos métodos MCMC, é interessante avaliarmos o tempo que estes levam para atingir a convergência. Neste trabalho utilizamos como critério de comparação o tamanho efetivo de amostra (ESS)<sup>1</sup>, como feito em Sahu (2002). O ESS é definido para cada parâmetro como o tamanho das amostras MCMC dividido pela autocorrelação,  $\gamma = 1 + 2 \sum_{k=1}^{\infty} \rho_k$ , onde  $\rho_k$  é a autocorrelação com defasagem  $k$ . Sendo assim, quanto maior for o valor de ESS, mais próxima de uma amostra aleatória da distribuição a posteriori conjunta estará a amostra gerada e mais representativa esta será, veja Gamerman and Lopes (2006). Seguindo Sahu (2002), utilizamos como estimativa de  $\gamma$  o limite superior  $(1 + \rho^*)/(1 - \rho^*)$  onde  $\rho^* = |\rho_1|$ . Nesse estudo utilizamos o conjunto de dados referente à avaliação do progresso matemático de estudantes de algumas escolas primárias de Londrina. Este conjunto de dados será descrito com mais detalhes no Capítulo 5. Realizamos 50000 iterações de cada algoritmo e descartamos as 20000 primeiras. Com as amostras obtidas, calculamos os ESS e também os ESS por minuto (ESS/m), dividindo o ESS pelo tempo de computação, para cada um dos parâmetros. A Tabela 4.3 mostra as médias dos ESS e ESS/m para os dois algoritmos considerando as quatro prioris adotadas. Na última coluna reportamos a razão entre os ESS/m dos dois algoritmos. Em termos do ESS, o algoritmo ADGS teve a melhor performance, exceto quando foi utilizada a priori 3. Em termos do ESS/m, observamos que o algoritmo ADGS também obteve a melhor performance, exceto quando foi utilizada a priori 4. Na maioria das situações o algoritmo ADGS foi mais rápido (quase duas vezes mais), principalmente, quando foi utilizado a priori 1 (que corresponde a priori de Jeffreys para os parâmetros populacionais).

Tabela 4.3: Performance dos algoritmos ADMHGS e ADGS no conjunto de dados referente à avaliação do progresso matemático de estudantes de escolas primárias de Londrina

| Priori | ADMHGS   |        | ADGS     |        | Razão |
|--------|----------|--------|----------|--------|-------|
|        | ESS      | ESS/m  | ESS      | ESS/m  |       |
| 1      | 4474,178 | 25,228 | 4771,270 | 36,361 | 1,441 |
| 2      | 4549,746 | 20,190 | 4724,984 | 36,014 | 1,784 |
| 3      | 5054,523 | 32,316 | 4955,711 | 35,533 | 1,010 |
| 4      | 4569,048 | 27,968 | 5103,164 | 23,491 | 0,840 |

<sup>1</sup>Sigla do inglês: *effective sample size*

### 4.3 Estudo de replicação

Nesta seção apresentamos os resultados de um estudo de replicação feito com intuito de avaliar a acurácia das estimativas e, também, o impacto de importantes fatores na recuperação dos parâmetros. Realizamos este estudo utilizando o algoritmo ADGS. Esta escolha foi baseada nos resultados obtidos na seção anterior, que apontaram este algoritmo como sendo o de melhor performance, em termos de convergência.

Para nossas inferências, consideramos amostras a posteriori de tamanho 1000, obtidas após 50000 iterações do algoritmo ADGS, descartando-se as 20000 primeiras e retendo valores a cada 30 iterações. Os fatores considerados com os níveis entre parênteses foram os seguintes: número de respondentes por grupo (NR) (1500, 3000), número de itens por grupo (NI) (20, 40), número de itens comuns (NIC) (25%, 50%), priori (P) (priori1, priori2, priori3, priori4) estas priors são as mesmas que foram definidas na subseção 4.2.1, vide Tabela 4.2. Resolvemos trabalhar com três grupos, nomeadamente 1, 2 e 3, sendo o grupo 1 o de referência. Incluímos mais um fator referente à distribuição dos traços latentes do grupo de referência e o chamamos de assimetria do grupo de referência (AGR) com níveis (1, 2, 3). Tais níveis contemplam as situações onde a assimetria da distribuição do grupo de referência é nula, fortemente positiva e fortemente negativa. Para isso escolhemos para  $\gamma_{\theta_1}$  os valores 0; 0,6 e -0,6; respectivamente. Os demais parâmetros populacionais foram fixados da seguinte forma: as médias e desvios-padrão foram fixados em  $\mu_{\theta} = (0, -1, 1)^{\top}$  e  $\sigma_{\theta} = (1; 0,8; 1,2)^{\top}$  para os grupos 1, 2 e 3, respectivamente. Os coeficientes de assimetria dos grupos 2 e 3 foram fixados em  $\gamma_{\theta_2} = 0,6$  e  $\gamma_{\theta_3} = -0,6$ . Ou seja, estudaremos as situação onde o grupo de referência é simétrico ou assimétrico e os demais grupos são assimétricos. A Tabela 4.4 resume os valores dos parâmetros populacionais escolhidos para cada grupo e cada nível do fator AGR.

Tabela 4.4: Parâmetros populacionais escolhidos para cada grupo para cada nível do fator AGR

|         |                     | AGR= 1 | AGR= 2 | AGR= 3 |
|---------|---------------------|--------|--------|--------|
| Grupo 1 | $\mu_{\theta_1}$    | 0      | 0      | 0      |
|         | $\sigma_{\theta_1}$ | 1      | 1      | 1      |
|         | $\gamma_{\theta_1}$ | 0      | 0,6    | -0,6   |
| Grupo 2 | $\mu_{\theta_2}$    | -1     | -1     | -1     |
|         | $\sigma_{\theta_2}$ | 0,8    | 0,8    | 0,8    |
|         | $\gamma_{\theta_2}$ | 0,6    | 0,6    | 0,6    |
| Grupo 3 | $\mu_{\theta_3}$    | 1      | 1      | 1      |
|         | $\sigma_{\theta_3}$ | 1,2    | 1,2    | 1,2    |
|         | $\gamma_{\theta_3}$ | -0,6   | -0,6   | -0,6   |

Os parâmetros dos itens foram fixados de maneira análoga ao que foi feito na simulação

da seção 4.2, mas diferentemente daquela situação, aqui tivemos que fixar 4 conjuntos de itens com tamanhos diferentes, tendo em vista as diferentes combinações de número de indivíduos, número de itens e número de itens comuns. Tais parâmetros podem ser vistos no Apêndice C nas Tabelas C.1, C.2, C.3 e C.4. Considerando todos esses fatores temos um total de 96 situações. Para cada uma das situações, foram simulados 10 conjuntos de respostas, segundo o modelo, que chamaremos de réplicas. Em cada réplica, obteve-se um conjunto de estimativas para todos os parâmetros.

De modo a comparar as diferentes situações, utilizamos estatísticas convenientes. Seja  $\vartheta_l \in (\theta_{jk}, a_i, b_i)$ , onde  $l$  é um índice conveniente ( $i, j$  ou  $k$ ) e  $\hat{\vartheta}_{lr}$  sua respectiva estimativa obtida na réplica  $r$ ,  $r = 1, \dots, R$ . Defina também  $\hat{\bar{\vartheta}}_l = \frac{1}{R} \sum_{r=1}^R \hat{\vartheta}_{lr}$ . A raiz quadrada do erro quadrático médio (REQM) e o viés relativo absoluto (VRA) são definidos da seguinte forma:

$$\begin{aligned} \text{REQM} &= \sqrt{\frac{1}{R} \sum_{r=1}^R (\hat{\vartheta}_{lr} - \vartheta_l)^2} \text{ e} \\ \text{VRA} &= \frac{|\hat{\bar{\vartheta}}_l - \vartheta_l|}{|\vartheta_l|}. \end{aligned}$$

Os resultados desse estudo foram resumidos através de gráficos das estatísticas REQM, VRA e das medianas de REQM que chamamos de MREQM, calculados para cada conjunto de parâmetros. Os gráficos das estatísticas REQM e VRA estão todos no formato boxplot, exceto para os parâmetros populacionais, que possuem apenas três observações em cada situação. Para esses parâmetros utilizamos gráficos da mediana com amplitude total (veja Apêndice B para mais detalhes). Os valores de REQM e VRA para as estimativas dos traços latentes tiveram muitos valores discrepantes, principalmente, os de VRA. Para garantir a comparabilidade entre os boxplots consideramos apenas os valores de  $\text{REQM} < 0,9$  e  $\text{VRA} < 0,7$ . Os gráficos das estimativas de REQM e VRA para todas as situações e todos os conjuntos de parâmetros podem ser vistos no apêndice B.

As Figuras 4.17 a 4.22 apresentam os gráficos de linha das estatísticas MREQM calculados da seguinte forma. Para os traços latentes tomamos a média dos REQM's para cada nível dos fatores NI e NIC. Para os demais parâmetros tomamos as medianas dos REQM's para cada nível dos fatores NR e NIC. Podemos ver que os resultados não diferiram, significativamente, para as quatro prioris e os três níveis do fator AGR, exceto para o coeficiente assimetria que teve uma maior oscilação entre as prioris. Além disso, os valores de MREQM dos traços latentes parecem diminuir com o aumento do número de itens e número de itens comuns. Para os demais parâmetros, os valores de MREQM diminuem com o aumento do número de respondentes e

também do número de itens comuns. Tal comportamento pode ser visto, também, nos gráficos de REQM e VAR. Observe, por exemplo, as Figuras B.16 (a), (b) e B.34 (a),(b) referentes aos traços latentes e as Figuras B.10 e B.28 (a) e (c) referentes aos parâmetros de discriminação.

Observando as Figuras B.7 a B.9 e as Figuras B.25 a B.27 referentes ao coeficiente de assimetria, vemos que os melhores resultados foram obtidos com a priori 1, exceto em três situações, vide Figuras B.7 (a), (b) e (f) e Figuras B.25 (a), (b) e (f). Ou seja, quando o grupo de referência é simétrico e o número de indivíduos é pequeno a priori 2 apresentou os melhores resultados na estimação dos coeficientes de assimetria.

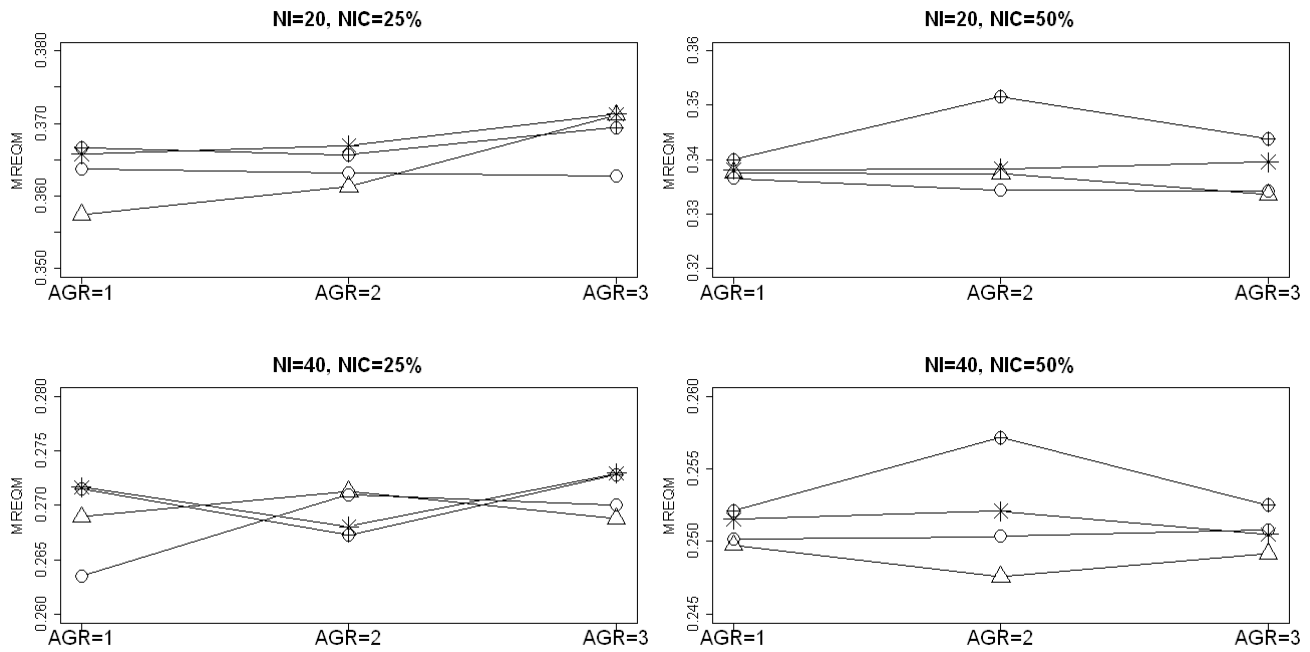


Figura 4.17: Mediana do REQM das estimativas dos traços latentes. Legenda: (—○— Priori 1, —△— Priori 2, —\*— Priori 3 e —⊕— Priori 4)

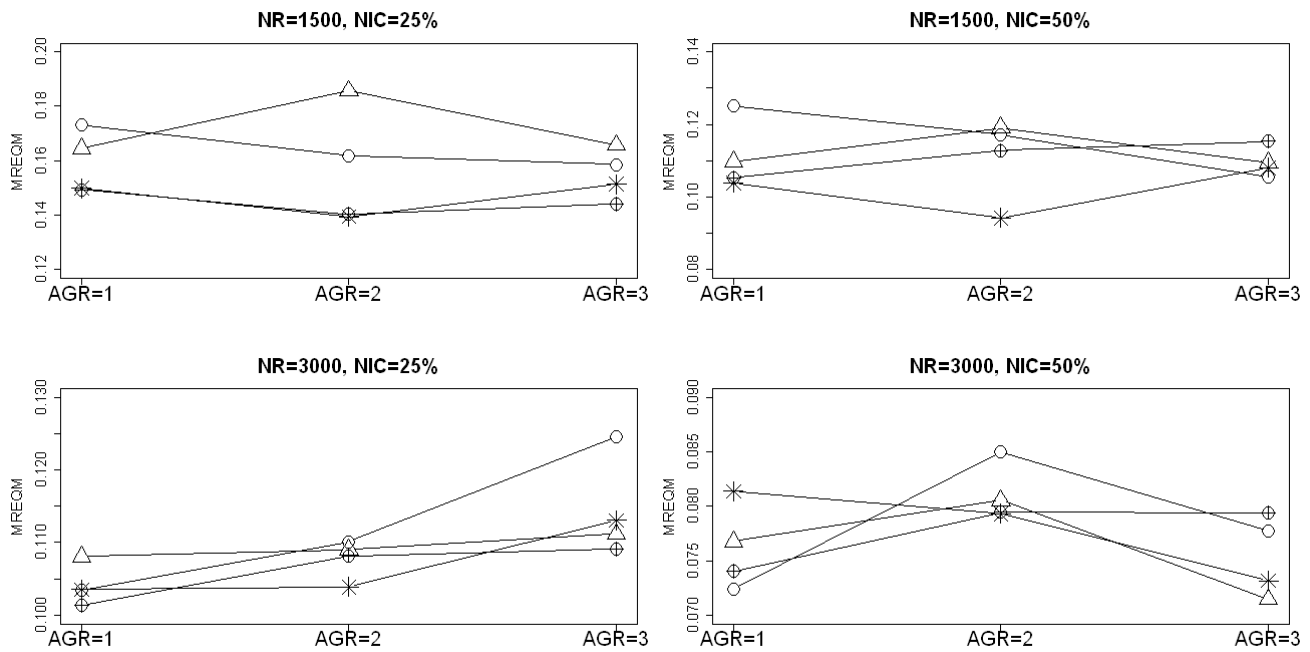


Figura 4.18: Mediana do REQM das estimativas dos parâmetros de discriminação. Legenda: (—○— Priori 1, —△— Priori 2, —\*— Priori 3 e —⊕— Priori 4)

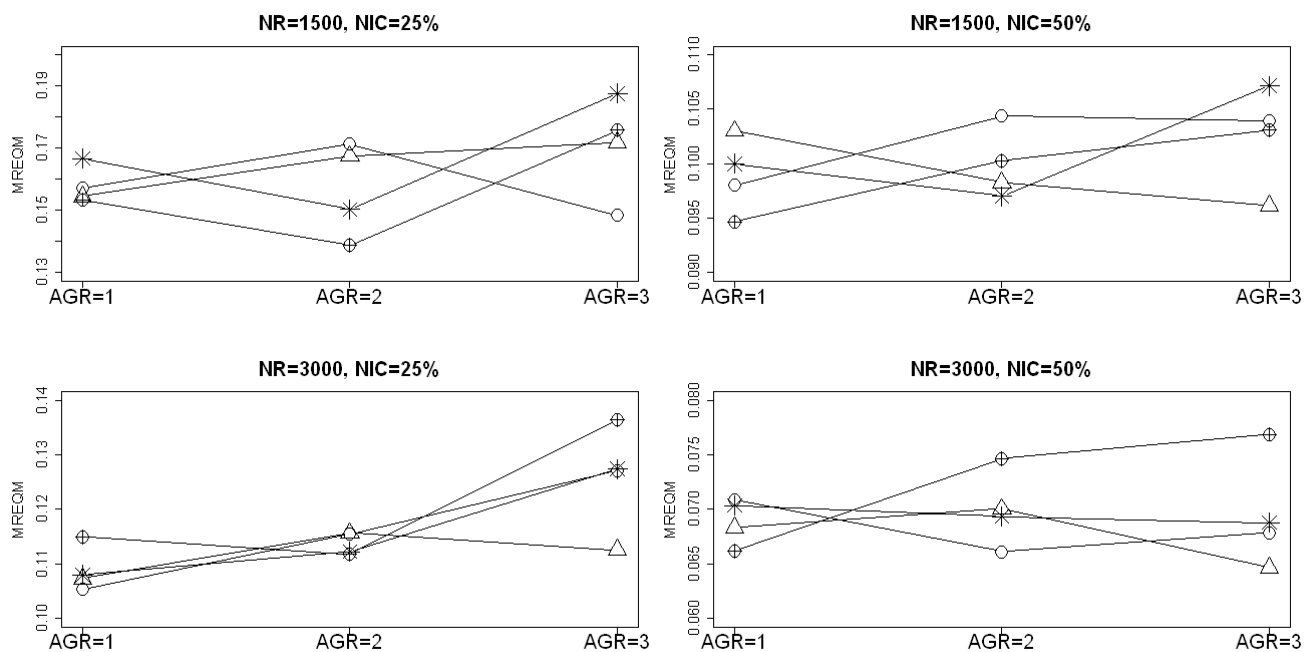


Figura 4.19: Mediana do REQM das estimativas dos parâmetros de dificuldade. Legenda: (—○— Priori 1, —△— Priori 2, —\*— Priori 3 e —⊕— Priori 4)

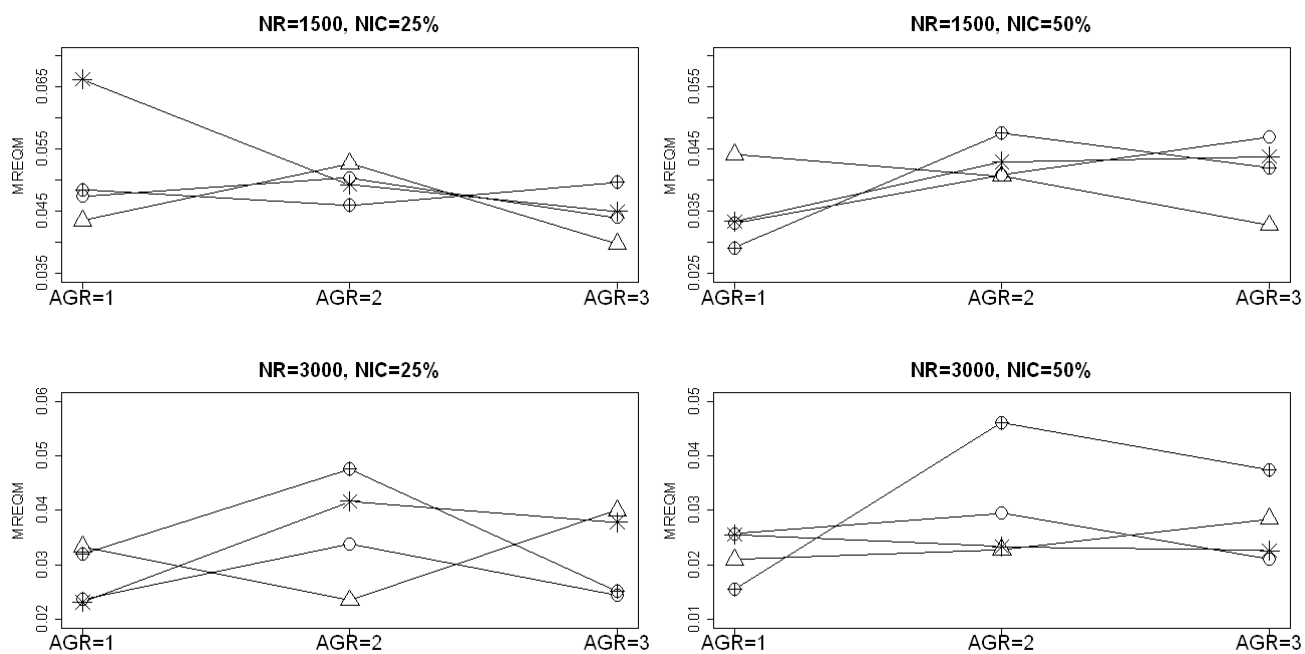


Figura 4.20: Mediana do REQM das estimativas das médias populacionais. Legenda: (—○— Priori 1, —△— Priori 2, —\*— Priori 3 e —⊕— Priori 4)

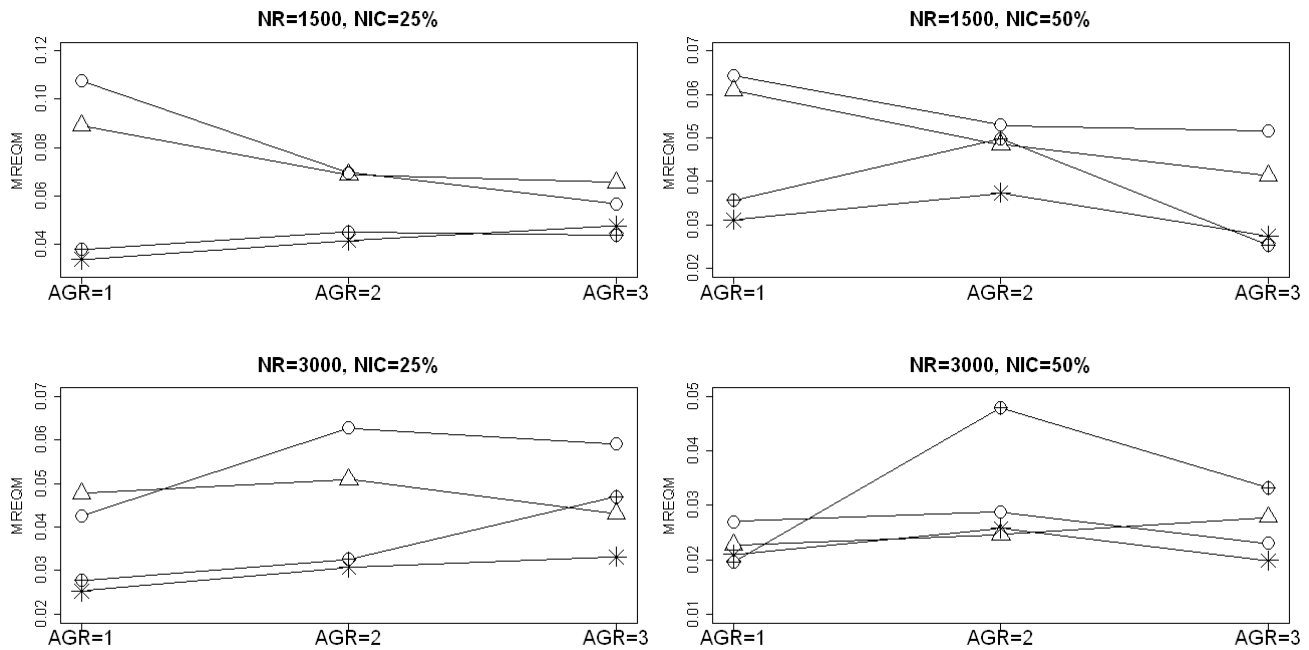


Figura 4.21: Mediana do REQM das estimativas dos desvios-padrão populacionais. Legenda: (—○— Priori 1, —△— Priori 2, —\*— Priori 3 e —⊕— Priori 4)

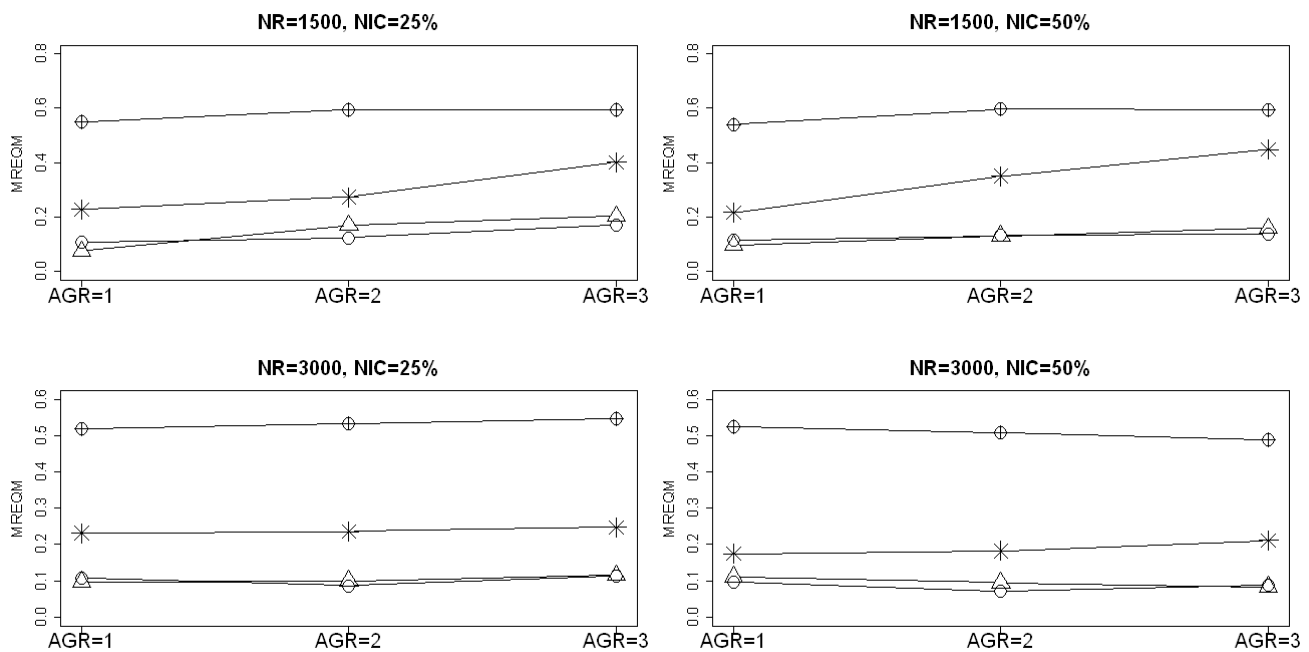


Figura 4.22: Mediana do REQM das estimativas dos coeficientes de assimetria populacionais. Legenda: (—○— Priori 1, —△— Priori 2, —\*— Priori 3 e —⊕— Priori 4)

Para avaliarmos o impacto de cada fator na precisão das estimativas dos parâmetros, fizemos uma análise de variância considerando como variável resposta,  $\ln(\text{REQM})$ ; veja De Ayala and Sava-Bolesta (1999) e DeMars (2003). Acreditamos que esta análise seja um complemento à análise gráfica, ajudando-nos a reforçar nossas conclusões. Considerando que temos muitas observações para cada combinação de níveis dos fatores, qualquer teste estatístico seria significativo. Também, algumas pressuposições da ANOVA, tais como, normalidade e homogeneidade, podem não ser verificadas considerando como resposta o logaritmo natural do REQM. Desta forma, consideramos a estatística  $\omega^2$  ao invés da estatística  $F$ , sem realizar qualquer teste estatístico. A estatística  $\omega^2$  é dada por

$$\omega^2 = \frac{SQ_{\text{fator}} - gl_{\text{fator}}QM_{\text{fator}}}{QM_{\text{resíduo}} + SQT}, \quad (4.1)$$

onde,  $SQ$  é a soma de quadrados,  $QM$  é o quadrado médio e  $gl$  são os graus de liberdade. As Tabelas 4.5 a 4.7 mostram as análise de variância do  $\ln(\text{REQM})$  das estimativas dos traços latentes, parâmetros de discriminação e parâmetros de dificuldade. Nestas mostramos apenas os efeitos principais. Todas as interações apresentaram valores de  $\omega^2$  menores do que  $10^{-4}$  e foram omitidas nestas tabelas, pois têm impacto praticamente nulo na precisão das estimativas. No entanto, o leitor pode ver as tabelas ANOVA completas no Apêndice C. Observando a Tabela 4.5, referente aos traços latentes, vemos que o número de itens por grupo é responsável por quase 20% da variabilidade de  $\ln(\text{REQM})$  seguido do número de itens comuns por grupo, que explica cerca de 1,3% da variabilidade. Com relação aos parâmetros de discriminação, vemos que o fator com maior impacto na precisão das estimativas é número de respondentes por grupo, seguido do número de itens comuns por grupo. Já para o parâmetro de dificuldade, a contribuição do fator NR é quase a igual a do fator NIC, sendo que o número de itens comuns por grupo explica quase 10% da variabilidade e o fator NR explica quase 7%. Vemos também que, o impacto dos fatores AGR e P foi irrisório nos três conjuntos de parâmetros analisados.

Tabela 4.5: ANOVA para os traços latentes

| Fonte de variação | gl     | SQ        | QM       | $\omega^2$ |
|-------------------|--------|-----------|----------|------------|
| NR                | 1      | 43,100    | 43,100   | 0,002      |
| NI                | 1      | 4781,160  | 4781,160 | 0,197      |
| NIC               | 1      | 334,860   | 334,860  | 0,014      |
| AGR               | 2      | 0,920     | 0,460    | <0,001     |
| P                 | 3      | 2,770     | 0,920    | <0,001     |
| Resíduos          | 215991 | 19122,790 | 0,089    |            |

Tabela 4.6: ANOVA para os parâmetros de discriminação

| Fonte de variação | gl   | SQ       | QM      | $\omega^2$ |
|-------------------|------|----------|---------|------------|
| NR                | 1    | 201,613  | 201,613 | 0,117      |
| NI                | 1    | 51,738   | 51,739  | 0,030      |
| NIC               | 1    | 161,045  | 161,045 | 0,093      |
| AGR               | 2    | 0,847    | 0,423   | <0,001     |
| P                 | 3    | 4,715    | 1,572   | 0,002      |
| Resíduos          | 6471 | 1306,068 | 0,202   |            |

Tabela 4.7: ANOVA para os parâmetros de dificuldade

| Fonte de variação | gl   | SQ       | QM      | $\omega^2$ |
|-------------------|------|----------|---------|------------|
| NR                | 1    | 191,632  | 191,632 | 0,068      |
| NI                | 1    | 50,454   | 50,454  | 0,018      |
| NIC               | 1    | 273,307  | 273,307 | 0,098      |
| AGR               | 2    | 4,313    | 2,156   | 0,001      |
| P                 | 3    | 0,308    | 0,103   | <0,001     |
| Resíduos          | 6471 | 2277,510 | 0,352   |            |

## 4.4 Comentários e conclusões

Todos nossos algoritmos foram implementados na linguagem orientada à objeto  $Ox^{TM}$ , veja <http://www.doornik.com/products.html>. A edição dos resultados (análises de variâncias, gráficos, etc) foi feita com o software R, veja <http://www.r-project.org/>. Vimos que, o algoritmo ADGS apresentou uma performance superior ao algoritmo ADMHGS, em termos de convergência. Excetuando-se o coeficiente de assimetria, vimos que as cadeias geradas por esse algoritmo convergiram após 1000 iterações. Para o coeficiente de assimetria notamos a necessidade de um período bem mais longo (20000 iterações). Reportamos na Tabela 4.8 os tempos de computação, para cada 1000 iterações do Algoritmo ADGS utilizando a priori 1, para diferentes combinações do número total de respondentes (NTR) e número total de itens (NTI). Os resultados foram obtidos em um computador com processador Intel(R) core(TM) i5 2,67 GHz, 4GB de memória RAM e sistema operacional de 64 Bits.

Tabela 4.8: Tempo de computação do algoritmo ADGS para cada 1000 iterações

| (NTR; NTI)  | Tempo de computação |
|-------------|---------------------|
| (1500; 40)  | 51s                 |
| (1500; 50)  | 1m10s               |
| (1500; 80)  | 1m42s               |
| (1500; 100) | 2m07s               |
| (3000; 40)  | 1m43s               |
| (3000; 50)  | 2m03s               |
| (3000; 80)  | 3m16s               |
| (3000; 100) | 3m38s               |

As estimativas dos coeficientes de assimetria sofreram maior influência da escolha da priori. As prioris 1 e 2 apresentaram os melhores resultados e dentre estas, a priori 1 apresentou os melhores resultados na grande maioria das situações. Dessa forma, o modelo de grupos múltiplos assimétrico com priori de Jeffreys (priori 1) recuperou bem todos os parâmetros. Com relação ao impacto de fatores na recuperação dos parâmetros concluimos que, o número de itens é o fator de maior impacto na estimação dos traços latentes, sendo que, o aumento do número de itens está relacionado com o aumento na precisão das estimativas. O número de respondentes é o que mais impacta na recuperação dos parâmetros de discriminação, sendo que o aumento destes, está relacionado com uma maior precisão nas estimativas. Já a precisão das estimativas dos parâmetros de dificuldade parecem não depender tanto do número de respondentes. Acreditamos que este fato esteja relacionado à própria definição de parâmetro de discriminação e parâmetro de dificuldade. Ou seja, é razoável pensar que o número de observações necessárias para se estimar bem a inclinação de uma curva, é maior do que o número de observações ne-

cessárias para se estimar bem a sua posição. Constatamos, também, que o número de itens comuns possui uma notada relevância aparecendo como o segundo fator de maior impacto na precisão das estimativas dos conjuntos de parâmetros analisados. Sendo que, o aumento destes, também está relacionado à estimativas mais precisas.

Não foi observado nenhuma mudança na precisão das estimativas quando variamos a assimetria da distribuição dos traços latentes do grupo de referência, com os demais grupos assimétricos. Ou seja, o modelo de grupos múltiplos assimétrico forneceu boas estimativas, independentemente da simetria ou assimetria do grupo de referência.

# Capítulo 5

## Análise de Dados Reais

### 5.1 Introdução

Neste Capítulo utilizamos o modelo de grupos múltiplos assimétrico para analisar um conjunto de dados referente a um estudo longitudinal realizado com alunos da primeira a quarta série do ensino fundamental de algumas escolas públicas da cidade de Londrina. Baseados nos resultados do Capítulo 4, escolhemos o algoritmo ADGS com priori de Jeffreys para realizar a estimação dos parâmetros. Avaliamos a qualidade de nosso ajuste utilizando diagnóstico baseado na distribuição preditiva de medidas de discrepância adequadas. O presente capítulo está dividido da seguinte forma: na seção 1.2 apresentamos algumas medidas de diagnóstico. Na seção 1.3 analisamos dados reais comparando os resultados obtidos com os modelo assimétrico e simétrico. Por último, na seção 1.3 apresentamos conclusões e comentários.

### 5.2 Medidas de diagnóstico

Em modelos bayesianos, uma forma de verificar a qualidade do ajuste é comparar a distribuição dos dados observados com a distribuição preditiva, veja Gelman et.al (2004). Para modelos de resposta ao item bayesianos existe, na literatura, diversos diagnósticos baseados em distribuições preditivas. Discussões gerais podem ser encontradas em, Stern and Sinharay (2005), Sinharay (2006), Sinharay et.al (2006), Fox (2004), Fox and Glas (2005) e Fox (2010). Azevedo et.al (2012) adaptaram alguns diagnósticos para o caso de grupos múltiplos. Tais diagnósticos são baseados em medidas de discrepância escolhidas de modo que se possa avaliar alguma suposição ou o ajuste geral do modelo. Neste trabalho utilizamos duas medidas de discrepância. Uma delas é baseada nos resíduos de Pearson. Este tipo de resíduo representa a diferença entre os valores observados e esperados dividida pelo desvio-padrão das observações. A outra será

definida a partir do resíduo componente do desvio (ou resíduo *deviance*), veja Paula (2010). Vamos agora definir algumas notações importantes para o desenvolvimento deste Capítulo.

Seja  $\mathbf{y}^{obs}$  a matriz de respostas observadas, e  $\mathbf{y}^{rep}$  a matriz de dados replicados gerados a partir da sua distribuição preditiva. A distribuição preditiva das respostas do grupo  $k$  é representada por

$$p(\mathbf{y}_k^{rep}|\mathbf{y}_k^{obs}) = \int p(\mathbf{y}_k^{rep}|\boldsymbol{\vartheta}_k)p(\boldsymbol{\vartheta}_k|\mathbf{y}_k^{obs})d\boldsymbol{\vartheta}_k, \quad (5.1)$$

onde  $\boldsymbol{\vartheta}_k$  denota o conjunto de parâmetros do modelo correspondentes ao grupo  $k$ . Gelman et al. (2004) sugerem vários gráficos para se comparar os dados replicados com os observados segundo uma medida de discrepância. Uma outra opção é o valor-p bayesiano definido como a probabilidade de observarmos dados replicados maiores do que, ou iguais aos dados observados, segundo a medida de discrepância adotada. Ou seja,

$$p_0(\mathbf{y}_k^{rep}|\mathbf{y}_k^{obs}) = P(D(\mathbf{y}_k^{rep}, \boldsymbol{\vartheta}_k) \geq D(\mathbf{y}_k^{obs}, \boldsymbol{\vartheta}_k) | \mathbf{y}_k^{obs}), \quad (5.2)$$

onde a probabilidade é tomada sob a posteriori conjunta de  $(\mathbf{y}_k^{rep}, \boldsymbol{\vartheta}_k)$ . Na prática, a distribuição preditiva é calculada utilizando simulação. Se tivermos  $T$  simulações da distribuição a posteriori de  $\boldsymbol{\vartheta}_k$  nós calculamos  $\mathbf{y}_k^{rep}$  da distribuição preditiva para cada  $\boldsymbol{\vartheta}_k$  simulado; teremos então uma amostra de tamanho  $T$  da distribuição a posteriori conjunta de  $(\mathbf{y}_k^{rep}, \boldsymbol{\vartheta}_k|\mathbf{y}_k^{obs})$ . A estimativa do valor-p é dada pela proporção de valores nas  $T$  simulações tal que,  $D((\mathbf{y}_k^{rep})^{(t)}, \boldsymbol{\vartheta}_k^{(t)}) \geq D((\mathbf{y}_k^{obs})^{(t)}, \boldsymbol{\vartheta}_k^{(t)})$ ,  $t = 1, \dots, T$ .

Béguin and Glas (2001), propõem um método gráfico para comparar a distribuição dos escores observados com a distribuição preditiva dos escores. Sob a estrutura de grupos múltiplos a distribuição dos escores pode ser avaliada por grupo.

A medida de discrepância baseada nos resíduos de Pearson para um grupo específico é definida como

$$D_p(\mathbf{y}_k) = \sum_l \frac{(n_{l,k} - E(n_{l,k}))^2}{V(n_{l,k})}, \quad (5.3)$$

em que  $n_{l,k}$  representa o número de indivíduos com escore  $l$  no grupo  $k$ . Ou seja,  $D_p(\mathbf{y}_k)$  é definida como o quadrado dos resíduos, o que corresponde a bem conhecida estatística qui-quadrado de Pearson. A esperança e variância de  $n_{l,k}$  são estimadas da seguinte forma: para cada conjunto de parâmetros simulados (traços latentes e parâmetros dos itens) um novo conjunto de respostas é também simulado e, a partir destes, são calculados os valores de  $n_{l,k}$ . Tomamos a média e variância sob todos  $n_{l,k}^{(t)}$  para toda simulação  $t = 1, \dots, T$ , escore  $l = 1, \dots, L$  e grupo

$k$ , em que  $L$  é o escore máximo observado. Esta medida também pode ser definida em nível populacional somando-se as medias de cada grupo, ou seja,  $D(\mathbf{y}) = \sum_k D(\mathbf{y}_k)$ .

Para o resíduo componente do desvio definimos a seguinte medida

$$D_v(\mathbf{y}_k) = \sum_l d(n_{l,k}), \quad (5.4)$$

em que  $d(n_{l,k})$  representa a contribuição da verossimilhança da observação. Na prática, a contribuição da log-verossimilhança também pode ser utilizada.

De maneira similar, a distribuição dos escores observados por item pode ser também avaliada por grupo. Considerando a estatística de Pearson, a medida de discrepância para um item  $i$  apresentado ao grupo  $k$  é definida como

$$D_p(\mathbf{y}_{ik}) = \sum_l \frac{(n_{l,ik} - E(n_{l,ik}))^2}{V(n_{l,ik})}, \quad (5.5)$$

em que,  $n_{l,ik}$  é o número de indivíduos com escore  $l$  que responderam corretamente ao item  $i$  no grupo  $k$ . A média e variância de  $n_{l,ik}$  são calculadas da mesma forma que é feito para  $n_{lk}$ . Para os resíduos componente do desvio, tomando a contribuição da log-verossimilhança temos

$$D_v(\mathbf{y}_i) = \sum_k d(n_k), \quad (5.6)$$

em que,

$$d(n_k) = \sum_{j \in \Omega(n_k)} \sqrt{2|\log(P_{ijk})|} - \sum_{j \notin \Omega(n_k)} \sqrt{2|\log(1 - P_{ijk})|}. \quad (5.7)$$

Essas medidas também podem ser generalizadas de modo a resumir todos os grupos. Sendo  $D$  uma das duas medidas de discrepância definidas temos

$$D(\mathbf{y}_i) = \sum_k D(\mathbf{y}_{ik}). \quad (5.8)$$

Para resumir todos os itens basta tomar

$$D(\mathbf{y}) = \sum_i D(\mathbf{y}_i). \quad (5.9)$$

### 5.3 Estudo longitudinal em Matemática

Nosso conjunto de dados faz parte de uma grande avaliação educacional longitudinal coordenada pela universidade de Exeter, Inglaterra. Vários países participam deste estudo, incluindo o Brasil através da Universidade Estadual de Londrina. O projeto IPMA (*International Project on Mathematical Attainment*) monitora o progresso matemático de alunos em escolas primárias.

No Brasil, 568 estudantes da primeira série do ensino fundamental foram selecionados de oito escolas, localizadas em diferentes lugares da cidade de Londrina. Das oito escolas selecionadas, seis são municipais e duas são estaduais. Os estudantes foram acompanhados até a quarta série. Durante a sua realização, vários estudantes abandonaram o estudo por diferentes motivos.

O presente conjunto de dados consiste do seguinte número de estudantes da primeira à quarta série: 556, 556, 401 e 294. Os estudantes da primeira série responderam a um teste de 20 itens, a segunda série respondeu a 40 itens incluindo os 20 itens do teste da primeira série. O mesmo ocorreu com a terceira e quarta série, ou seja, estes alunos também responderam a 40 itens incluindo os últimos 20 itens da série precedente mais 20 itens inéditos (veja Tabela C.8). Temos então um total de 80 itens.

#### 5.3.1 Estimação dos parâmetros e avaliação do ajuste

Azevedo et.al (2012) analisam este mesmo conjunto de dados utilizando o modelo de grupos múltiplos simétrico. Eles avaliam a existência de correlações entre as séries. Seus resultados indicaram correlação significativa apenas entre a segunda e quarta série (correlação de Pearson igual a 0,215). Portanto, parece ser razoável supormos que os traços latentes são não correlacionados ao longo das séries.

Ajustamos o modelo de grupos múltiplos assimétrico (MGMA) aos dados considerando cada série como um grupo. Ajustamos também o modelo de grupos múltiplos simétrico (MGM) com o objetivo de fazer uma comparação dos resultados. O MGM também foi ajustado através do algoritmo ADGS modificado no passo 5.3, veja Apêndice A.2. Nas Figuras 5.1 e 5.2, são apresentadas as distribuições dos escores preditos e observados em cada série, segundo os modelos assimétrico e simétrico, respectivamente. Podemos ver que em cada série os escores preditos estão bem próximos dos observados e estes encontram-se dentro da região de credibilidade. Isso sugere que os modelos estão bem ajustados aos dados. Podemos ver, também, que nem todas as distribuições dos escores são simétricas, principalmente, a da segunda série. Isso sugere que as distribuições dos traços latentes de alguns grupos podem ser assimétricas.

As estimativas dos parâmetros populacionais para os modelos simétrico e assimétrico podem ser vistas na Tabela 5.1. Nesta podemos ver as estimativas dos parâmetros (média a

posteriori), os erros-padrão (EP), intervalos de credibilidade de 95% (IC (95%)) e HPD's de 95%. Denotamos por  $\mu_{\theta_k}$ ,  $\sigma_{\theta_k}$  e  $\gamma_{\theta_k}$ ,  $k = 1, \dots, 4$  os parâmetros populacionais da primeira a quarta série. Vemos que o MGMA confirmou nossas suspeitas a respeito da assimetria da distribuição dos traços latentes. Os resultados indicam que há uma forte assimetria à esquerda na distribuição dos traços latentes da segunda série (coeficiente de assimetria igual a -0,805). Já os traços latentes da primeira e terceira séries, apresentaram fraca assimetria à esquerda e os da quarta série apresentaram fraca assimetria à direita. A média e o desvio padrão da primeira série foram fixadas em zero e um, respectivamente, para identificação da escala. Na Figura 5.3 podemos ver as estimativas dos parâmetros populacionais com seus respectivos intervalos de credibilidade obtidas com os dois modelos. Vemos que as estimativas das médias e desvios-padrão obtidas com o MGMA são um pouco menores, em relação as estimativas do MGM, principalmente nos grupos 2 e 3. Ou seja, o modelo que ignora a assimetria, pode estar superestimando os parâmetros populacionais.

Pudemos ver que houve um incremento na proficiência média dos alunos da primeira até a quarta série. Isso indica uma melhora significativa na performance dos estudante a cada ano. Os desvios-padrão da terceira e quarta série são menores comparados aos da primeira e segunda série. Significa que os estudantes estão mais parecidos nos dois últimos anos e essa semelhança é maior do que a sugerida pelo modelo simétrico. Ou seja, após dois anos no projeto, os alunos parecem estar mais parecidos com relação a sua performance matemática.

As medidas de discrepância (5.5) e (5.6) foram utilizadas para avaliar o ajuste dos parâmetros dos itens. A Figura 5.4 mostra os valores-p bayesianos para cada item e para os dois modelos ajustados. A estatística *deviance* parece ser mais conservadora no sentido de detectar itens que não se ajustaram bem aos dados. Observando os valores-p baseados na estatística qui-quadrado vemos que, de modo geral os itens estão devidamente ajustados com exceção dos itens, 19, 39 e 57 que foram apontados pelos dois modelos. O item 23 parece não ter se ajustado bem segundo o MGM com  $p = 0,078$  ajustando-se melhor segundo o MGMA ( $p = 0,178$ ). O ajuste "pobre" desses itens pode ser devido a mau especificação da função de resposta ao item, má formulação do item, etc.

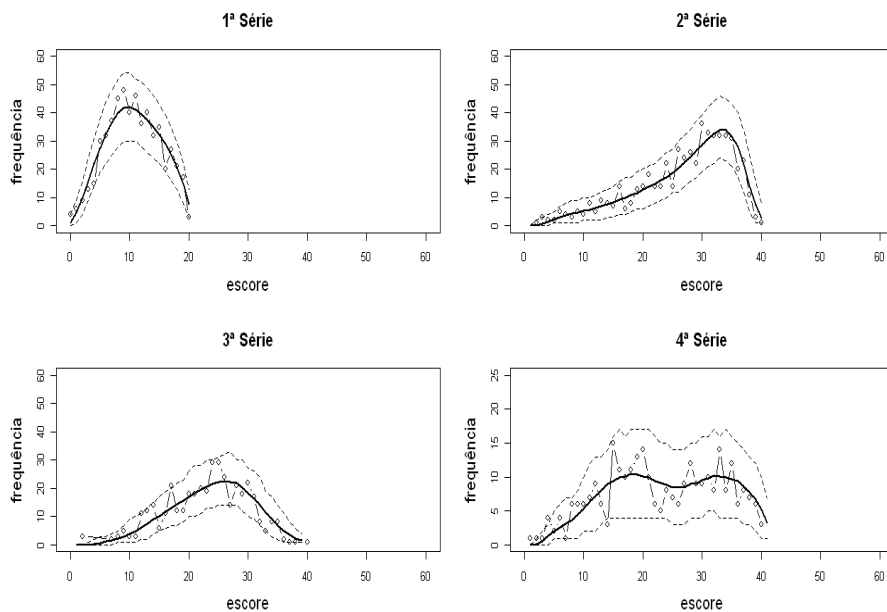


Figura 5.1: Distribuição dos escores segundo o MGMA. Legenda: linha cheia: escores preditos,  $-o-$  escores observados,  $- - -$  intervalo de credibilidade de 95%

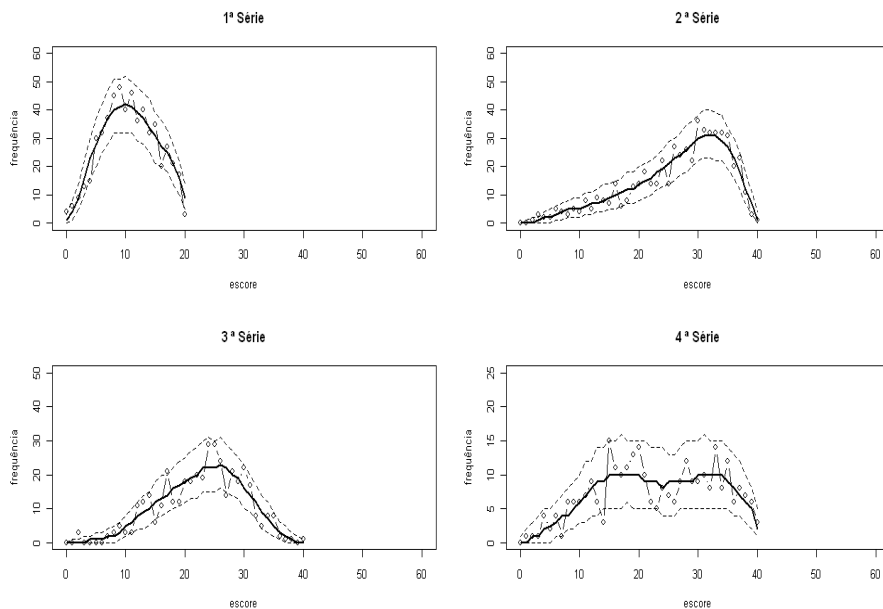


Figura 5.2: Distribuição dos escores segundo o MGM. Legenda: linha cheia: escores preditos,  $-o-$  escores observados,  $- - -$  intervalo de credibilidade de 95%

As Figuras 5.5 e 5.6 mostram as estimativas (médias a posteriori) dos parâmetros dos itens com intervalos de credibilidade de 95% para o MGMA e MGM. Podemos ver que os 20 primeiros

Tabela 5.1: Resumo a posteriori dos parâmetros populacionais de cada série segundo os modelos simétrico e assimétrico

| Parâmetros          | Assimétrico |       |          |        |           |        | Simétrico |       |          |       |           |       |
|---------------------|-------------|-------|----------|--------|-----------|--------|-----------|-------|----------|-------|-----------|-------|
|                     | Média       | EP    | IC (95%) |        | HPD (95%) |        | Média     | EP    | IC (95%) |       | HPD (95%) |       |
| $\mu_{\theta_1}$    | 0           | -     | -        | -      | -         | -      | 0         | -     | -        | -     | -         | -     |
| $\mu_{\theta_2}$    | 1,049       | 0,050 | 0,956    | 1,150  | 0,958     | 1,153  | 1,078     | 0,054 | 0,974    | 1,186 | 0,981     | 1,193 |
| $\mu_{\theta_3}$    | 1,427       | 0,057 | 1,317    | 1,539  | 1,316     | 1,536  | 1,463     | 0,060 | 1,356    | 1,588 | 1,356     | 1,589 |
| $\mu_{\theta_4}$    | 2,093       | 0,088 | 1,926    | 2,270  | 1,917     | 2,258  | 2,156     | 0,091 | 1,986    | 2,346 | 1,996     | 2,354 |
| $\sigma_{\theta_1}$ | 1           | -     | -        | -      | -         | -      | 1         | -     | -        | -     | -         | -     |
| $\sigma_{\theta_2}$ | 0,828       | 0,045 | 0,740    | 0,917  | 0,750     | 0,923  | 0,887     | 0,042 | 0,811    | 0,970 | 0,810     | 0,967 |
| $\sigma_{\theta_3}$ | 0,588       | 0,036 | 0,520    | 0,659  | 0,519     | 0,659  | 0,623     | 0,037 | 0,556    | 0,703 | 0,555     | 0,702 |
| $\sigma_{\theta_4}$ | 0,744       | 0,055 | 0,642    | 0,869  | 0,636     | 0,853  | 0,772     | 0,056 | 0,673    | 0,892 | 0,679     | 0,895 |
| $\gamma_{\theta_1}$ | -0,283      | 0,153 | -0,587   | -0,001 | -0,552    | 0,001  | 0,000     | -     | -        | -     | -         | -     |
| $\gamma_{\theta_2}$ | -0,805      | 0,092 | -0,960   | -0,605 | -0,975    | -0,630 | 0,000     | -     | -        | -     | -         | -     |
| $\gamma_{\theta_3}$ | -0,160      | 0,160 | -0,484   | 0,045  | -0,448    | 0,060  | 0,000     | -     | -        | -     | -         | -     |
| $\gamma_{\theta_4}$ | 0,140       | 0,184 | -0,079   | 0,569  | -0,098    | 0,537  | 0,000     | -     | -        | -     | -         | -     |

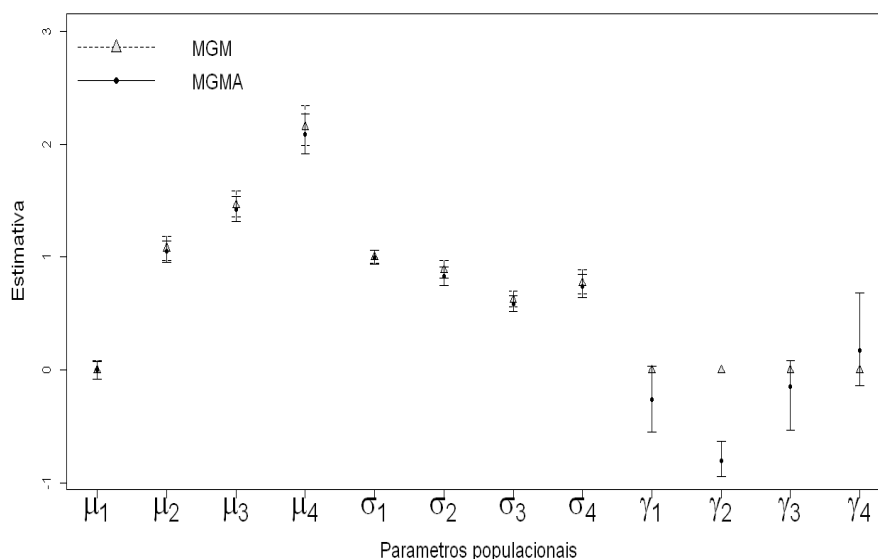


Figura 5.3: Estimativas dos parâmetros populacionais com intervalos de credibilidade de 95% para o MGM e MGMA

itens, que foram respondidos pelos estudantes da primeira e segunda série foram os mais fáceis e pra estes não notamos diferenças significativas entre os dois modelos. Alguns itens, dentre os itens 40 a 60, que foram respondidos pela terceira e quarta série parecem ter sido muito fáceis, uma vez que estes têm dificuldade comparável aos itens que foram respondidos pela primeira série. Os vinte últimos itens foram respondidos somente pela quarta série, e estes são os mais difíceis. No entanto, esta dificuldade é menor do que a sugerida pelo modelo simétrico. Ou seja, os estudantes nessa série precisaram de mais conhecimento matemático para responder aos itens. Vemos que tanto as estimativas dos parâmetros de dificuldade, quanto as estimativas dos parâmetros de discriminação crescem ao longo das séries. Ou seja, os itens mais difíceis

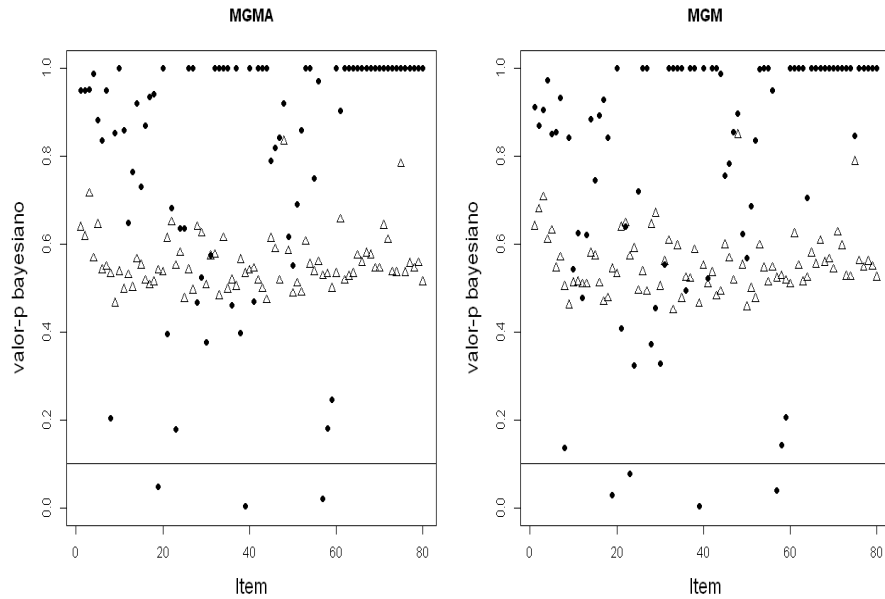


Figura 5.4: Valores-p Bayesianos correspondentes as medidas de discrepância baseadas nas estatísticas de Pearson e *deviance*. Legenda: ● Estatística qui-quadrado, △ Estatística *Deviance*

são também os mais discriminantes. Além disso, vemos que o MGMA sugeriu que o teste teve um poder de discriminação mais do que o sugerido pelo MGM. Todas as estimativas dos parâmetros dos itens segundo o MGMA podem ser vistas na Tabela C.8.

Nas Figuras 5.7 a 5.10 apresentamos uma comparação entre as estimativas dos traços latentes obtidas segundo os modelos simétrico e assimétrico. Na Figura 5.7 podemos que as estimativas não diferem significativamente, exceto na segunda série, onde notamos alguma diferença para os valores mais altos da proficiência. A Figura 5.8 apresenta as distribuições dos traços latentes estimados em forma histogramas alisados segundo os dois modelos. Como já havíamos mencionado, houve um aumento da proficiência média dos alunos ao longo das séries. Já nas Figuras 5.9 e 5.10 apresentamos as distribuições dos traços latentes estimados para cada série e cada modelo com as suas respectivas curvas teóricas. Podemos ver que o MGMA caracteriza melhor a distribuição dos traços latentes principalmente na presença de forte assimetria.

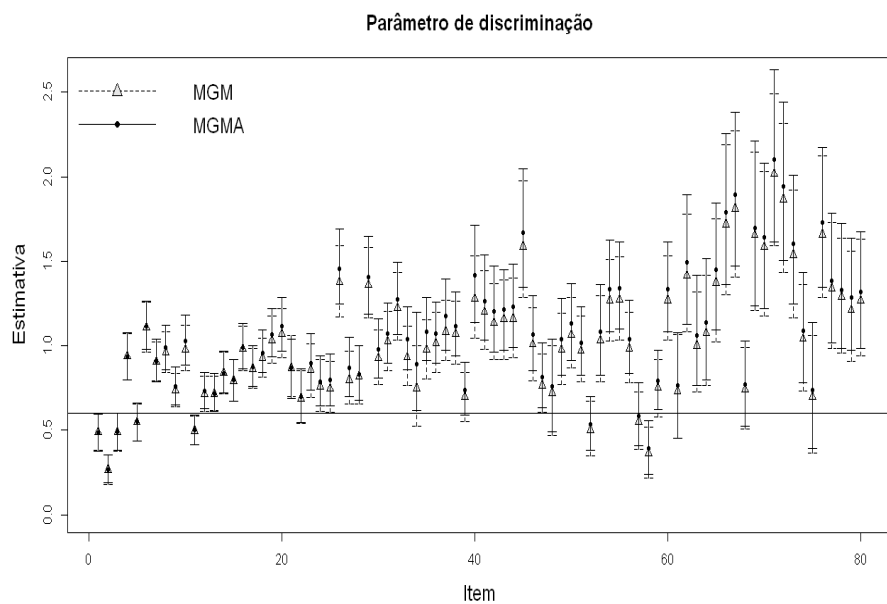


Figura 5.5: Estimativas dos parâmetros de discriminação com intervalos de credibilidade de 95%

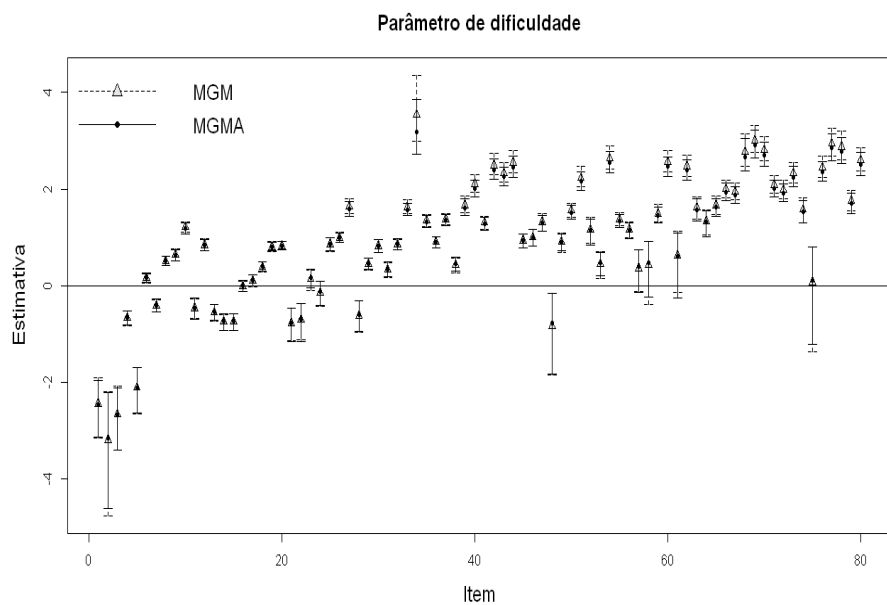


Figura 5.6: Estimativas dos parâmetros de dificuldade com intervalos de credibilidade de 95%

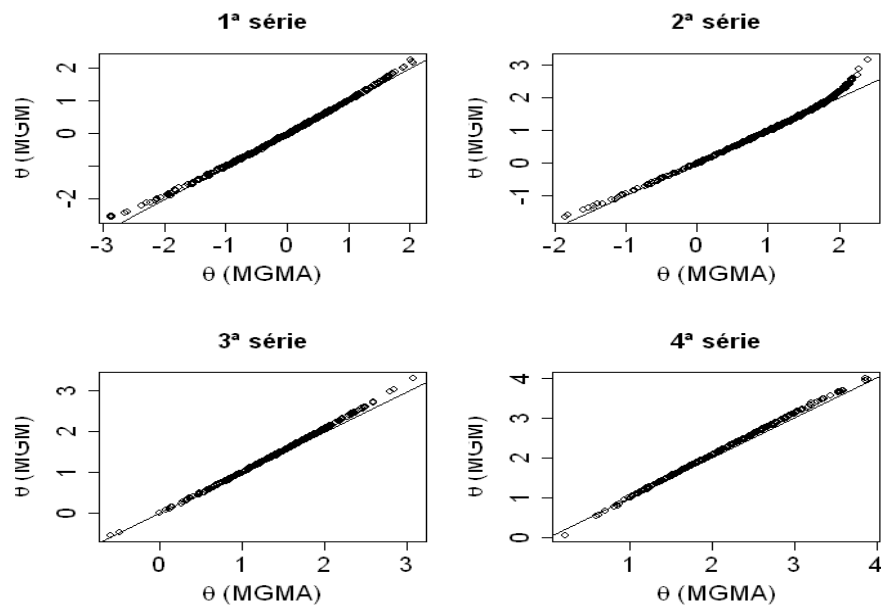


Figura 5.7: Gráfico de dispersão das estimativas dos traços latentes segundo o MGMA e MGM

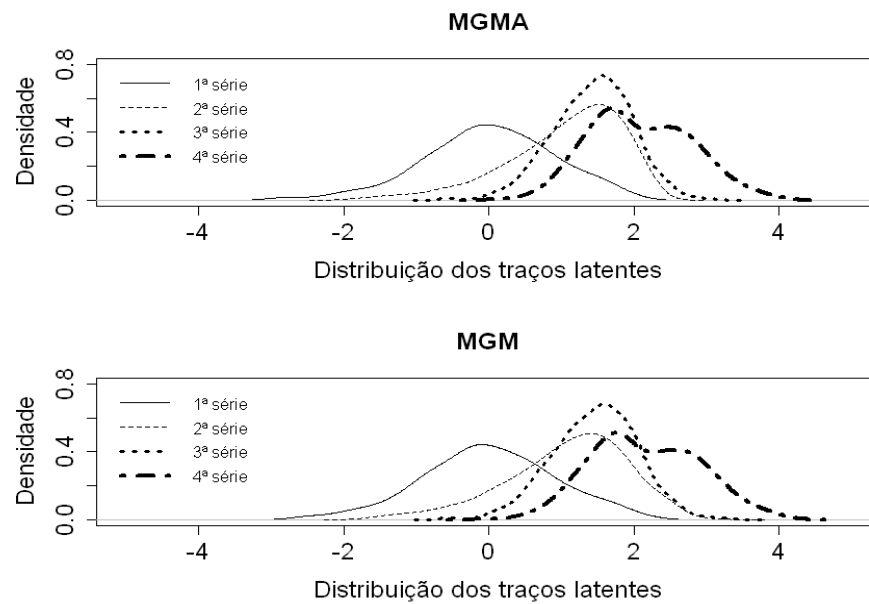


Figura 5.8: Distribuições dos traços latentes por série em escala comum

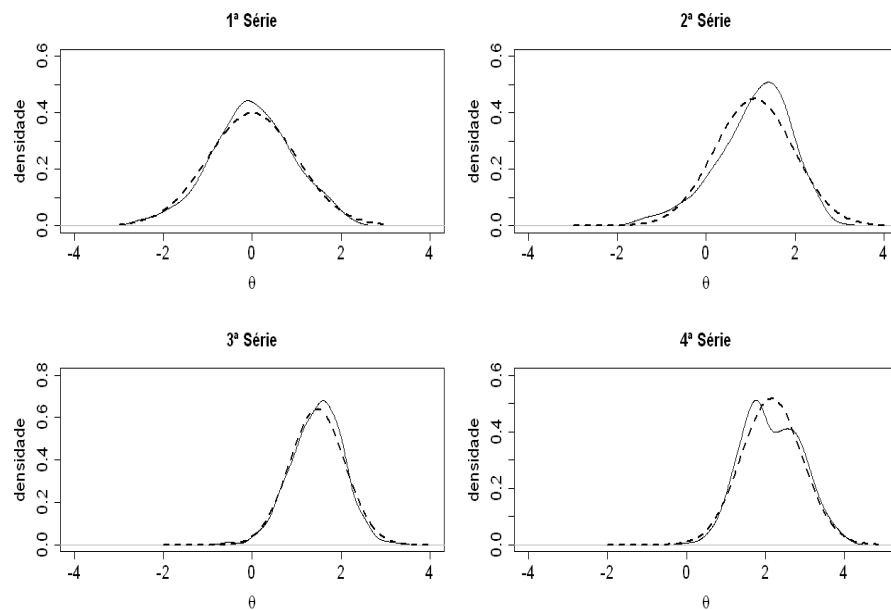


Figura 5.9: Distribuições dos traços latentes com curvas teóricas segundo o modelo simétrico. Legenda: (— curva estimada, - - - curva teórica)

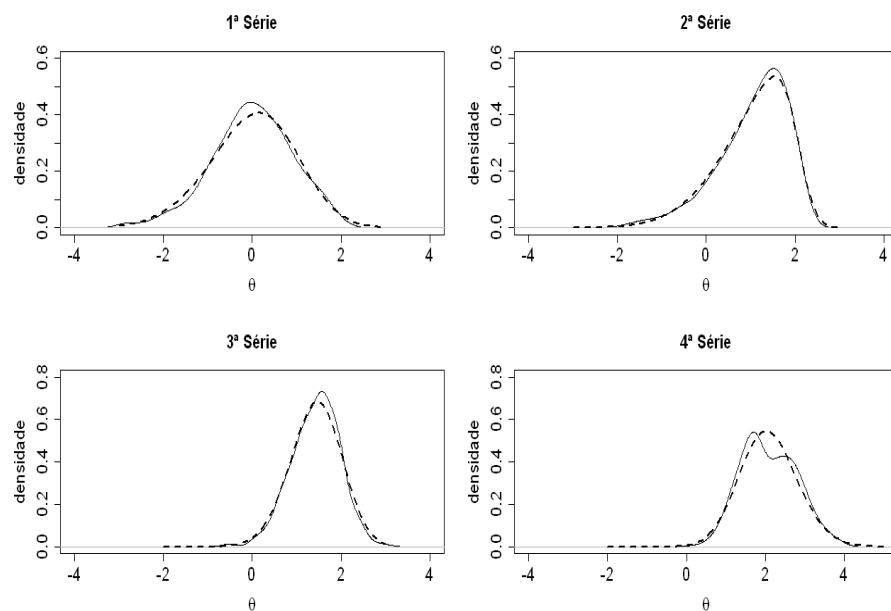


Figura 5.10: Distribuições dos traços latentes com curvas teóricas segundo o modelo assimétrico. Legenda: (— curva estimada, - - - curva teórica)

De modo a comparar o ajuste global dos modelos utilizamos as estatísticas de ajuste consideradas por Bazán et al. (2006) e Azevedo et al. (2011). Tais estatísticas foram estudadas em detalhes por Spiegelhalter et al. (2002). São elas o Critério de informação deviance (DIC), e os valores esperados do Critério de informação de Akaike (EAIC) e do Critério de informação bayesiano (EBIC). Esses critérios são baseados na estatística  $\rho_D$  definida da seguinte forma:

$$\rho_D = \overline{D(\vartheta)} - D(\bar{\vartheta}), \quad (5.10)$$

em que,  $\vartheta$  representa algum parâmetro de interesse e  $D(\vartheta) = -2(\log\text{-verossimilhança})$ . No nosso caso temos que

$$D(\vartheta) = -2L(\theta_{..}, \zeta_{..}, \boldsymbol{\eta}_{\theta_{..}}) = -2L(\theta_{..}, \zeta_{..})P(\theta_{..}|\boldsymbol{\eta}_{\theta_{..}}). \quad (5.11)$$

Na prática, podemos utilizar amostras MCMC para estimar  $\rho_D$ . Tendo-se  $T$  simulações de cada um dos parâmetros, a estatística  $\overline{D(\vartheta)}$  é estimada pela média das estatísticas avaliadas em cada uma das simulações. Ou seja,

$$\widehat{\overline{D(\vartheta)}} = \frac{1}{T} \sum_{t=1}^T D(\vartheta^{(t)}) \quad \forall \quad t = 1, \dots, T. \quad (5.12)$$

Já para estimar  $D(\bar{\vartheta})$ , utilizamos (mas não necessariamente) as médias a posteriori dos parâmetros. As medianas ou modas a posteriori podem também ser usadas. Dessa forma, as estimativas de DIC, EAIC e EBIC são dadas por

$$\widehat{\text{DIC}} = D(\bar{\vartheta}) + 2\hat{\rho}_D, \quad (5.13)$$

$$\widehat{\text{EAIC}} = \overline{D(\vartheta)} + 2p, \quad (5.14)$$

$$\widehat{\text{EBIC}} = \overline{D(\vartheta)} + p \log(N), \quad (5.15)$$

em que  $p$  é número de parâmetros e  $N$  é o número de observações, que no nosso caso, corresponde ao número total de respostas, ou seja,  $N = n \times I$ . Spiegelhalter et al. (2002) sugerem utilizar  $\hat{\rho}_D$  ao invés de  $p$ . Segundo eles, isto pode dar ainda mais potencial aos critérios. Seguimos esta sugestão neste trabalho. Os resultados podem ser vistos na Tabela 5.2. Estes indicam que o MGMA não só se ajustou devidamente aos dados e foi melhor do que o MGM, segundo todos os critérios de comparação.

Tabela 5.2: Critérios para comparação de modelos

| Modelo | $\widehat{D(\vartheta)}$ | $\widehat{D(\vartheta)}$ | $\widehat{\rho}_D$ | $\widehat{\text{DIC}}$ | $\widehat{\text{EAIC}}$ | $\widehat{\text{EBIC}}$ |
|--------|--------------------------|--------------------------|--------------------|------------------------|-------------------------|-------------------------|
| MGM    | 58103,00                 | 60073,00                 | 1969,80            | 62042,60               | 64012,60                | 83477,08                |
| MGMA   | 57873,00                 | 59836,00                 | 1962,60            | 61798,20               | 63761,20                | 83154,54                |

## 5.4 Conclusões e comentários

Na análise dos dados referente ao conhecimento matemático de alunos do ensino fundamental, o ajuste do modelo de grupos múltiplos assimétrico foi melhor do que o ajuste do modelo de grupos múltiplos usual. O MGMA detectou assimetria em 3 dos 4 grupos estudados. O diagnóstico desenvolvido na seção 1.2 foi utilizado para verificar o ajuste dos parâmetros dos itens e traços latentes. O valor-p bayesiano mostrou ser uma boa ferramenta de diagnóstico para verificação de ajuste dos itens.



# Capítulo 6

## Comentários finais e sugestões para futuros trabalhos

Nesse capítulo adicionamos mais alguns comentários a respeito da nossa metodologia e discutimos algumas possíveis extensões.

### 6.1 Comentários finais

O modelo desenvolvido neste trabalho pode ser aplicado em um grande número de situações reais, uma vez que, é comum observar, em estudos com grupos múltiplos, comportamentos diferentes entre as distribuições dos traços latentes. Comportamentos assimétricos ocorrem, muitas vezes, quando os indivíduos são selecionados de acordo com resultados de pré-teste, veja Genton (2004), ou quando são submetidos à diferentes programas de ensino, por exemplo. Em estudos longitudinais, onde pode ocorrer a eliminação ou inserção de indivíduos, a distribuição dos traços latentes pode se alterar ao longo do tempo. O conjunto de dados analisado no Capítulo 5 é um exemplo típico dessa situação. Portanto, o uso de distribuições assimétricas para representar o comportamento dos traços latentes, amplia o número de aplicações do modelo de grupos múltiplos e nos permite obter resultados mais fidedignos. A utilização de outras distribuições assimétricas, como por exemplo,  $t$  assimétrica poderia ser estudada. Com a distribuição  $t$  assimétrica, contemplaríamos também, os casos onde a distribuição dos traços latentes possui caudas pesadas.

Alguns aspectos que não foram explorados neste trabalho, bem como possíveis extensões, são discutidos a seguir.

## 6.2 Sugestões para futuros trabalhos

### 6.2.1 Uso de outras funções de resposta ao item

O modelo probito de 2 parâmetros, não representa de modo satisfatório a probabilidade de acerto de indivíduos com habilidade muito baixa, em itens em que o acerto ao acaso é possível. Uma abordagem mais apropriada é a utilização do modelo probito de 3 parâmetros. Uma importante extensão de nosso modelo seria feita com a inclusão do parâmetro de acerto casual na FRI utilizando a metodologia de dados aumentados desenvolvida em Sahu (2002) e Béguin and Glas (2001).

Por outro lado, em testes com itens de múltipla escolha ou dissertativos, pode ser mais apropriado utilizar modelos de resposta policotômicas. Ou seja, utilizar FRI's que contemplem a probabilidade de escolha de cada categoria, sejam elas nominais ou ordinais. Considerando um único grupo, nos modelos policotômicos, a probabilidade de escolha da categoria  $s$  do item  $i$  pelo indivíduo  $j$ , pode ser representada pelo modelo Bernoulli multivariado, ou seja,

$$P(\mathbf{Y}_{ij.} = \mathbf{y}_{ij.} | \theta, \boldsymbol{\zeta}_i) = \prod_{s=1}^{m_i} P_{ijs}^{y_{ijs}}, \quad (6.1)$$

em que,  $P_{ijs}$  é uma função de resposta ao item relacionada a algum modelo policotômico. Os modelos policotômicos assimétricos seriam baseados na seguinte suposição

$$\theta_j | \boldsymbol{\eta}_\theta \sim \text{NA}_c(\mu_\theta, \sigma_\theta, \gamma_\theta). \quad (6.2)$$

Patz and Junker (1999a) e Azevedo and Andrade (2010), propõem um amostrador de *Gibbs* com passos de *Metropolis-Hastings* para estimação dos modelos de crédito parcial generalizado e resposta nominal, respectivamente. No caso assimétrico, podemos usar, por exemplo, a priori de Jeffreys sob estrutura hierárquica para os parâmetros populacionais. Como vimos, esta escolha nos permite encontrar as condicionais completas dos parâmetros populacionais com forma conhecida, vide Capítulo 3 subseção 3.3.2.

Poderíamos também utilizar as FRI's assimétricas propostas por Bazán (2005). Dessa forma teríamos um modelo que permitiria assimetria nos itens e também nos traços latentes.

### 6.2.2 Outras classes de modelos de resposta ao item

Outras classes de modelos de potencial interesse na TRI, poderiam também ser estendidas para o caso assimétrico. Algumas classes importantes são: a classe dos modelos multidimensionais, veja Nojosa (2001), os quais assumem que é necessário mais de um traço latente para

modelar apropriadamente as probabilidades de resposta; os Modelos Multivariados, veja Matos (2001), que consideram estruturas de dependência dos traços latentes medidos em diferentes áreas de conhecimento; Modelos longitudinais, veja Azevedo (2008), Azevedo et al (2011 c), Tavares (2001), Andrade e Tavares (2005) e Tavares e Andrade (2006), que modelam dependências longitudinais dos traços latentes e os Modelos Multiníveis, os quais assumem que as unidades amostrais pertencem a diferentes grupos e estes, por sua vez, podem pertencer ou não a outros tipos de grupos. A distribuição normal assimétrica centralizada multivariada, veja Arellano-Valle and Azzalini (2008), poderia ser utilizada para modelar a dependência entre os traços latentes. Por exemplo, o modelo probito de 2 parâmetros longitudinal de grupos múltiplos teria a seguinte estrutura

**Modelo de resposta ao item (parte funcional)**

$$\begin{aligned} Y_{ijkt} | \theta_{jkt}, \zeta_i &\sim \text{Bernoulli}(P_{ijkt}), \\ P_{ijkt} = P(Y_{ijkt} | \theta_{jkt}, \zeta_i) &= \Phi(a_i \theta_{jkt} - b_i). \end{aligned} \quad (6.3)$$

**Distribuição Latente (parte estrutural)**

$$\theta_{jk.} | \eta_{\theta_k} \sim \text{NA}_c(\mu_{\theta_k}, \Sigma_{\theta_k}, \gamma_{\theta_k}), \quad (6.4)$$

em que  $Y_{ijkt}$  denota a resposta do indivíduo  $j$  da população  $k$  ao item  $i$ , no instante  $t$  e  $\text{NA}_c$  denota a distribuição normal assimétrica centralizada multivariada de parâmetros  $\mu_{\theta_k}, \Sigma_{\theta_k}$  e  $\gamma_{\theta_k}$ . Azevedo (2008), desenvolveu um amostrador de *Gibbs* sob a estrutura de dados aumentados que pode ser estendido para estimação do modelo longitudinal assimétrico.

Matos (2008), desenvolve um modelo probito de 3 parâmetros multidimensional com distribuições  $t$  assimétricas para os traços latentes, considerando um único grupo. Uma extensão interessante, seria a adaptação deste modelo para o caso de grupos múltiplos. Essa adaptação seria semelhante a que foi feita por Béguin and Glas (2001), para o modelo multidimensional simétrico.

Em termos de dados aumentados e considerando alguma FRI binária, seguindo Albert (1992) e ou Fox (2010) o modelo de grupos múltiplos assimétrico de 2 parâmetros com estrutura multinível é apresentado por,

$$Z_{ijk} = \mathbf{X}_{ijk} \beta_{ik} + a_{ik} \theta_{jk} - b_{ik} + \xi_{ijk} \quad (6.5)$$

$$\theta_{jk} = \mathbf{W}_{jk} \delta_k + \epsilon_{jk} \quad (6.6)$$

$$\delta_k = \mathbf{T}_k \tau + \mathbf{u}_k, \quad (6.7)$$

em que,  $\epsilon_{jk}$  possui distribuição normal assimétrica centralizada de parâmetros zero,  $\sigma_{\theta_k}$  e  $\gamma_{\theta_k}$ ; e  $\delta_k$  possui distribuição normal assimétrica centralizada multivariada com médias zero, matriz de covariância  $\boldsymbol{\delta}$  e vetor de coeficientes de assimetria  $\boldsymbol{\gamma}_{\delta}$ . Ou seja,

$$\epsilon_{jk} \sim \text{NA}_c(0, \sigma_{\theta_k}, \gamma_{\theta_k}), \quad (6.8)$$

$$\delta_k \sim \text{NA}_c(\mathbf{0}, \Sigma_{\boldsymbol{\delta}}, \boldsymbol{\gamma}_{\delta}). \quad (6.9)$$

As priors dos parâmetros podem ser condicionalmente modeladas como

$$\log(a_{ik}) = \mathbf{H}_{ik}^{(a)} \kappa_k^{(a)} + \epsilon_{ik}^{(a)} \quad (6.10)$$

$$b_{ik} = \mathbf{H}_{ik}^{(b)} \kappa_k^{(b)} + \epsilon_{ik}^{(b)}. \quad (6.11)$$

Os modelos propostos por Fox and Glas (2001), Azevedo et al. (2011), Soares et al. (2009) e da-Silva and Gomes (2011), podem ser vistos como casos particulares desse modelo, veja discussão em Azevedo et al. (2012).

### 6.2.3 Outros métodos de estimação

Apesar dos algoritmos MCMC constituírem uma ferramenta poderosa na utilização da inferência bayesiana e na validação de modelos, seus longos períodos de computação fazem com que métodos de estimação alternativos sejam buscados. Uma alternativa interessante seria o uso do algoritmo EM condicional de dados aumentados proposto por Azevedo and Andrade (2007) ou o algoritmo EM estocástico, veja Celeux and Diebolt (1985). Estes métodos são uma variação do algoritmo EM e permitem ajustar modelos complexos com razoável facilidade e com baixo custo de processamento computacional. Uma outra possibilidade seria a estimação por máxima verossimilhança marginal. Salientamos que, a estrutura hierárquica desenvolvida neste trabalho, pode facilitar a implementação dos supracitados algoritmos.

### 6.2.4 Medidas de diagnóstico

As medidas de diagnóstico estudadas neste trabalho (valor-p bayesiano e escores preditos) são ferramentas úteis na avaliação do ajuste de modelos de resposta ao item. No entanto, outras medidas podem ser estudadas. Por exemplo, a razão de chances e a estatística de Mantel-Haenzel, poderiam ser usados para detectar afastamentos da suposição de unidimensionalidade, veja Sinharay et al (2006). Considerando a estrutura hierárquica desenvolvida na subseção 3.3.2,

poderíamos desenvolver algum tipo de resíduo bayesiano para os traços latentes, baseados na distribuição de  $H_{jk}$ . Estes poderiam ser usados para avaliar a pressuposição de normalidade assimétrica.



# Apêndice A

## Detalhes dos algoritmos MCMC

### A.1 Algoritmo ADMHWGS

A seguir temos a descrição do algoritmo ADMHWGS (utilizando densidade original):

- Passo 1: Simule as variáveis aumentadas  $Z_{ijk}^{(t)}$  de  $Z_{ijk}|(\cdot) \sim N(a_i^{(t-1)}\theta_{jk}^{(t-1)} - b_i^{(t-1)}, 1)\mathbb{1}_{(y_{ijk}, z_{ijk})}$ ,  $\forall k = 1, \dots, K, j = 1, \dots, n_k$  e  $i \in I_{jk}$  de forma mutuamente independente.
- Passo 2: Simule  $\theta_{jk}^{(t)}$  de  $\theta_{jk}|(\cdot)$  de forma mutuamente independente, considerando:
  - (a) simule  $\theta_{jk}^{(*)}|\theta_{jk}^{(t-1)} \sim N(\theta_{jk}^{(t-1)}, \sigma_0^2)$ ,
  - (b) Aceite  $\theta_{jk}^{(t)} = \theta_{jk}^{(*)}$  com probabilidade

$$\pi_{jk}(\theta_{jk}^{(t-1)}, \theta_{jk}^{(*)}) = \min \{R_{\theta_{jk}}, 1\} \text{ em que,}$$

$$\begin{aligned}
 R_{\theta_{jk}} &= \frac{\exp \left\{ -\frac{1}{2} \left\{ (\theta_{jk}^{(*)})^2 \sum_{i \in I_{jk}} (a_i^{(t-1)})^2 - 2\theta_{jk}^{(*)} \left[ \sum_{i \in I_{jk}} (z_{ijk}^{(t)} a_i^{(t-1)}) + \sum_{i \in I_{jk}} (a_i^{(t-1)} b_i^{(t-1)}) \right] \right\} \right\}}{\exp \left\{ -\frac{1}{2} \left\{ (\theta_{jk}^{(t-1)})^2 \sum_{i \in I_{jk}} (a_i^{(t-1)})^2 - 2\theta_{jk}^{(t-1)} \left[ \sum_{i \in I_{jk}} (z_{ijk}^{(t)} a_i^{(t-1)}) + \sum_{i \in I_{jk}} (a_i^{(t-1)} b_i^{(t-1)}) \right] \right\} \right\}} \\
 &\times \frac{\exp \left\{ -\frac{1}{2(\omega_{\theta_k}^{(t-1)})^2} \left[ (\theta_{jk}^{(*)})^2 - 2\theta_{jk}^{(*)} \xi_{\theta_k}^{(t-1)} \right] \right\} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(*)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t-1)}} \right) \right]}{\exp \left\{ -\frac{1}{2(\omega_{\theta_k}^{(t-1)})^2} \left[ (\theta_{jk}^{(t-1)})^2 - 2\theta_{jk}^{(t-1)} \xi_{\theta_k}^{(t-1)} \right] \right\} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t-1)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t-1)}} \right) \right]}. \tag{A.1}
 \end{aligned}$$

Caso contrário, faça  $\theta_{jk}^{(t)} = \theta_{jk}^{(t-1)}$ .

- Passo 3: Simule os parâmetros dos itens  $\zeta_i^{(t)}$  de  $\zeta_i | (.) \sim N(\widehat{\Psi}_{\zeta_i}^{(t)} \widehat{\zeta}_i^{(t)}, \widehat{\Psi}_{\zeta_i}^{(t)})$ , veja Azevedo (2008) e Lindley and Smith (1972), de forma mutuamente independente, em que

$$\begin{aligned}\widehat{\zeta}_i^{(t)} &= \left(\Theta_{i.}^{(t)}\right)^\top \mathbf{z}_{i.}^{(t)} + \Psi_{\zeta}^{-1} \mu_{\zeta}, \\ \widehat{\Psi}_{\zeta_i}^{(t)} &= \left[\left(\Theta_{i.}^{(t)}\right)^\top (\Theta_{i.}) + \Psi_{\zeta}^{-1}\right]^{-1}, \\ \Theta_{i.} &= \left[\boldsymbol{\theta}^{(t)} - \mathbf{1}_n\right] \bullet \mathbb{1}_i,\end{aligned}$$

onde  $\mathbb{1}_i$  é uma matriz  $(n \times 2)$  com elementos, em cada linha, iguais a 1 ou 0, de acordo com a resposta ou omissão de resposta do indivíduo  $j$  da população  $k$  ao item  $i$  e  $\bullet$  denota o produto de *Hadamard*, veja Horn and Johnson (1991).

- Passo 4: Escolha os passos: 4.1 a 4.3 ou os passos: 4.4 a 4.6.

#### (Estimação dos parâmetros populacionais utilizando a priori (3.15))

- Passo 4.1: Simule  $\xi_{\theta_k}^{(t)}$  de  $\xi_{\theta_k} | (.)$ , considerando:

- simule  $\xi_{\theta_k}^{(*)} | \xi_{\theta_k}^{(t-1)} \sim N(\xi_{\theta_k}^{(t-1)}, \sigma_0^2)$
- Aceite  $\xi_{\theta_k}^{(t)} = \xi_{\theta_k}^{(*)}$  com probabilidade

$$\pi_k \left[ \xi_{\theta_k}^{(t-1)}, \xi_{\theta_k}^{(*)} \right] = \min \left\{ R_{\xi_{\theta_k}}, 1 \right\} \text{ onde,}$$

$$R_{\xi_{\theta_k}} = \frac{\exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(*)}}{\omega_{\theta_k}^{(*)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(*)}}{\omega_{\theta_k}^{(*)}} \right) \right]}{\exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t-1)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t-1)}} \right) \right]}$$

Caso contrário, faça  $\xi_{\theta_k}^{(t)} = \xi_{\theta_k}^{(t-1)}$

- Passo 4.2: Simule  $\omega_{\theta_k}^{(t)}$  de  $\omega_{\theta_k} | (.)$  considerando:

- simule  $\omega_{\theta_k}^{(*)} | \omega_{\theta_k}^{(t-1)} \sim \text{Log-Normal}(\log(\omega_{\theta_k}^{(t-1)}), \sigma_0^2)$ ,
- Aceite  $\omega_{\theta_k}^{(t)} = \omega_{\theta_k}^{(*)}$  com probabilidade

$$\pi_k(\omega_{\theta_k}^{(t-1)}, \omega_{\theta_k}^{(*)}) = \min \left\{ R_{\omega_{\theta_k}}, 1 \right\} \text{ em que,}$$

$$R_{\omega_{\theta_k}} = \frac{(\omega^{(*)})^{-n_k} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(*)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(*)}} \right) \right]}{(\omega^{(t-1)})^{-n_k} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(t-1)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(t-1)}} \right) \right]}$$

- Passo 4.3: Simule  $\lambda_{\theta_k}^{(t)}$  de  $\lambda_{\theta_k} | (.)$  considerando:

- (a) simule  $\lambda_{\theta_k}^{(*)} | \lambda_{\theta_k}^{(t-1)} \sim N(\lambda_{\theta_k}^{(t-1)}, \sigma_0^2)$ ,
- (b) Aceite  $\lambda_{\theta_k}^{(t)} = \lambda_{\theta_k}^{(*)}$  com probabilidade

$$\pi_k(\lambda_{\theta_k}^{(t-1)}, \lambda_{\theta_k}^{(*)}) = \min \left\{ R_{\lambda_{\theta_k}}, 1 \right\} \text{ em que,}$$

$$R_{\lambda_{\theta_k}} = \frac{\left( 1 + \frac{(\lambda_{\theta_k}^{(*)})^2}{d\varphi^2} \right)^{-\frac{d+1}{2}} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(*)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(t)}} \right) \right]}{\left( 1 + \frac{(\lambda_{\theta_k}^{(t-1)})^2}{d\varphi^2} \right)^{-\frac{d+1}{2}} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(t)}} \right) \right]}.$$

Caso contrário, faça  $\lambda_{\theta_k}^{(t)} = \lambda_{\theta_k}^{(t-1)}$ .

(Estimação dos parâmetros populacionais utilizando as prioris (3.19), (3.20) e (3.21))

- Passo 4.4: Simule  $\xi_{\theta_k}^{(t)}$  de  $\xi_{\theta_k} | (.)$  considerando:

- (a) simule  $\xi_{\theta_k}^{(*)} | \xi_{\theta_k}^{(t-1)} \sim N(\xi_{\theta_k}^{(t-1)}, \sigma_0^2)$ ,
- (b) Aceite  $\xi_{\theta_k}^{(t)} = \xi_{\theta_k}^{(*)}$  com probabilidade

$$\pi_k(\xi_{\theta_k}^{(t-1)}, \xi_{\theta_k}^{(*)}) = \min \left\{ R_{\xi_{\theta_k}}, 1 \right\} \text{ em que,}$$

$$R_{\xi_{\theta_k}} = \frac{\exp \left\{ -\frac{1}{2\sigma_{\xi_{\theta}}^2} (\xi_{\theta_k}^{(*)} - \mu_{\xi_{\theta}})^2 \right\} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(*)}}{\omega_{\theta_k}^{(t)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(*)}}{\omega_{\theta_k}^{(t)}} \right) \right]}{\exp \left\{ -\frac{1}{2\sigma_{\xi_{\theta}}^2} (\xi_{\theta_k}^{(t-1)} - \mu_{\xi_{\theta}})^2 \right\} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t)}} \right) \right]}.$$

Caso contrário, faça  $\xi_{\theta_k}^{(t)} = \xi_{\theta_k}^{(t-1)}$ .

- Passo 4.5: Simule  $(\omega_{\theta_k}^2)^{(t)}$  de  $\omega_{\theta_k}^2 | (\cdot)$  considerando:

- simule  $(\omega_{\theta_k}^2)^{(*)} | (\omega_{\theta_k}^2)^{(t-1)} \sim \text{Log-Normal}(\log((\omega_{\theta_k}^2)^{(t-1)}), \sigma_0^2)$ ,
- Aceite  $(\omega_{\theta_k}^2)^{(t)} = (\omega_{\theta_k}^2)^{(*)}$  com probabilidade

$$\pi_k((\omega_{\theta_k}^2)^{(t-1)}, (\omega_{\theta_k}^2)^{(*)}) = \min \left\{ R_{\omega_{\theta_k}^2}, 1 \right\} \text{ em que,}$$

$$R_{\omega_{\theta_k}^2} = \frac{((\omega_{\theta_k}^2)^{(*)})^{-(\alpha_{\omega_{\theta_k}} + \frac{n_k}{2})} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(*)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(*)}} \right) \right]}{((\omega_{\theta_k}^2)^{(t-1)})^{-(\alpha_{\omega_{\theta_k}} + \frac{n_k}{2})} \exp \left\{ -\frac{1}{2} \sum_{j=1}^{n_k} \left( \frac{\theta_{jk}^{(t-1)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t-1)}} \right)^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t-1)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t-1)}} \right) \right]}.$$

Caso contrário, faça  $(\omega_{\theta_k}^2)^{(t)} = (\omega_{\theta_k}^2)^{(t-1)}$ .

- Passo 4.6: Simule  $\lambda_{\theta_k}^{(t)}$  de  $\lambda_{\theta_k} | (\cdot)$  considerando:

- simule  $\lambda_{\theta_k}^{(*)} | \lambda_{\theta_k}^{(t-1)} \sim N(\lambda_{\theta_k}^{(t-1)}, \sigma_0^2)$ ,
- Aceite  $\lambda_{\theta_k}^{(t)} = \lambda_{\theta_k}^{(*)}$  com probabilidade

$$\pi_k(\lambda_{\theta_k}^{(t-1)}, \lambda_{\theta_k}^{(*)}) = \min \left\{ R_{\lambda_{\theta_k}}, 1 \right\} \text{ em que,}$$

$$R_{\lambda_{\theta_k}} = \frac{\exp \left\{ -\frac{1}{2\sigma_{\lambda_{\theta_k}}^2} (\lambda_{\theta_k}^{(*)} - \mu_{\lambda_{\theta_k}})^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(*)} \left( \frac{\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}}{\omega_{\theta_k}^{(t)}} \right) \right]}{\exp \left\{ -\frac{1}{2\sigma_{\lambda_{\theta_k}}^2} (\lambda_{\theta_k}^{(t-1)} - \mu_{\lambda_{\theta_k}})^2 \right\} \prod_{j=1}^{n_k} \Phi \left[ \lambda_{\theta_k}^{(t-1)} \left( \frac{\theta_{jk}^{(t-1)} - \xi_{\theta_k}^{(t-1)}}{\omega_{\theta_k}^{(t-1)}} \right) \right]}.$$

Caso contrário, faça  $\lambda_{\theta_k}^{(t)} = \lambda_{\theta_k}^{(t-1)}$ .

**Observação:** As estimativas dos parâmetros populacionais  $(\mu_{\theta_k}, \sigma_{\theta_k}, \gamma_{\theta_k})$  são obtidas a cada iteração a partir das estimativas dos parâmetros  $(\xi_{\theta_k}, \omega_{\theta_k}, \lambda_{\theta_k})$  da seguinte forma:

$$\begin{aligned}
\gamma_{\theta_k}^{(t)} &= r \left( \delta_{\theta_k}^{(t)} \right)^3 \left( \frac{4}{\pi} - 1 \right) \left( 1 - r^2 \left( \delta_{\theta_k}^{(t)} \right)^2 \right)^{-3/2}, \\
\sigma_{\theta_k}^{(t)} &= \frac{\omega_{\theta_k}^{(t)}}{\sqrt{1 + \left( \gamma_{\theta_k}^{(t)} \right)^{2/3} s^2}}, \\
\mu_{\theta_k}^{(t)} &= \xi_{\theta_k}^{(t)} + \sigma_{\theta_k}^{(t)} \left( \gamma_{\theta_k}^{(t)} \right)^{1/3} s,
\end{aligned}$$

para todo  $t = 1, 2, \dots, B, \dots, M$ , onde  $B$  é o *burn-in* e  $M$  é o tamanho de amostra gerado e,

$$\begin{aligned}
\delta_{\theta_k}^{(t)} &= \frac{\lambda_{\theta_k}^{(t)}}{\sqrt{1 + \left( \lambda_{\theta_k}^{(t)} \right)^2}}, \\
s &= \left( \frac{2}{4 - \pi} \right)^{1/3}.
\end{aligned}$$

## A.2 Algoritmo ADGS

A seguir temos a descrição do algoritmo amostrador de *Gibbs* completo (utilizando representação estocástica).

- Passo 1: Simule as variáveis aumentadas  $Z_{ijk}^{(t)}$  de  $Z_{ijk}|(\cdot) \sim N(a_i^{(t-1)}\theta_{jk}^{(t-1)} - b_i^{(t-1)}, 1)\mathbb{1}_{(y_{ijk}, z_{ijk})}$ , de forma mutuamente independente  $\forall k = 1, \dots, K, j = 1, \dots, n_k$  e  $i \in I_{jk}$ .
- Passo 2: Simule  $\theta_{jk}^{(t)}$  de  $\theta_{jk}|(\cdot) \sim N(\hat{\tau}_{\theta_{jk}}^{(t-1)}\hat{\theta}_{jk}^{(t-1)}, \hat{\tau}_{\theta_{jk}}^{(t-1)}) \forall k = 1, \dots, K, j = 1, \dots, n_k$  de forma mutuamente independente, em que:

$$\begin{aligned}
\hat{\theta}_{jk}^{(t-1)} &= \sum_{i \in I_{jk}} z_{ijk}^{(t)} a_i^{(t-1)} + \sum_{i \in I_{jk}} a_i^{(t-1)} b_i^{(t-1)} + \frac{\xi_{\theta_k}^{(t-1)} + \tau_{\theta_k}^{(t-1)} h_{jk}^{(t-1)}}{(\varsigma_{\theta_k}^{(t-1)})^2}, \\
\hat{\tau}_{\theta_{jk}}^{(t-1)} &= \left( \frac{1}{(\varsigma_{\theta_k}^{(t-1)})^2} + \sum_{i \in I_{jk}} (a_i^{(t-1)})^2 \right)^{-1}.
\end{aligned}$$

- Passo 3: Simule as variáveis latentes  $H_{jk}^{(t)}$  de

$$H_{jk}|(.) \sim \text{HN} \left( \frac{\tau_{\theta_k}^{(t-1)}(\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t-1)})}{(\tau_{\theta_k}^{(t-1)})^2 + (\xi_{\theta_k}^{(t-1)})^2}, \frac{(\xi_{\theta_k}^{(t-1)})^2}{(\xi_{\theta_k}^{(t-1)})^2 + (\tau_{\theta_k}^{(t-1)})^2} \right),$$

de forma mutuamente independente  $\forall k = 1, \dots, K$  e  $j = 1, \dots, n_k$ .

- Passo 4: Simule os parâmetros dos itens  $\zeta_i^{(t)}$  de  $\zeta_i|(.) \sim N(\widehat{\Psi}_{\zeta_i}^{(t)} \widehat{\zeta}_i^{(t)}, \widehat{\Psi}_{\zeta_i}^{(t)})$ , veja Azevedo (2008) e Lindley and Smith (1972), de forma mutuamente independente, em que

$$\begin{aligned} \widehat{\zeta}_i^{(t)} &= \left( \Theta_{i..}^{(t)} \right)^\top \mathbf{z}_{i..}^{(t)} + \Psi_{\zeta}^{-1} \mu_{\zeta}, \\ \widehat{\Psi}_{\zeta_i}^{(t)} &= \left[ \left( \Theta_{i..}^{(t)} \right)^\top \Theta_{i..}^{(t)} + \Psi_{\zeta}^{-1} \right]^{-1}, \\ \Theta_{i..} &= \left[ \theta^{(t)} - \mathbf{1}_n \right] \bullet \mathbb{1}_i, \end{aligned}$$

onde  $\mathbb{1}_i$  é uma matriz  $(n \times 2)$  com elementos, em cada linha, iguais a 1 ou 0, de acordo com a resposta ou omissão de resposta do indivíduo  $j$  da população  $k$  ao item  $i$  e  $\bullet$  denota o produto de *Hadamard*, veja Horn and Johnson (1991).

- Passo 5: Escolha os passos 5.1 a 5.4 ou os passos 5.5 e 5.6.

**(Estimação dos parâmetros populacionais utilizando a priori (3.34)).**

- Passo 5.1: Simule as variáveis latentes  $T_k^{(t)}$  de

$$T_k|(.) \sim \text{Gama} \left( \frac{d+1}{2}, \frac{(\tau_{\theta_k}^{(t-1)})^2}{(\xi_{\theta_k}^{(t-1)})^2 \varphi^2} + d \right),$$

de forma mutuamente independente  $\forall k = 1, \dots, K$ .

- Passo 5.2: Simule  $\xi_{\theta_k}^{(t)}$  de

$$\xi_{\theta_k}|(.) \sim N \left( \frac{1}{n_k} \sum_{j=1}^{n_k} (\theta_{jk}^{(t)} - \tau_{\theta_k}^{(t-1)} h_{jk}^{(t)}), \frac{(\xi_{\theta_k}^{(t-1)})^2}{n_k} \right)$$

de forma mutuamente independente  $\forall k = 1, \dots, K$ .

- Passo 5.3: Simule  $\tau_{\theta_k}^{(t)}$  de  $\tau_{\theta_k}|(.) \sim N(\hat{\mu}_{\tau_{\theta_k}}^{(t)}, (\hat{\sigma}_{\tau_{\theta_k}}^{(t)})^2)$  de forma mutuamente independente, em que:

$$\begin{aligned}\hat{\mu}_{\tau_{\theta_k}}^{(t)} &= \frac{\sum_{j=1}^{n_k} (\theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)}) h_{jk}^{(t)}}{\sum_{j=1}^{n_k} (h_{jk}^{(t)})^2 + \frac{t_k^{(t)}}{(\varsigma_{\theta_k}^{(t-1)})^2 \varphi^2}}, \\ (\hat{\sigma}_{\tau_{\theta_k}}^{(t)})^2 &= \frac{(\varsigma_{\theta_k}^{(t-1)})^2}{\sum_{j=1}^{n_k} (h_{jk}^{(t)})^2 + \frac{t_k^{(t)}}{(\varsigma_{\theta_k}^{(t-1)})^2 \varphi^2}},\end{aligned}$$

$\forall k = 1, \dots, K$ .

- Passo 5.4: Simule  $(\varsigma_{\theta_k}^{(t)})^2$  de  $\varsigma_{\theta_k}^2 \sim \text{Gama-Inv.}(\hat{\alpha}_{\varsigma_{\theta_k}}^{(t)}, \hat{\beta}_{\varsigma_{\theta_k}}^{(t)})$  de forma mutuamente independente, em que:

$$\begin{aligned}\hat{\alpha}_{\varsigma_{\theta_k}}^{(t)} &= \frac{n_k + 4}{2}, \\ \hat{\beta}_{\varsigma_{\theta_k}}^{(t)} &= \frac{1}{2} \left[ \sum_{j=1}^{n_k} \left( \theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)} - \tau_{\theta_k}^{(t)} h_{jk}^{(t)} \right)^2 + \frac{(\tau_{\theta_k}^{(t)})^2 t_k^{(t)}}{\varphi^2} \right],\end{aligned}$$

$\forall k = 1, \dots, K$ .

**(Estimação dos parâmetros populacionais utilizando as priors (3.35) e (3.36)).**

- Passo 5.5: Simule  $\beta_{\theta_k}^{(t)} = (\xi_{\theta_k}^{(t)}, \tau_{\theta_k}^{(t)})^\top$  de  $\beta_{\theta_k}|(.) \sim N[\hat{\boldsymbol{\tau}}_{\theta_k}^{(t)} \hat{\boldsymbol{\xi}}_{\theta_k}^{(t)}, \hat{\boldsymbol{\tau}}_{\theta_k}^{(t)}]$  de forma mutuamente independente, em que:

$$\begin{aligned}
\hat{\boldsymbol{\tau}}_{\theta_k}^{(t)} &= \left[ \left( \mathbf{H}_{.k}^{(t)} \right)^\top \left( \mathbf{D}_{n_k}^{(t-1)} \right)^{-1} \left( \mathbf{H}_{.k}^{(t)} \right) + \Sigma_{\beta_\theta}^{-1} \right]^{-1}, \\
\hat{\boldsymbol{\xi}}_{\theta_k}^{(t)} &= \left( \mathbf{H}_{.k}^{(t)} \right) \left( \mathbf{D}_{n_k}^{(t-1)} \right)^{-1} \boldsymbol{\theta}_{.k}^{(t)} + \Sigma_{\beta_\theta}^{-1} \boldsymbol{\mu}_{\beta_\theta}, \\
\mathbf{H}_{.k}^{(t)} &= \begin{bmatrix} \mathbf{1}_{n_k} & \mathbf{h}_{.k}^{(t)} \end{bmatrix} \text{ e} \\
\mathbf{D} &= \begin{pmatrix} \left( \varsigma_{\theta_k}^{(t-1)} \right)^2 & & \\ & \ddots & \\ & & \left( \varsigma_{\theta_k}^{(t-1)} \right)^2 \end{pmatrix},
\end{aligned}$$

é uma matriz diagonal de ordem  $n_k$ ,  $\forall k = 1, \dots, K$ .

- Passo 5.6: Simule  $\left( \varsigma_{\theta_k}^{(t)} \right)^2$  de

$$\varsigma_{\theta_k}^2 | (\cdot) \sim \text{Gama-Inv.} \left( \frac{n_k}{2} + \alpha_{\varsigma_\theta}, \frac{1}{2} \sum_{j=1}^{n_k} \left( \theta_{jk}^{(t)} - \xi_{\theta_k}^{(t)} - \tau_{\theta_k}^{(t)} h_{jk}^{(t)} \right)^2 + \beta_{\varsigma_\theta} \right),$$

de forma mutuamente independente  $\forall k = 1, \dots, K$ .

**Observação:** As estimativas dos parâmetros populacionais  $(\mu_{\theta_k}, \sigma_{\theta_k}, \gamma_{\theta_k})$  são obtidas a cada iteração a partir das estimativas dos parâmetros  $(\xi_{\theta_k}, \tau_{\theta_k}, \varsigma_{\theta_k})$  da seguinte forma:

$$\begin{aligned}
\gamma_{\theta_k}^{(t)} &= r \left( \delta_{\theta_k}^{(t)} \right)^3 \left( \frac{4}{\pi} - 1 \right) \left( 1 - r^2 \left( \delta_{\theta_k}^{(t)} \right)^2 \right)^{-3/2}, \\
\sigma_{\theta_k}^{(t)} &= \frac{\omega_{\theta_k}^{(t)}}{\sqrt{1 + \left( \gamma_{\theta_k}^{(t)} \right)^{2/3} s^2}}, \\
\mu_{\theta_k}^{(t)} &= \xi_{\theta_k}^{(t)} + \sigma_{\theta_k}^{(t)} \left( \gamma_{\theta_k}^{(t)} \right)^{1/3} s,
\end{aligned}$$

para todo  $t = 1, 2, \dots, B, \dots, M$ , onde  $B$  é o *burn-in* e  $M$  é o tamanho de amostra gerado e,

$$\delta_{\theta_k}^{(t)} = \frac{\tau_{\theta_k}^{(t)}}{\sqrt{\left(\tau_{\theta_k}^{(t)}\right)^2 + \left(\varsigma_{\theta_k}^{(t)}\right)^2}},$$

$$\omega_{\theta_k}^{(t)} = \frac{\tau_{\theta_k}^{(t)}}{\delta_{\theta_k}^{(t)}}.$$



# Apêndice B

## Gráficos: Estudo de simulação

### B.1 Gráficos das estatísticas REQM

Os gráficos referentes aos parâmetros populacionais apresentam a mediana (ponto sólido) e amplitude total (representada pela linha vertical tracejada) que é igual a diferença entre o máximo e o mínimo, de cada conjunto de valores. Os gráficos dos demais parâmetros são boxplots.

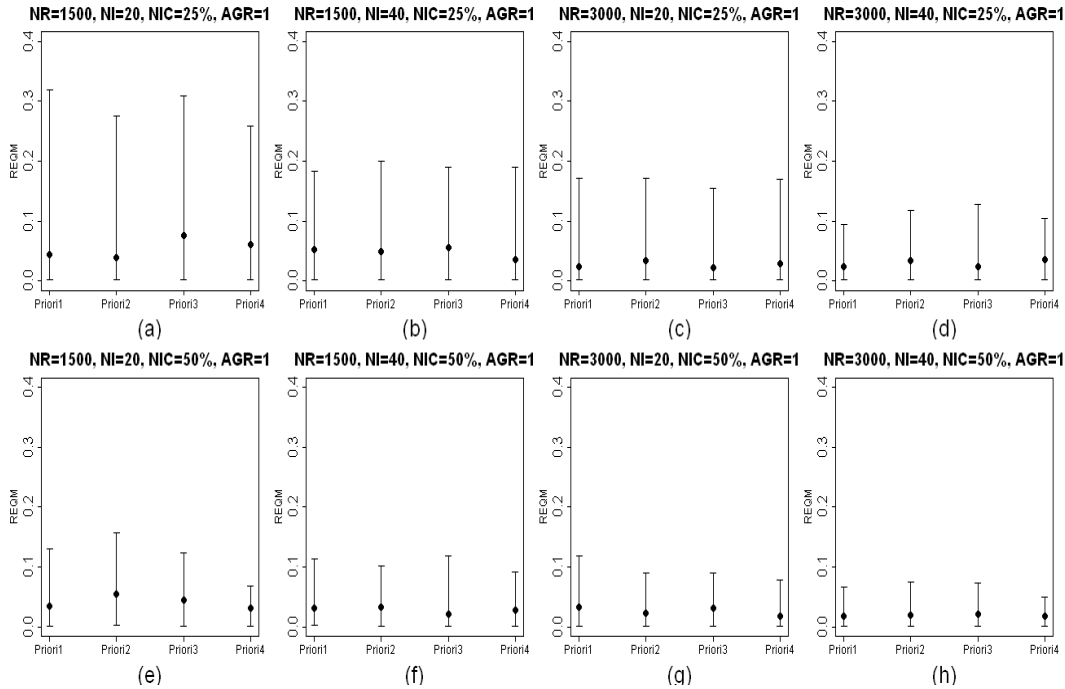


Figura B.1: Valores de REQM das estimativas das médias populacionais considerando AGR=1

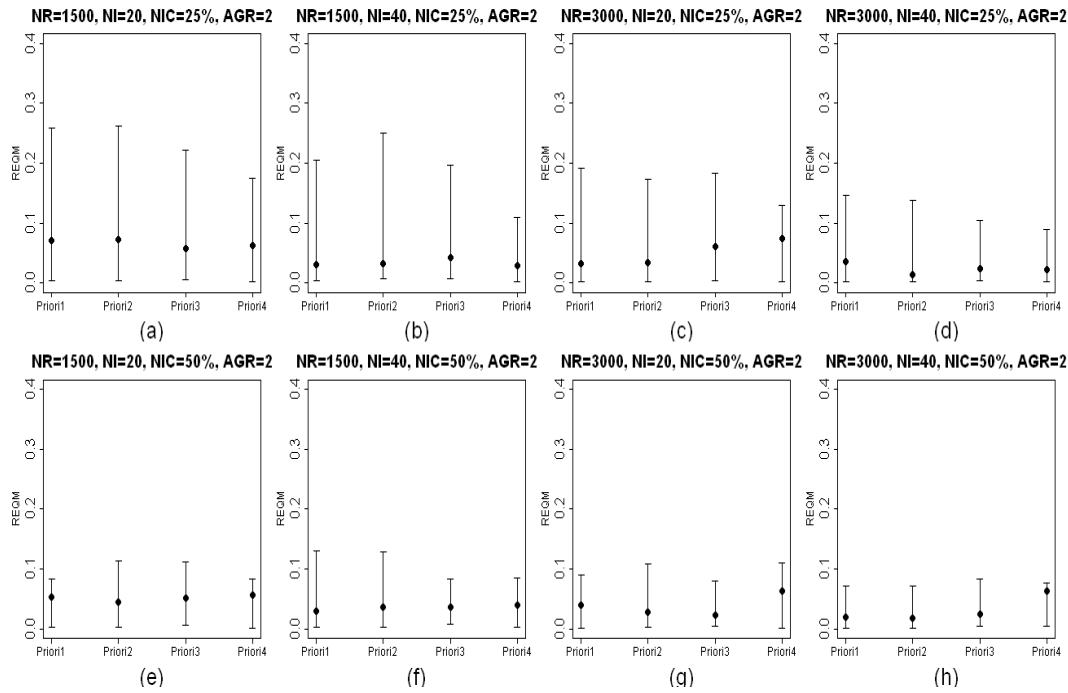


Figura B.2: Valores de REQm das estimativas das médias populacionais considerando AGR=2

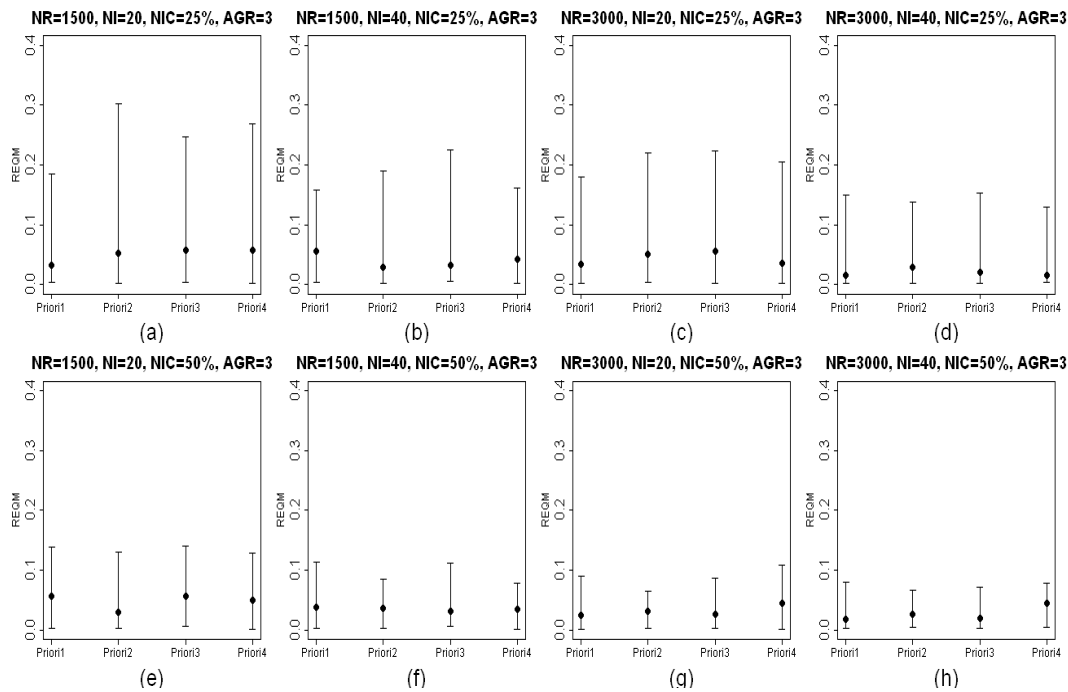


Figura B.3: Valores de REQm das estimativas das médias populacionais considerando AGR=3

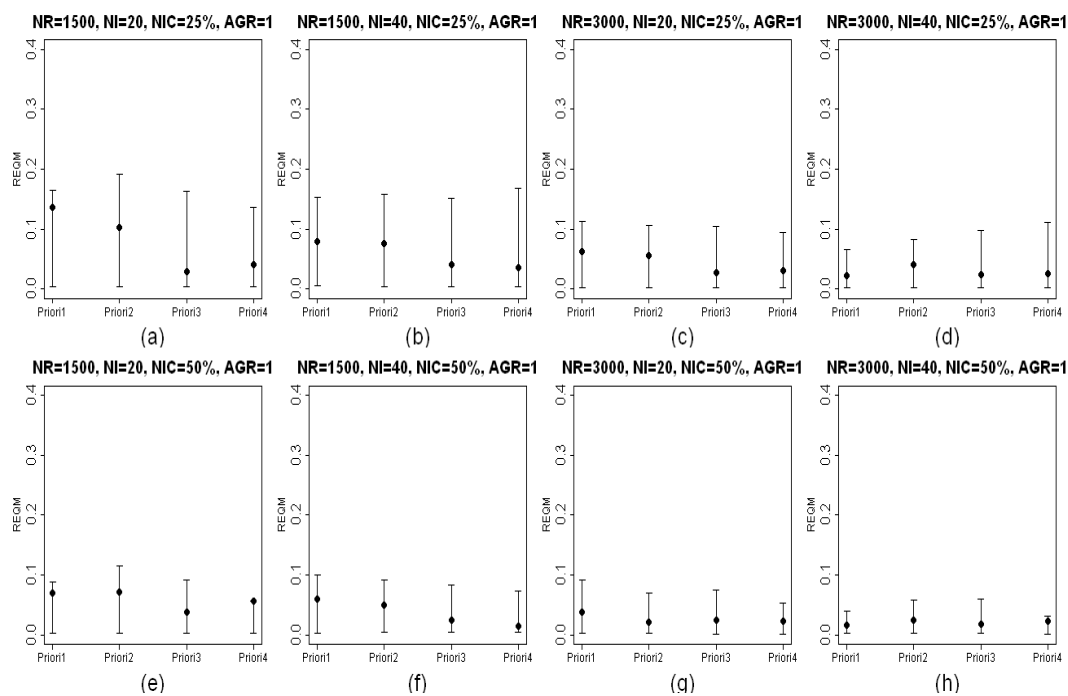


Figura B.4: Valores de REQM das estimativas dos desvios-padrão populacionais considerando AGR=1

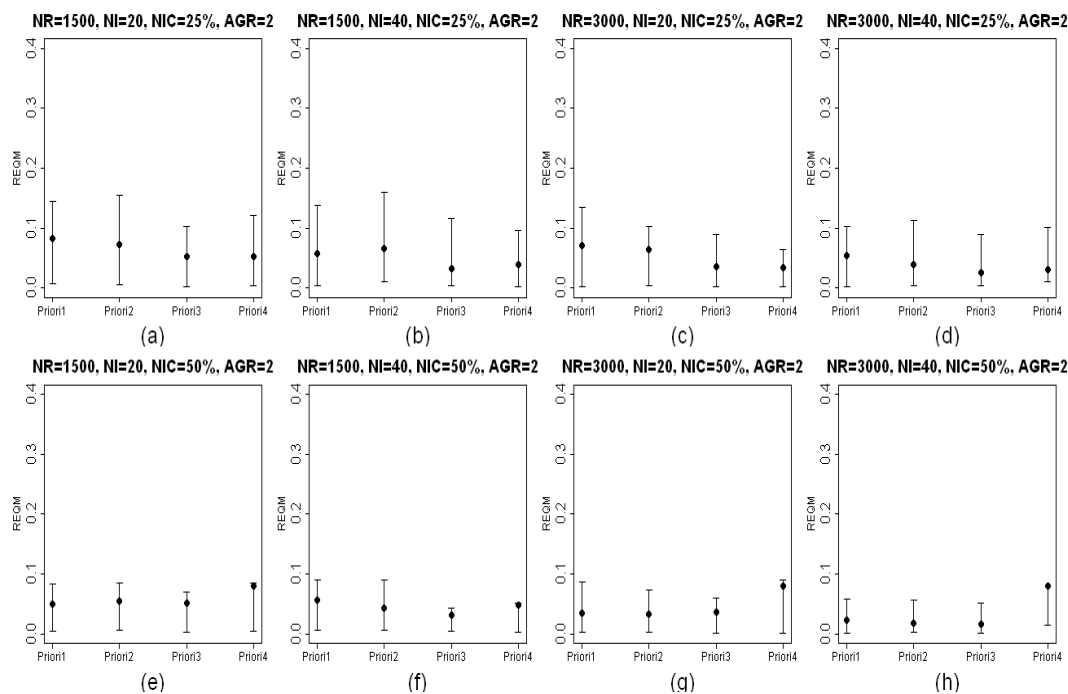


Figura B.5: Valores de REQM das estimativas dos desvios-padrão populacionais considerando AGR=2

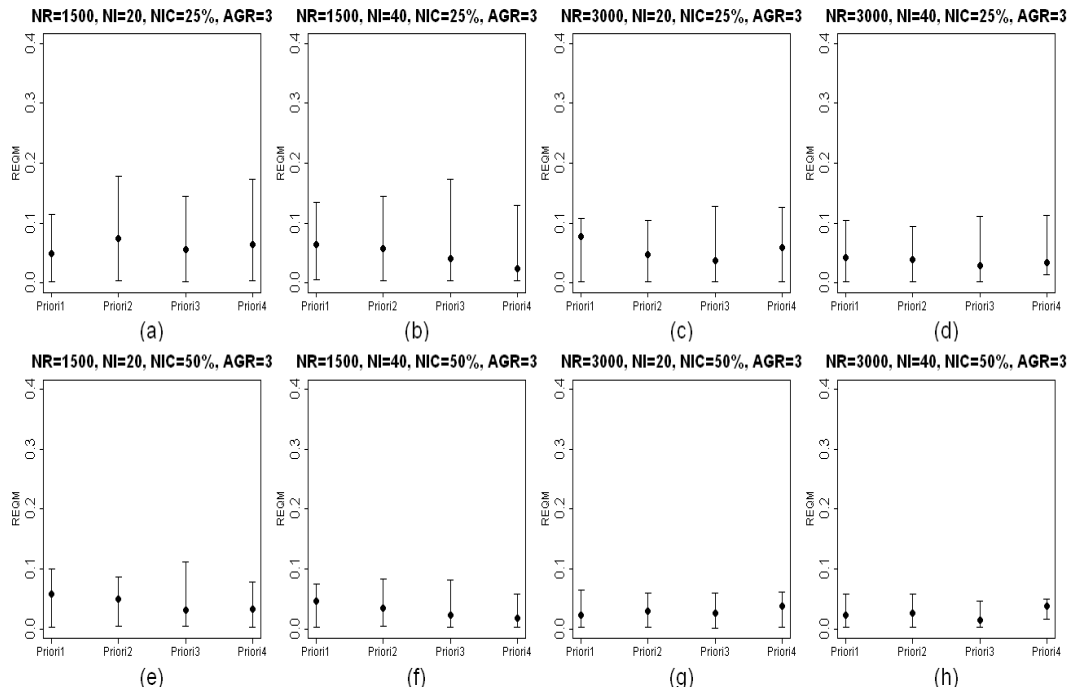


Figura B.6: Valores de REQM das estimativas dos desvios-padrão populacionais considerando  $AGR=3$

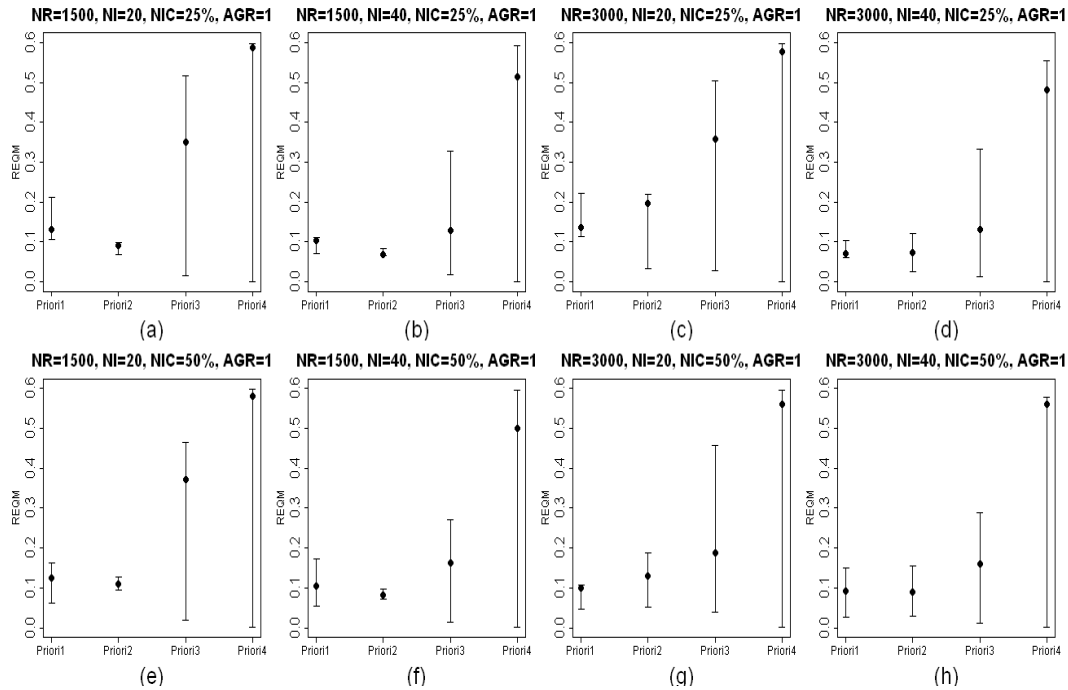


Figura B.7: Valores de REQM das estimativas dos coeficientes de assimetria populacionais considerando  $AGR=1$

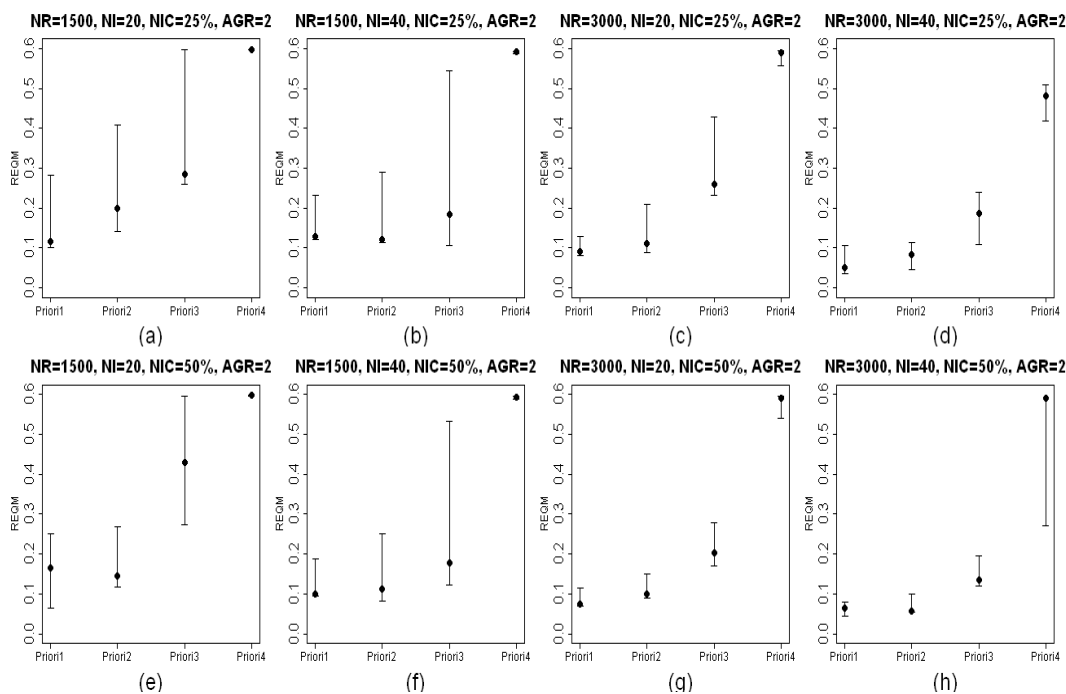


Figura B.8: Valores de REQM das estimativas dos coeficientes de assimetria populacionais considerando AGR=2

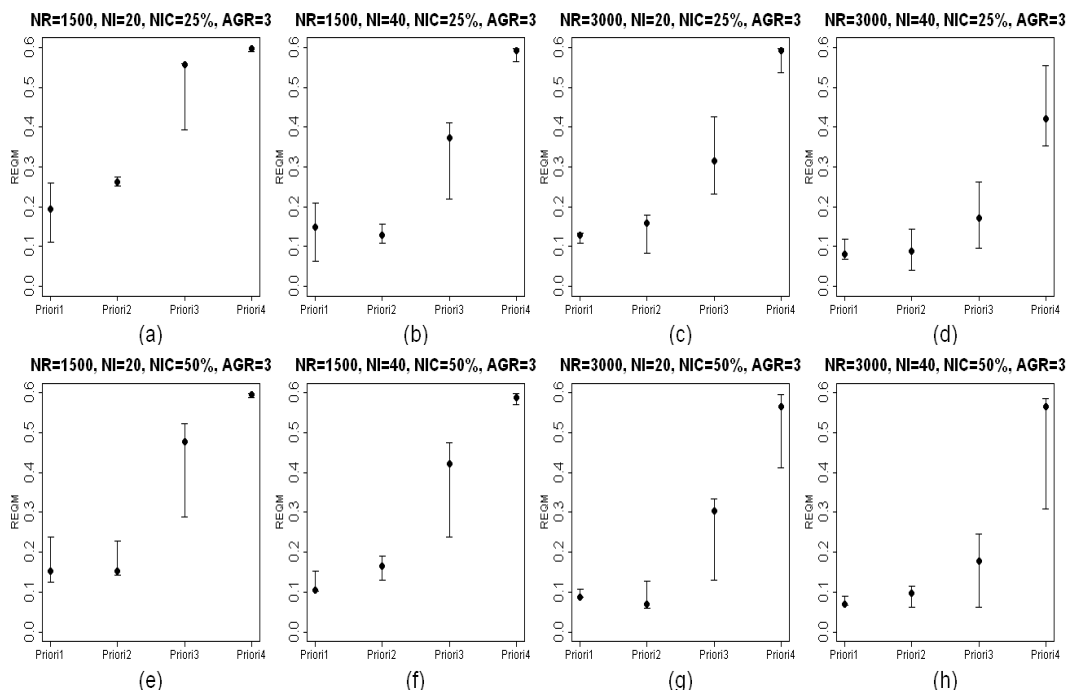


Figura B.9: Valores de REQM das estimativas dos coeficientes de assimetria populacionais considerando AGR=3

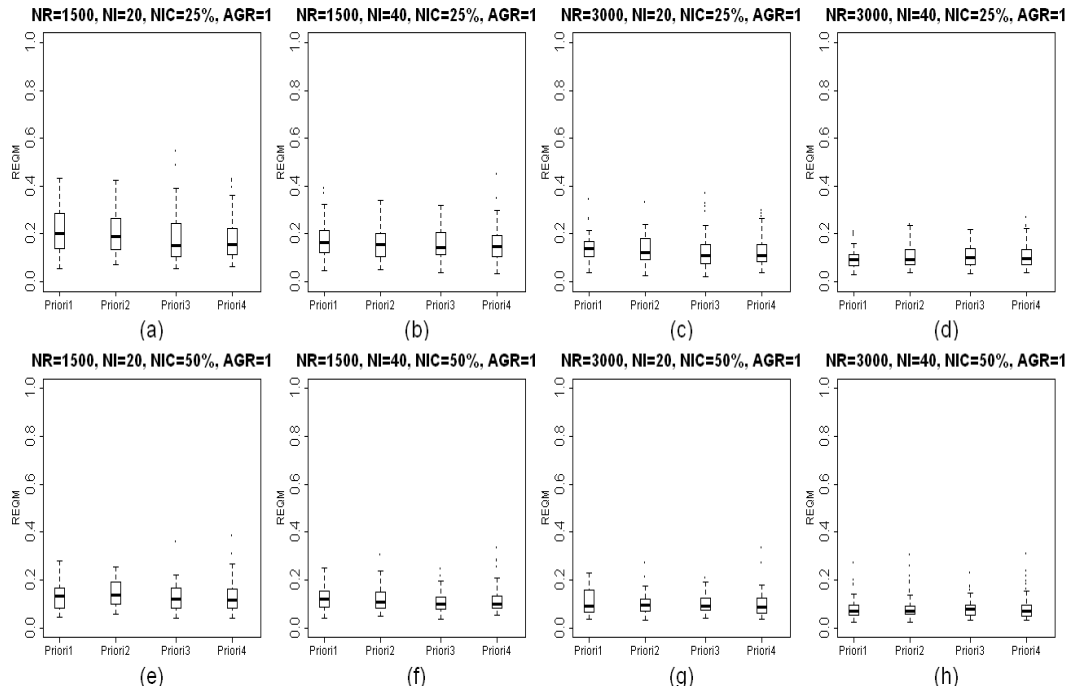


Figura B.10: Valores de REQM das estimativas dos parâmetros de discriminação considerando  $AGR=1$

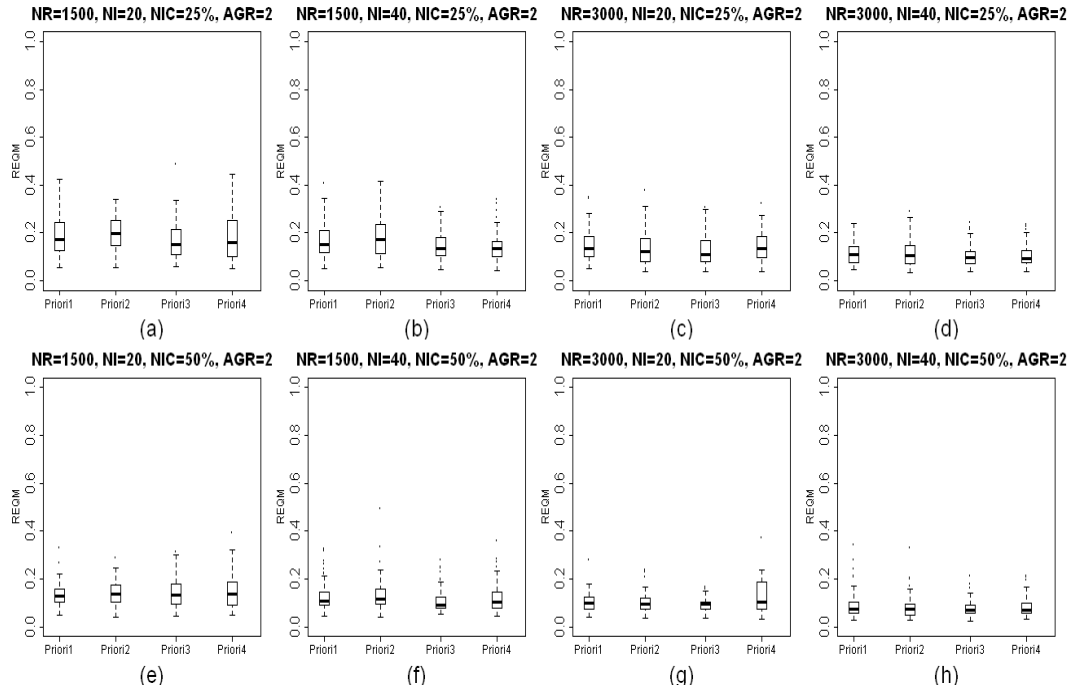


Figura B.11: Valores de REQM das estimativas dos parâmetros de discriminação considerando  $AGR=2$

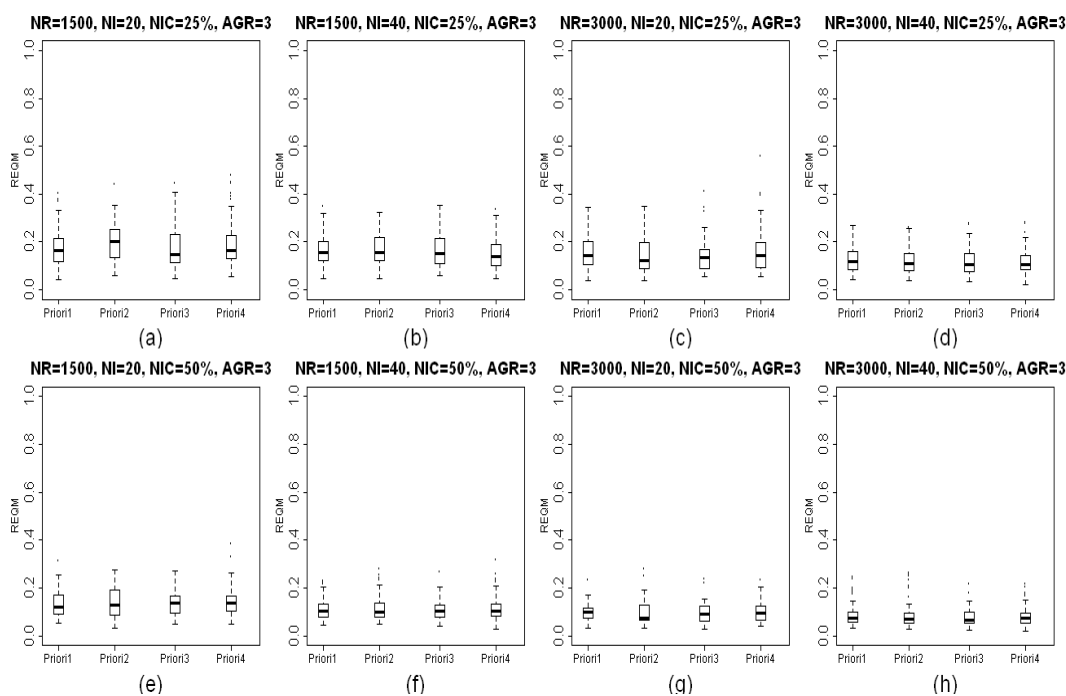


Figura B.12: Valores de REQM das estimativas dos parâmetros de discriminação considerando  $AGR=3$

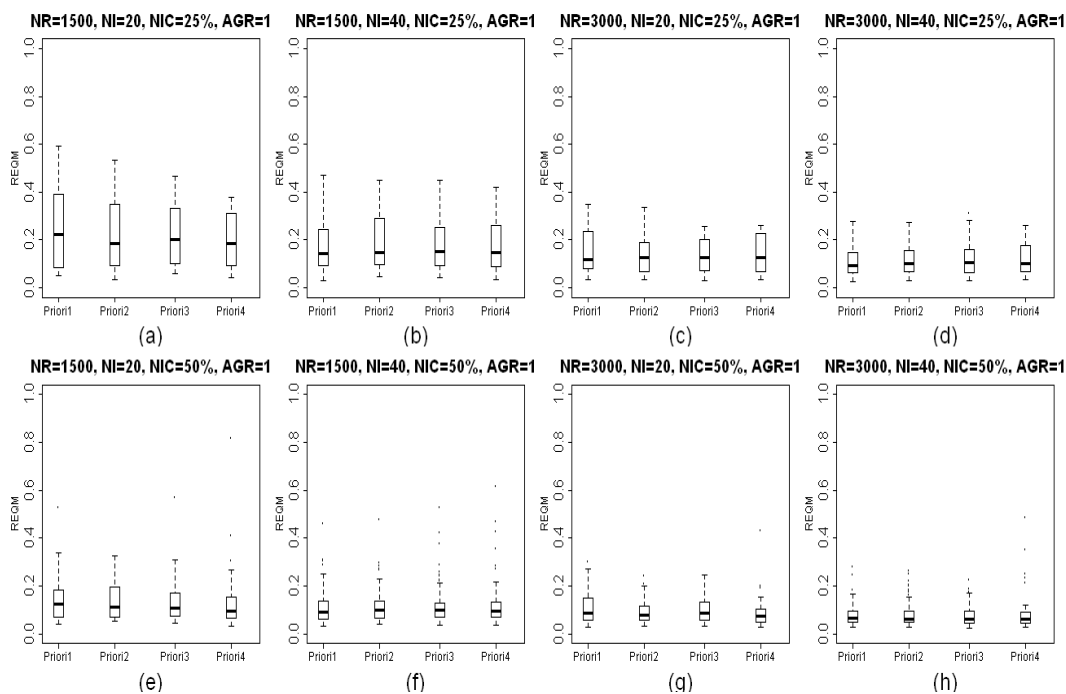


Figura B.13: Valores de REQM das estimativas dos parâmetros de dificuldade considerando  $AGR=1$

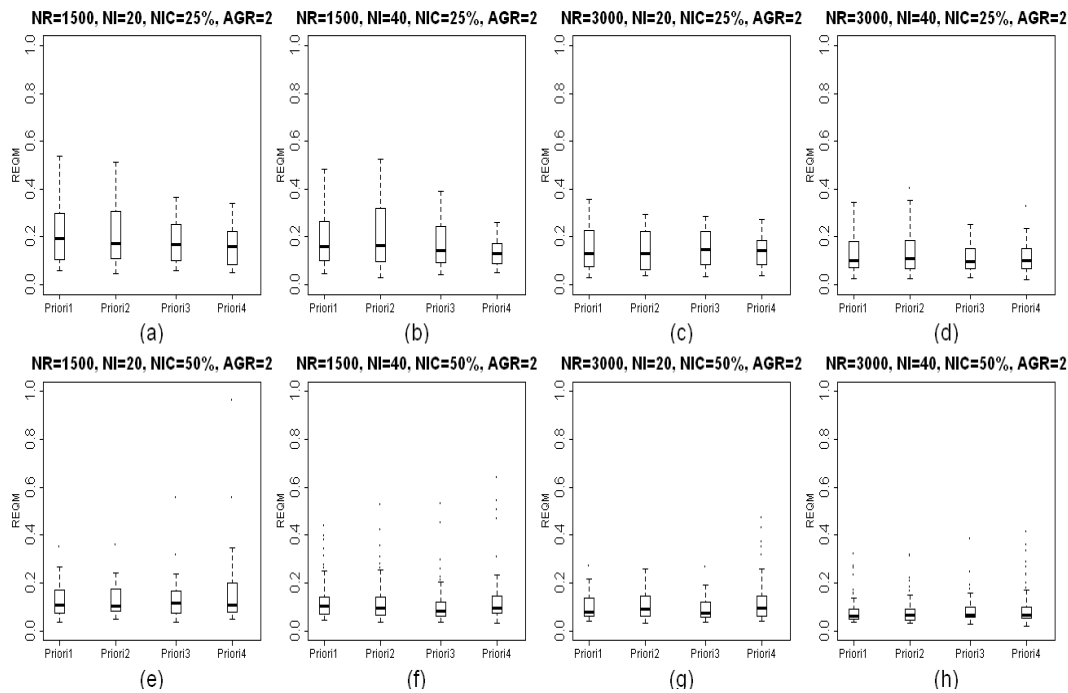


Figura B.14: Valores de REQM das estimativas dos parâmetros de dificuldade considerando  $AGR=2$

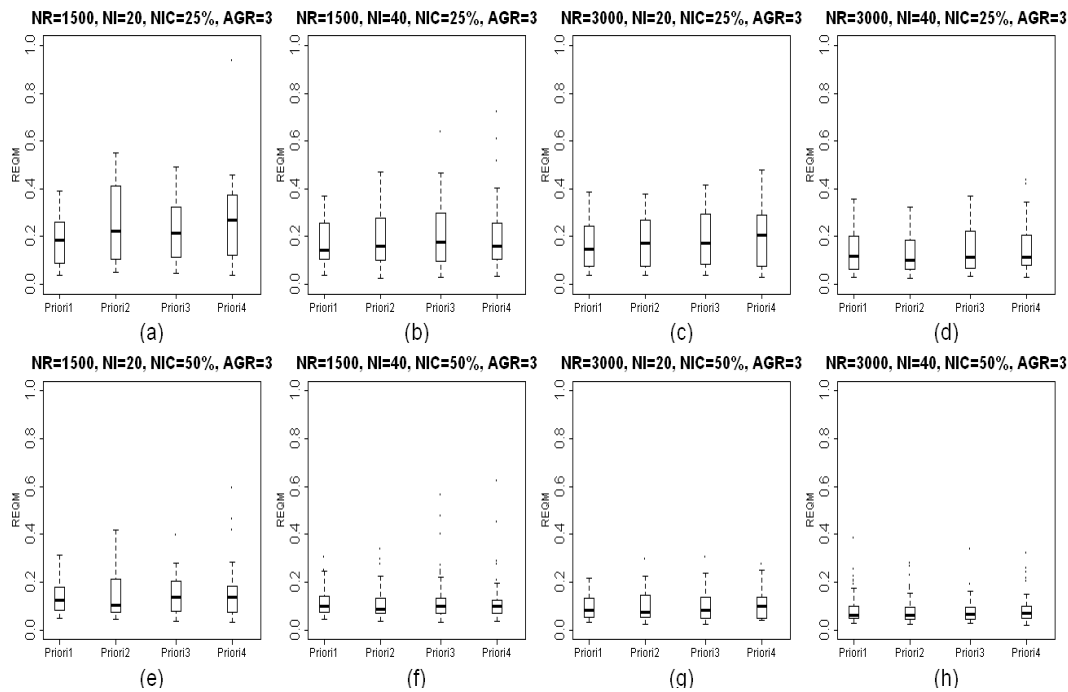


Figura B.15: Valores de REQM das estimativas dos parâmetros de dificuldade considerando  $AGR=3$

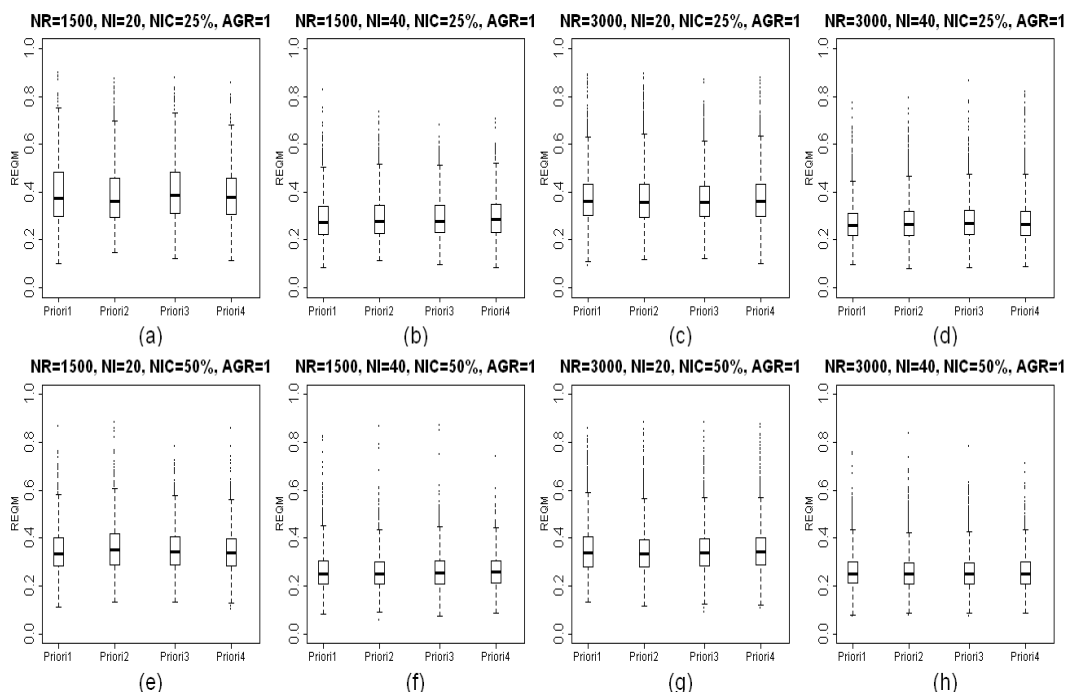


Figura B.16: Valores de REQM das estimativas dos traços latentes considerando AGR=1

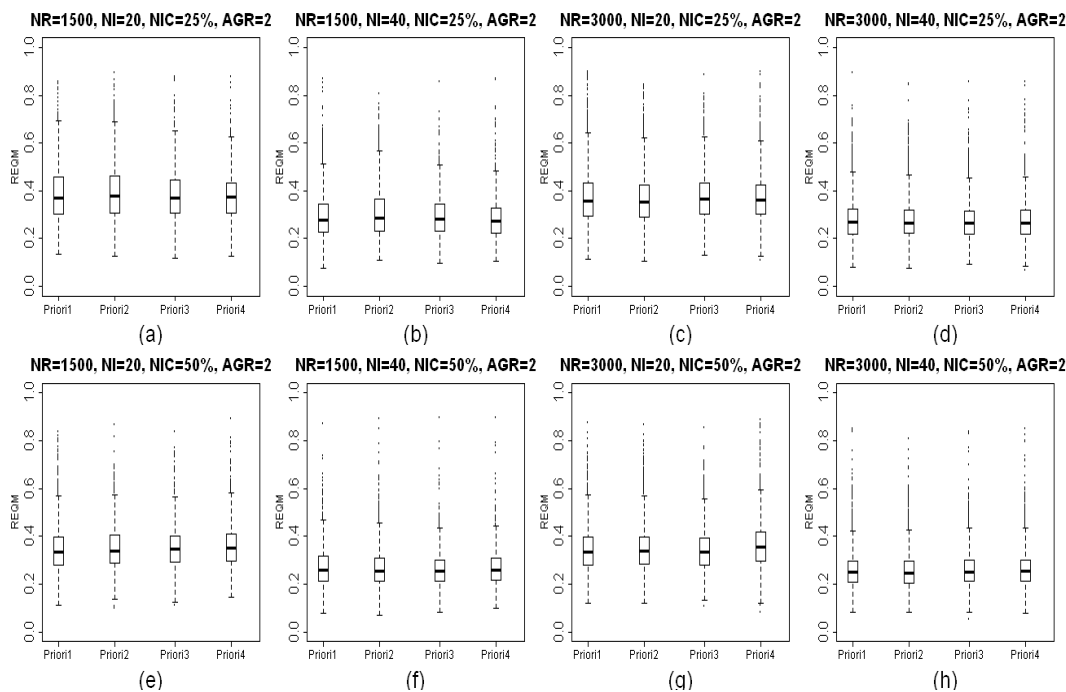


Figura B.17: Valores de REQM das estimativas dos traços latentes considerando AGR=2

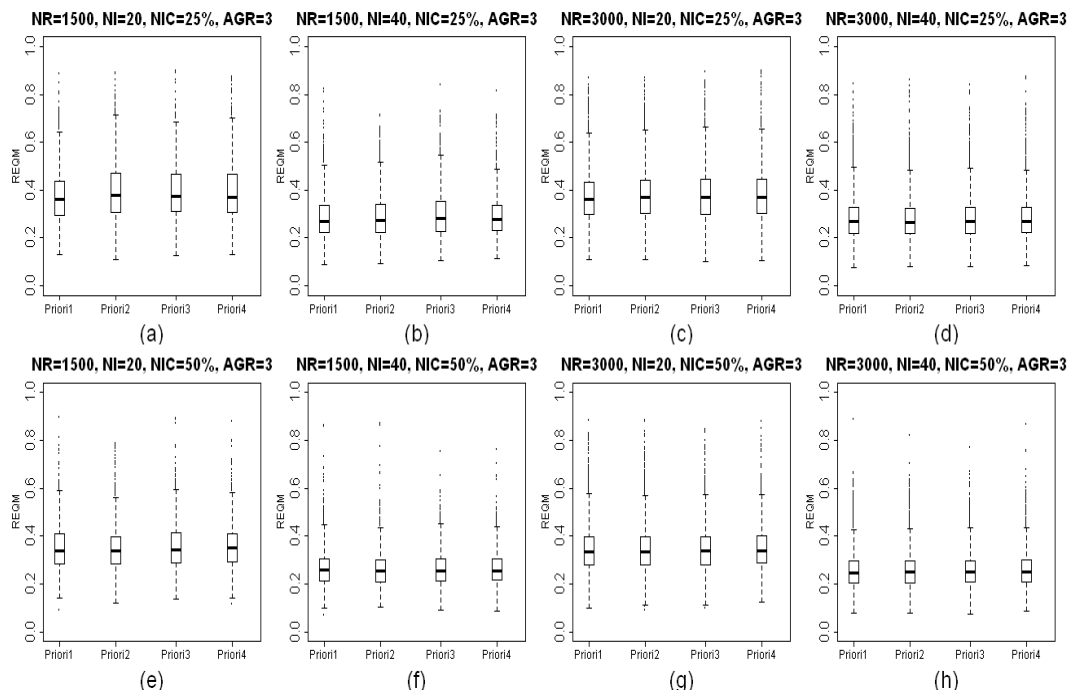


Figura B.18: Valores de REQM das estimativas dos traços latentes considerando  $\text{AGR}=3$

## B.2 Gráficos das estatísticas VRA

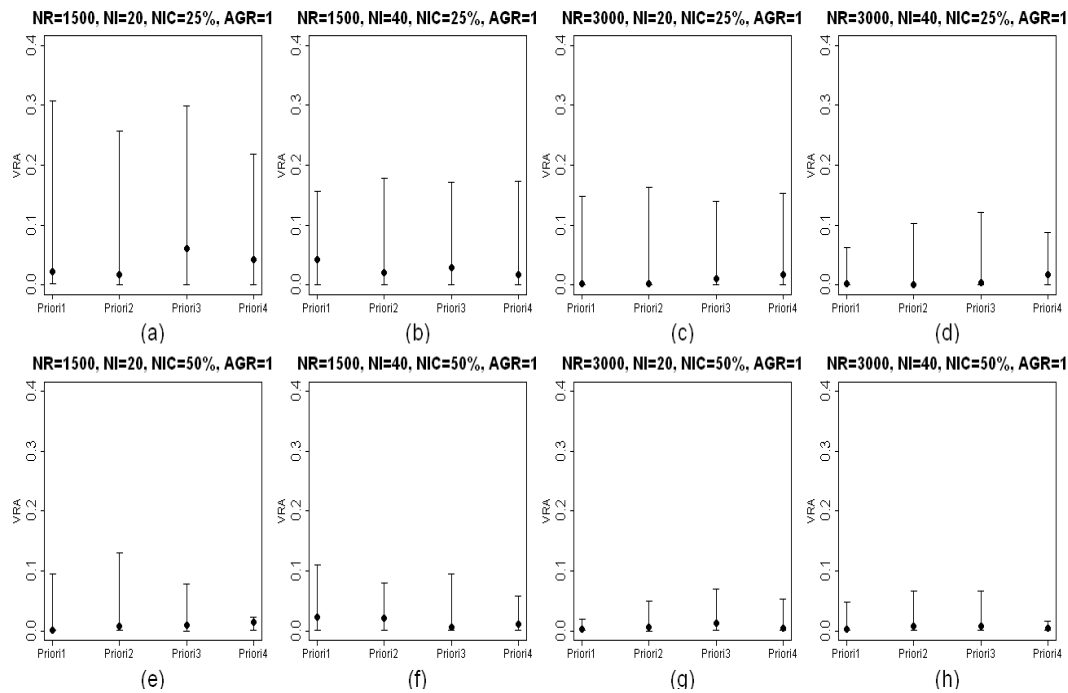


Figura B.19: Valores de VRA das estimativas das médias populacionais considerando  $AGR=1$

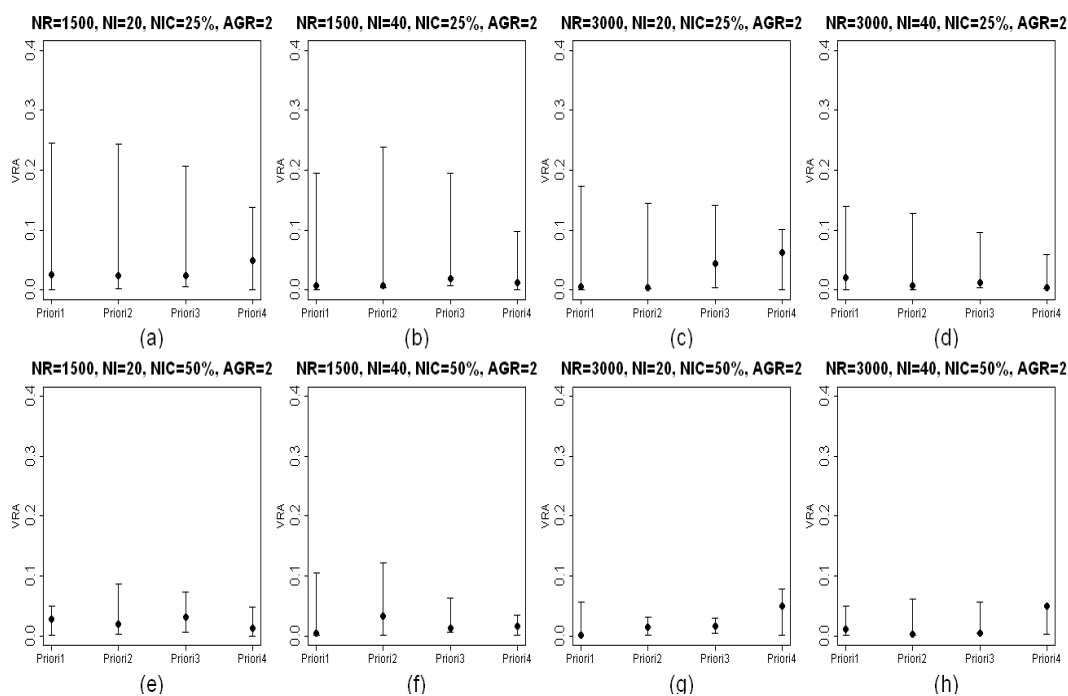


Figura B.20: Valores de VRA das estimativas das médias populacionais considerando  $AGR=2$

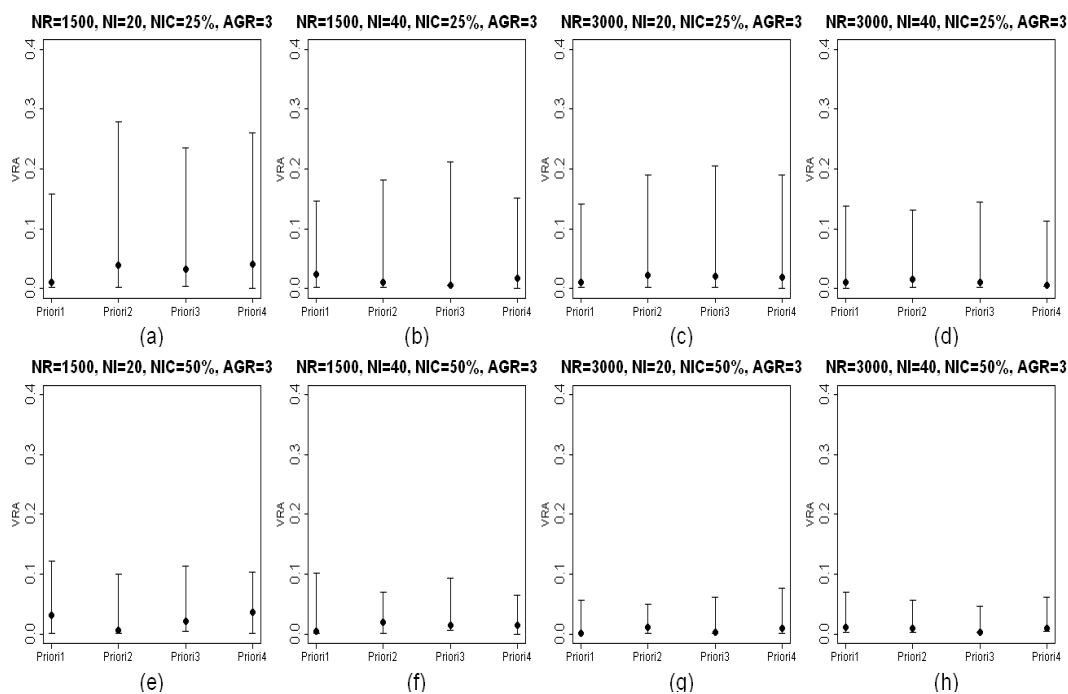


Figura B.21: Valores de VRA das estimativas das médias populacionais considerando  $AGR=3$

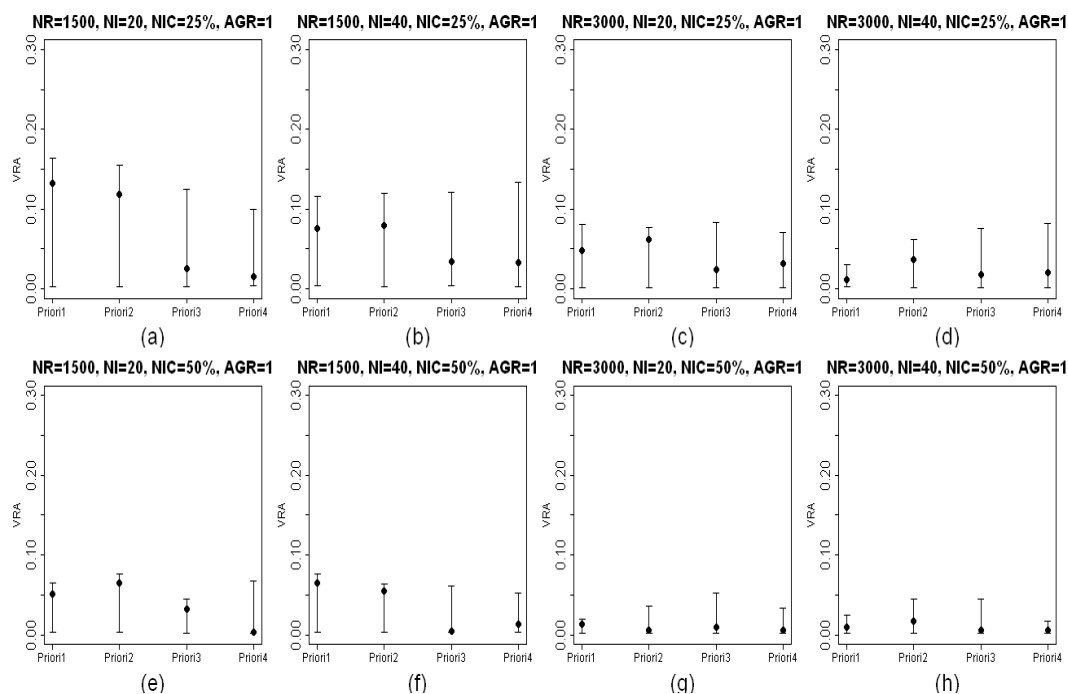


Figura B.22: Valores de VRA das estimativas dos desvios-padrão populacionais considerando  $AGR=1$

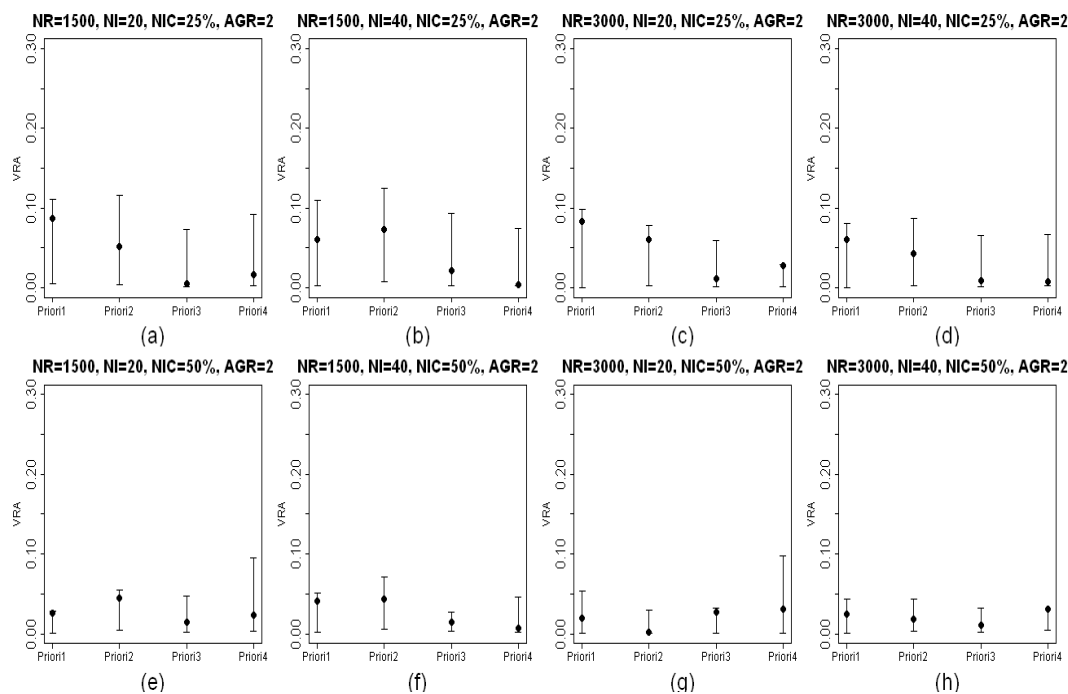


Figura B.23: Valores de VRA das estimativas dos desvios-padrão populacionais considerando  $AGR=2$

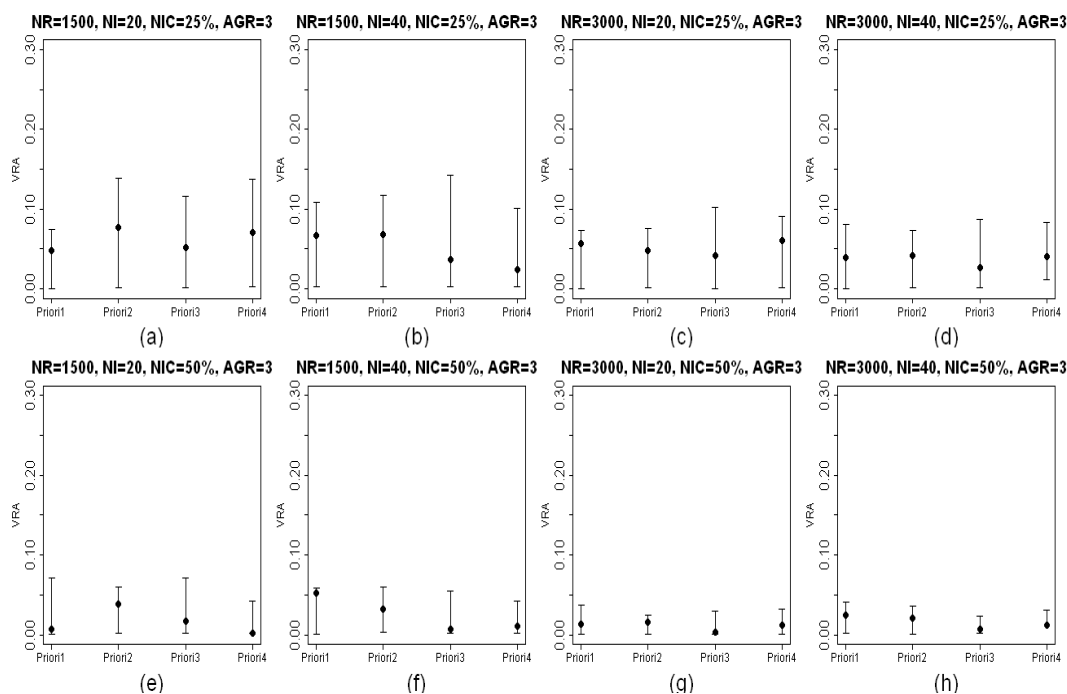


Figura B.24: Valores de VRA das estimativas dos desvios-padrão populacionais considerando  $AGR=3$

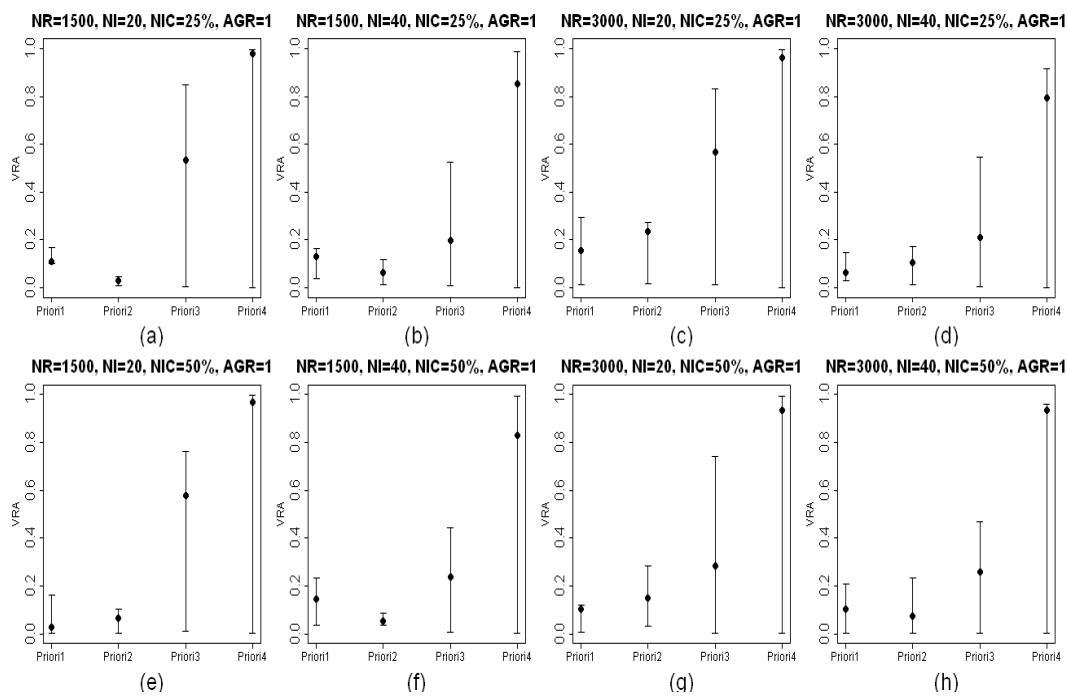


Figura B.25: Valores de VRA das estimativas dos coeficientes de assimetria populacionais considerando  $AGR=1$

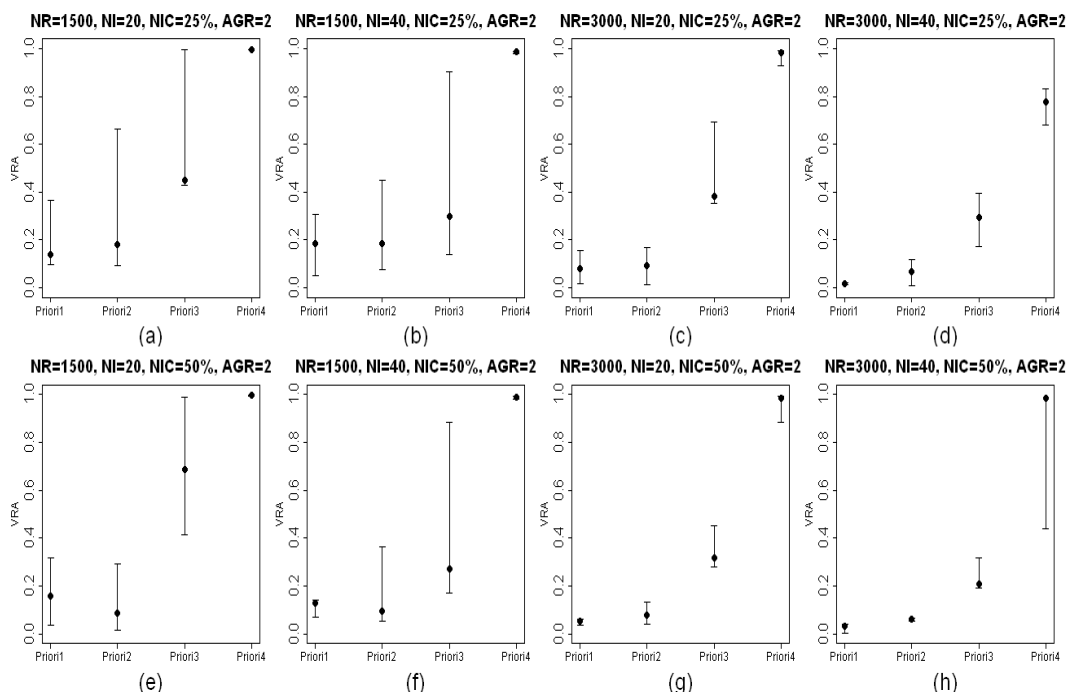


Figura B.26: Valores de VRA das estimativas dos coeficientes de assimetria populacionais considerando  $AGR=2$

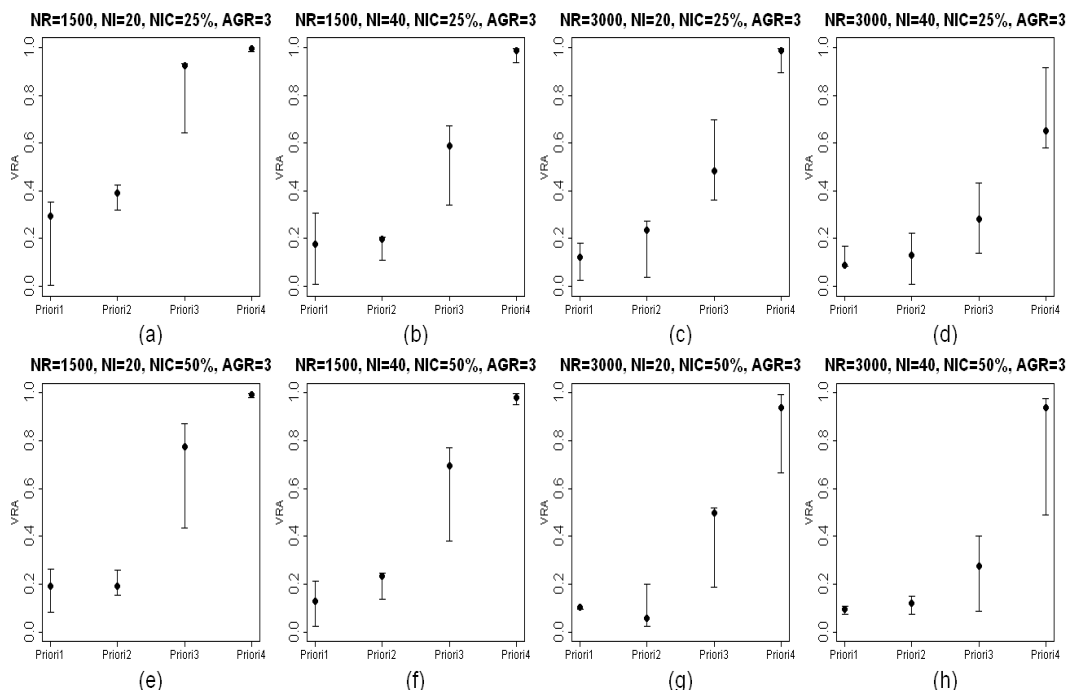


Figura B.27: Valores de VRA das estimativas dos coeficientes de assimetria populacionais considerando  $AGR=3$

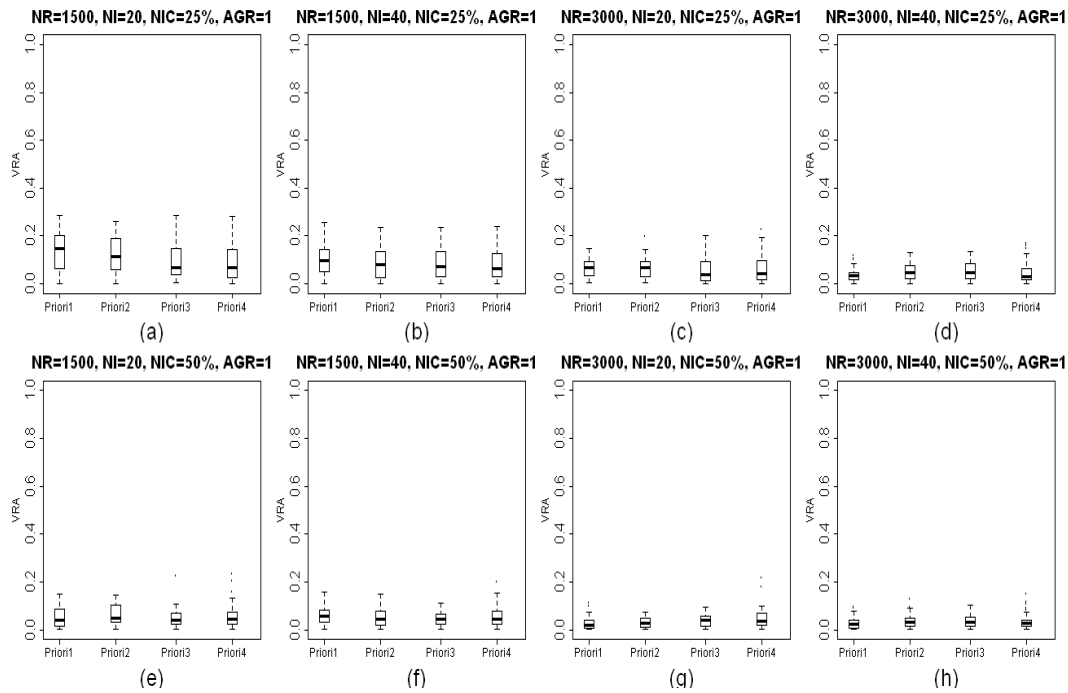


Figura B.28: Valores de VRA das estimativas dos parâmetros de discriminação considerando  $AGR=1$

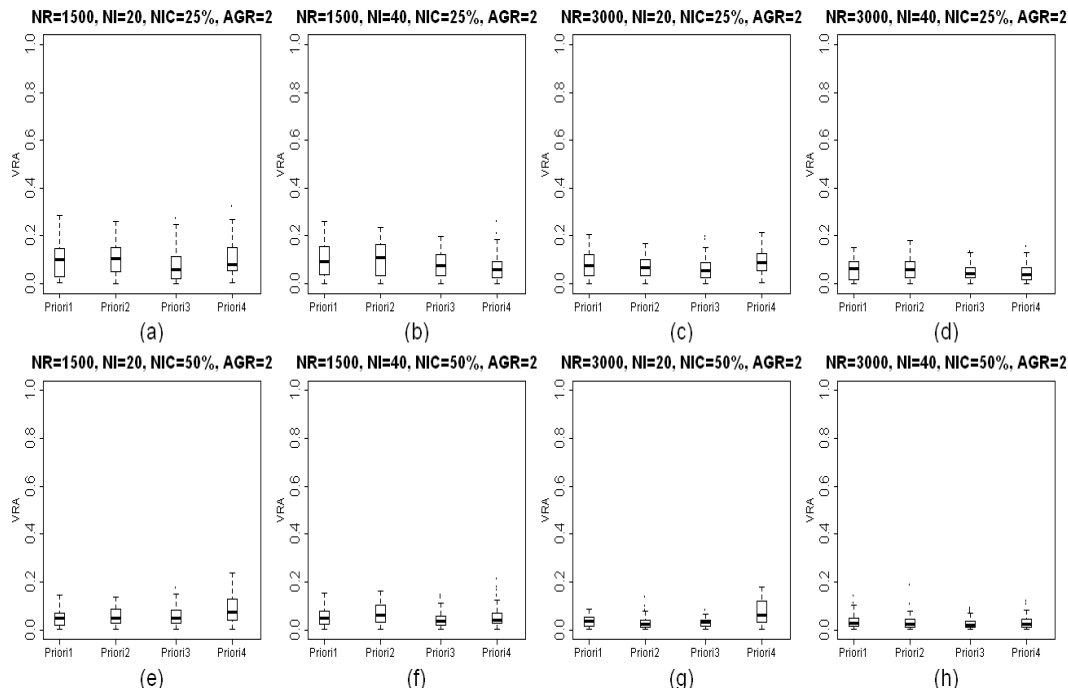


Figura B.29: Valores de VRA das estimativas dos parâmetros de discriminação considerando  $AGR=2$

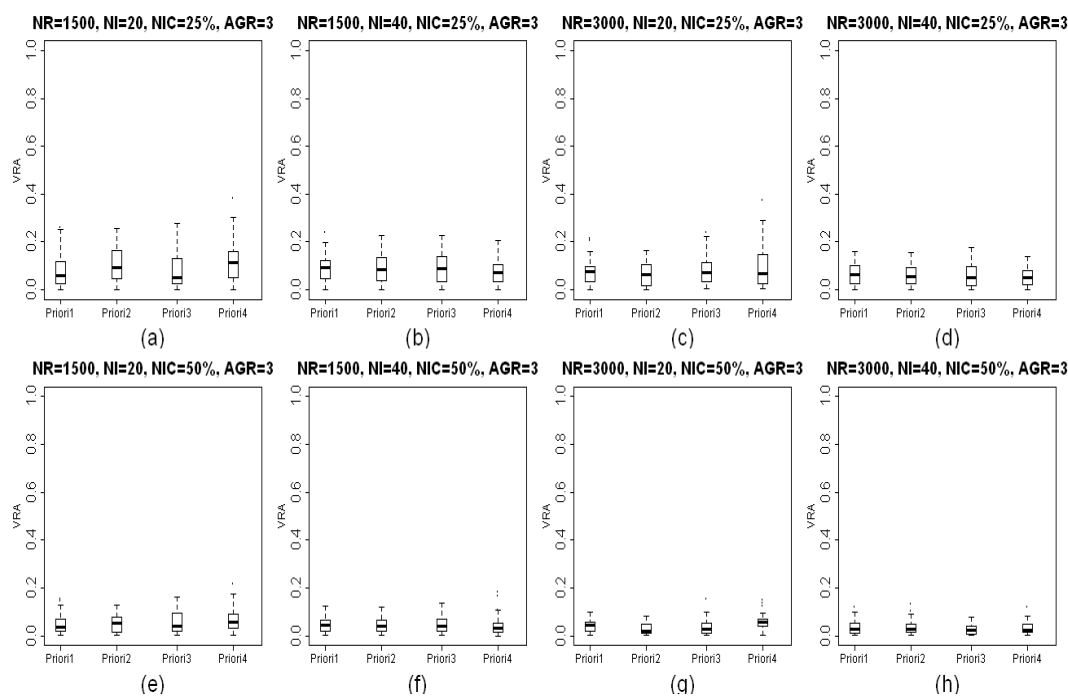


Figura B.30: Valores de VRA das estimativas dos parâmetros de discriminação considerando  $AGR=3$

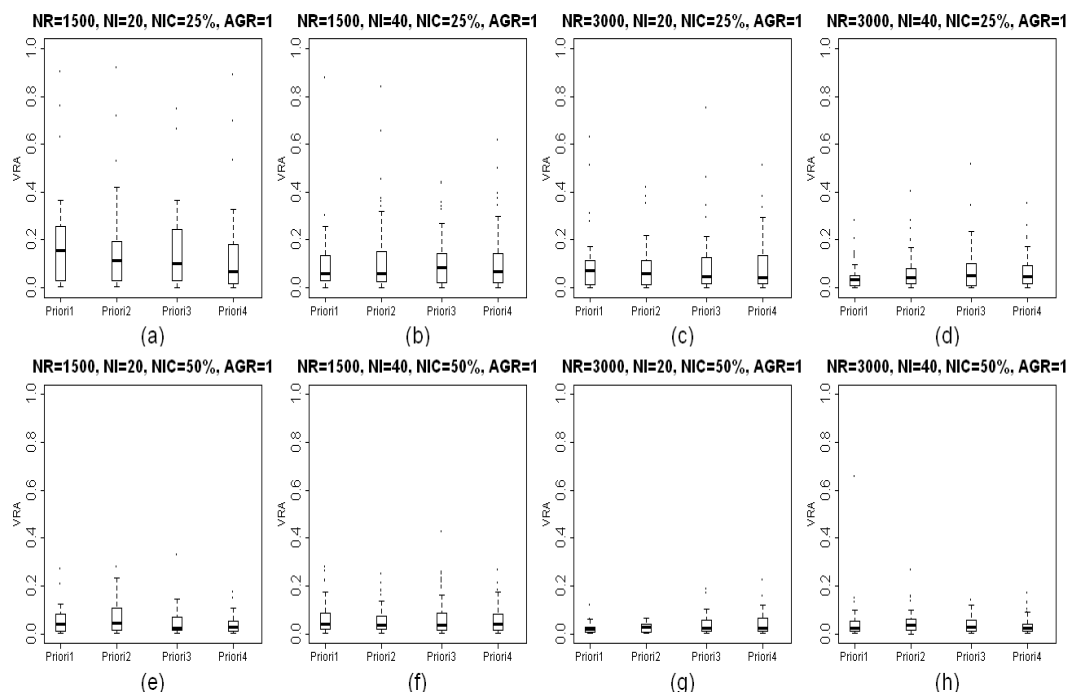


Figura B.31: Valores de VRA das estimativas dos parâmetros de dificuldade considerando  $AGR=1$

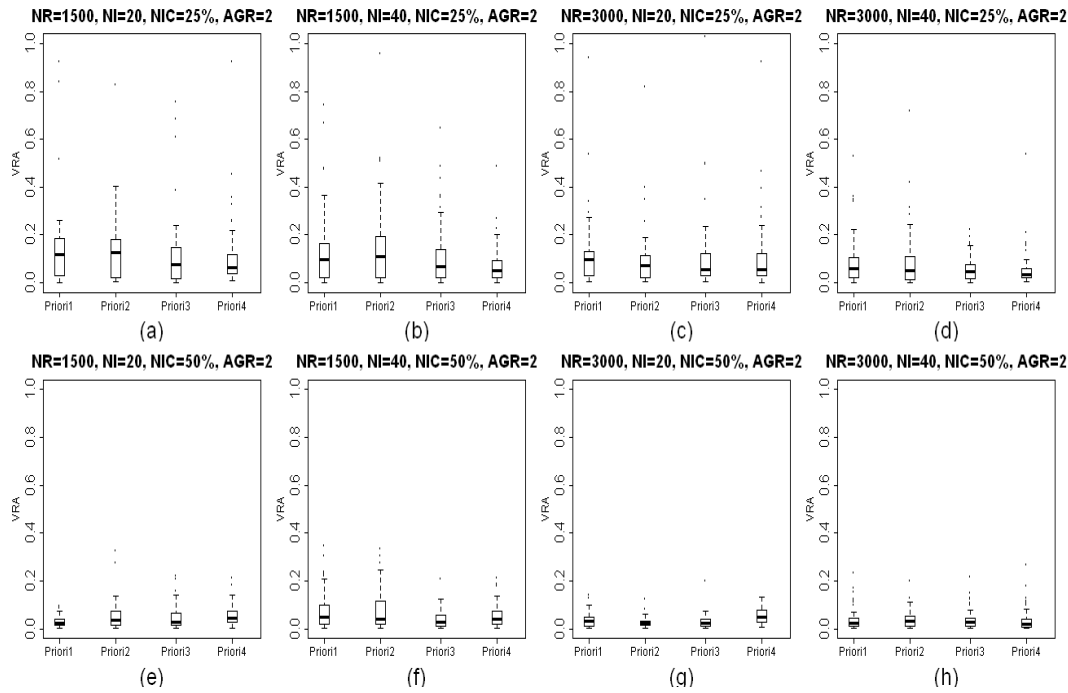


Figura B.32: Valores de VRA das estimativas dos parâmetros de dificuldade considerando  $AGR=2$

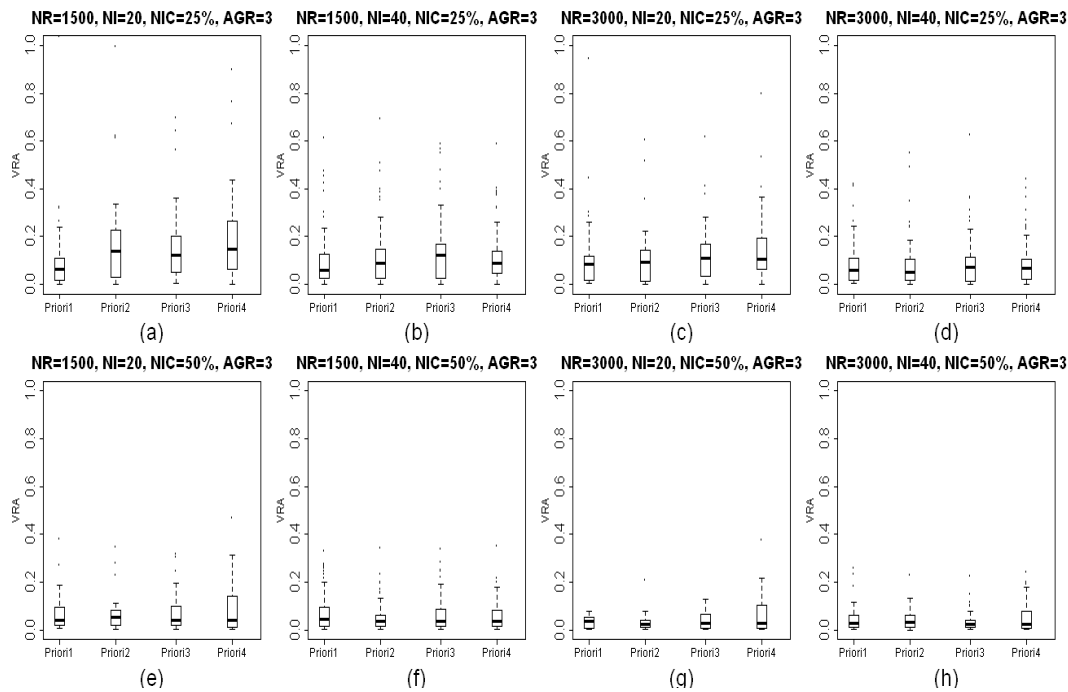


Figura B.33: Valores de VRA das estimativas dos parâmetros de dificuldade considerando  $AGR=3$

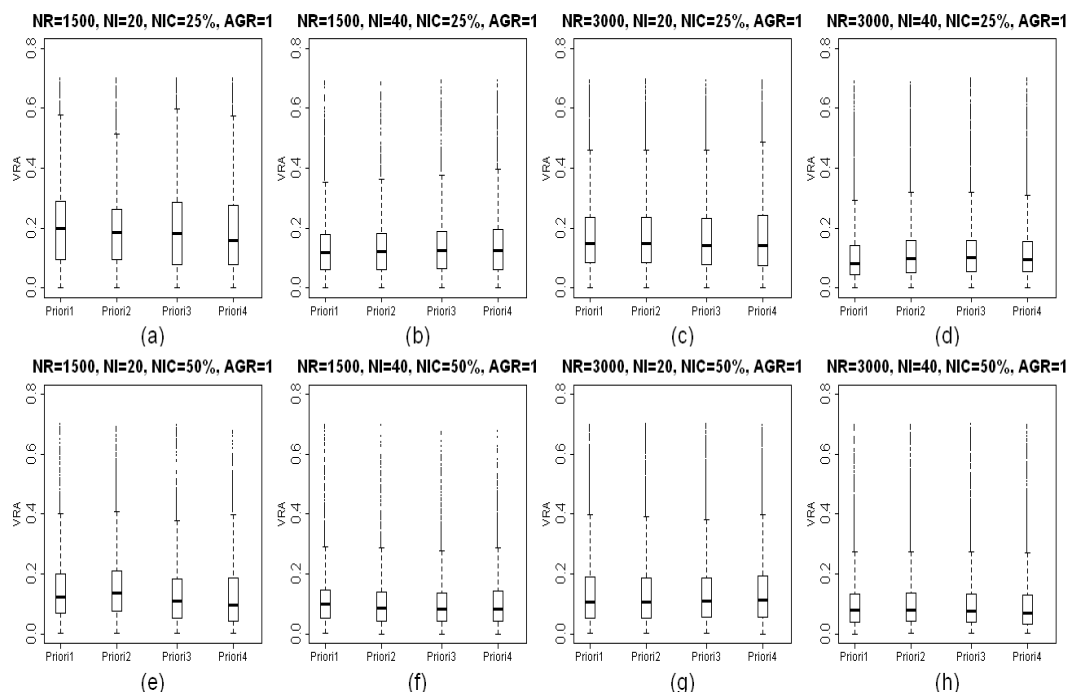


Figura B.34: Valores de VRA das estimativas dos traços latentes considerando  $AGR=1$

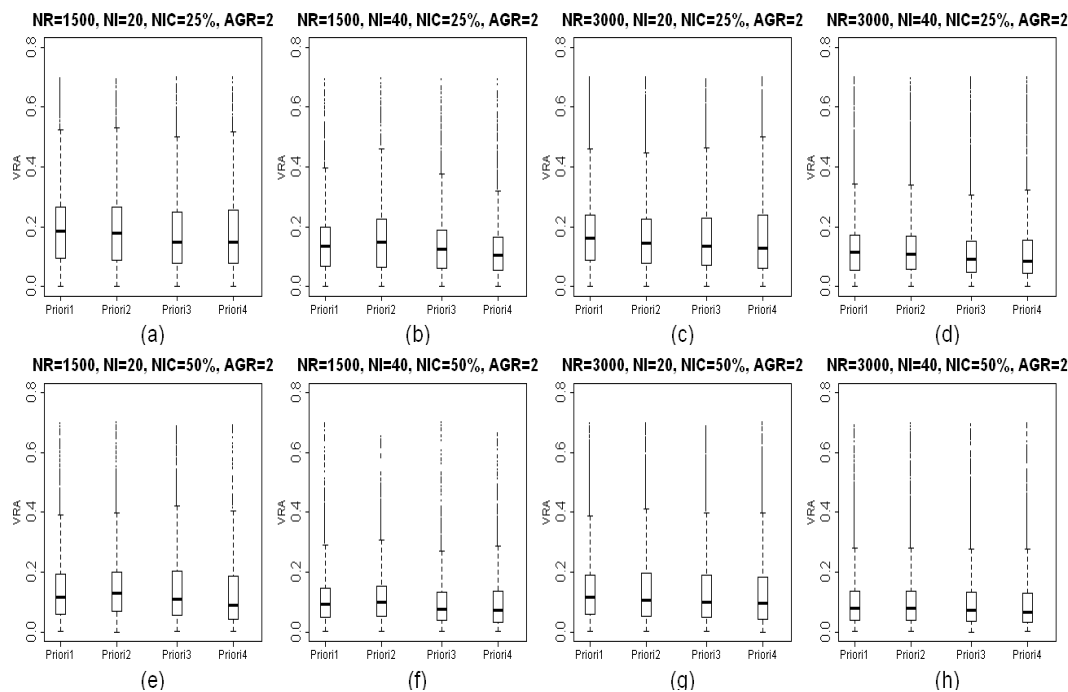


Figura B.35: Valores de VRA das estimativas dos traços latentes considerando  $AGR=2$

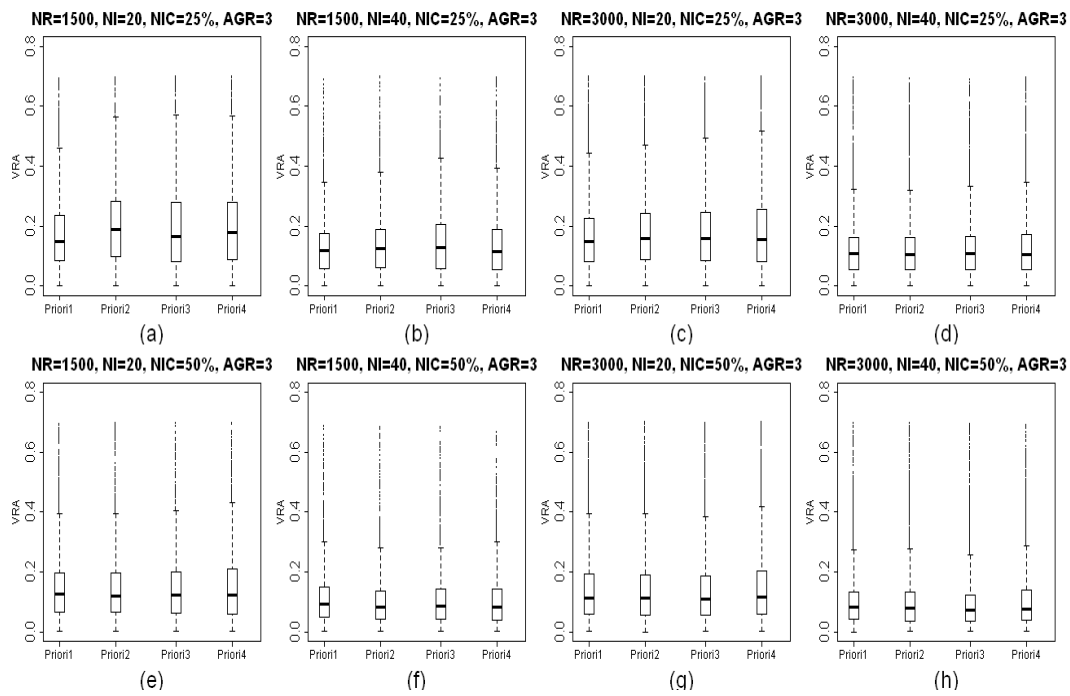


Figura B.36: Valores de VRA das estimativas dos traços latentes considerando  $AGR=3$

# Apêndice C

## Tabelas: estudo de simulação e análise de dados reais

Tabela C.1: Parâmetros dos itens escolhidos quando NI=20 e NIC=50%: Estudo de simulação

| Teste | Item | Parâmetros escolhidos |             | Teste | Item | Parâmetros escolhidos |             |
|-------|------|-----------------------|-------------|-------|------|-----------------------|-------------|
|       |      | Discriminação         | Dificuldade |       |      | Discriminação         | Dificuldade |
| 1     | 1    | 0,70                  | -1,30       | 2-3   | 21   | 1,20                  | 0,65        |
| 1     | 2    | 0,80                  | -1,20       | 2-3   | 22   | 1,25                  | 0,70        |
| 1     | 3    | 0,90                  | -1,10       | 2-3   | 23   | 1,13                  | 0,80        |
| 1     | 4    | 1,30                  | -0,50       | 2-3   | 24   | 1,40                  | 0,90        |
| 1     | 5    | 1,20                  | 1,00        | 2-3   | 25   | 1,10                  | -1,10       |
| 1     | 6    | 1,10                  | 1,30        | 2-3   | 26   | 0,65                  | -0,50       |
| 1     | 7    | 0,87                  | 1,77        | 2-3   | 27   | 0,87                  | -1,30       |
| 1     | 8    | 1,20                  | 1,64        | 2-3   | 28   | 1,14                  | 0,91        |
| 1     | 9    | 1,30                  | 1,90        | 2-3   | 29   | 1,32                  | 0,33        |
| 1     | 10   | 1,35                  | 2,00        | 2-3   | 30   | 0,66                  | 0,20        |
| 1-2   | 11   | 0,72                  | 0,50        | 3     | 31   | 0,89                  | 0,30        |
| 1-2   | 12   | 0,89                  | -0,30       | 3     | 32   | 0,92                  | 1,20        |
| 1-2   | 13   | 1,10                  | -1,30       | 3     | 33   | 1,20                  | 3,00        |
| 1-2   | 14   | 0,90                  | -2,10       | 3     | 34   | 1,03                  | 2,56        |
| 1-2   | 15   | 1,30                  | -1,30       | 3     | 35   | 1,20                  | 2,45        |
| 1-2   | 16   | 1,30                  | 0,90        | 3     | 36   | 1,30                  | 2,10        |
| 1-2   | 17   | 1,10                  | 0,82        | 3     | 37   | 1,36                  | 3,45        |
| 1-2   | 18   | 1,35                  | 0,88        | 3     | 38   | 1,20                  | 2,20        |
| 1-2   | 19   | 1,10                  | 0,90        | 3     | 39   | 1,30                  | 2,50        |
| 1-2   | 20   | 0,96                  | -1,30       | 3     | 40   | 1,40                  | 2,70        |

Tabela C.2: Parâmetros dos itens escolhidos quando NI=20 e NIC=25%: Estudo de simulação

| Teste | Item | Parâmetros escolhidos |             | Teste | Item | Parâmetros escolhidos |             |
|-------|------|-----------------------|-------------|-------|------|-----------------------|-------------|
|       |      | Discriminação         | Dificuldade |       |      | Discriminação         | Dificuldade |
| 1     | 1    | 0,70                  | -1,30       | 2     | 26   | 1,20                  | 0,65        |
| 1     | 2    | 0,80                  | -1,20       | 2     | 27   | 1,25                  | 0,70        |
| 1     | 3    | 0,90                  | -1,10       | 2     | 28   | 1,13                  | 0,80        |
| 1     | 4    | 1,30                  | -0,50       | 2     | 29   | 1,40                  | 0,90        |
| 1     | 5    | 1,20                  | 1,00        | 2     | 30   | 1,10                  | -1,10       |
| 1     | 6    | 1,10                  | 1,30        | 2-3   | 31   | 0,65                  | -0,50       |
| 1     | 7    | 1,20                  | 1,10        | 2-3   | 32   | 0,87                  | -1,30       |
| 1     | 8    | 1,20                  | 2,00        | 2-3   | 33   | 1,14                  | 0,91        |
| 1     | 9    | 1,40                  | 2,30        | 2-3   | 34   | 1,32                  | 0,33        |
| 1     | 10   | 1,35                  | 2,00        | 2-3   | 35   | 0,66                  | 0,20        |
| 1     | 11   | 0,90                  | 1,00        | 3     | 36   | 0,78                  | 1,50        |
| 1     | 12   | 0,89                  | 1,60        | 3     | 37   | 1,10                  | 1,64        |
| 1     | 13   | 1,10                  | 2,10        | 3     | 38   | 1,23                  | 2,30        |
| 1     | 14   | 1,21                  | 1,30        | 3     | 39   | 1,35                  | 2,70        |
| 1     | 15   | 1,30                  | 1,60        | 3     | 40   | 0,98                  | 1,65        |
| 1-2   | 16   | 0,72                  | 0,50        | 3     | 41   | 0,89                  | 0,30        |
| 1-2   | 17   | 0,89                  | -0,30       | 3     | 42   | 0,92                  | 0,10        |
| 1-2   | 18   | 1,10                  | -1,30       | 3     | 43   | 1,20                  | 3,00        |
| 1-2   | 19   | 0,90                  | -2,10       | 3     | 44   | 1,03                  | 2,56        |
| 1-2   | 20   | 1,30                  | -1,30       | 3     | 45   | 1,20                  | 2,45        |
| 2     | 21   | 1,30                  | 0,90        | 3     | 46   | 1,30                  | 2,10        |
| 2     | 22   | 1,10                  | 0,82        | 3     | 47   | 1,36                  | 3,00        |
| 2     | 23   | 1,35                  | 0,88        | 3     | 48   | 1,20                  | 2,20        |
| 2     | 24   | 1,10                  | -1,00       | 3     | 49   | 1,30                  | 2,50        |
| 2     | 25   | 0,96                  | -1,30       | 3     | 50   | 1,40                  | 2,70        |

Tabela C.3: Parâmetros dos itens escolhidos quando NI=40 e NIC=25%: Estudo de simulação

| Teste | Item | Parâmetros escolhidos |             | Teste | Item | Parâmetros escolhidos |             |
|-------|------|-----------------------|-------------|-------|------|-----------------------|-------------|
|       |      | Discriminação         | Dificuldade |       |      | Discriminação         | Dificuldade |
| 1     | 1    | 0,70                  | -1,30       | 2     | 51   | 0,85                  | -1,02       |
| 1     | 2    | 0,80                  | -1,20       | 2     | 52   | 0,75                  | -1,20       |
| 1     | 3    | 0,90                  | -1,10       | 2     | 53   | 0,70                  | -2,00       |
| 1     | 4    | 1,00                  | -0,50       | 2     | 54   | 1,10                  | 0,32        |
| 1     | 5    | 1,25                  | 1,00        | 2     | 55   | 1,10                  | -1,50       |
| 1     | 6    | 1,10                  | 1,30        | 2     | 56   | 0,78                  | 0,33        |
| 1     | 7    | 0,87                  | 1,77        | 2     | 57   | 1,12                  | 0,25        |
| 1     | 8    | 1,20                  | 1,10        | 2     | 58   | 1,04                  | 0,54        |
| 1     | 9    | 1,30                  | 1,90        | 2     | 59   | 1,06                  | 0,80        |
| 1     | 10   | 1,35                  | 2,00        | 2     | 60   | 0,87                  | 0,65        |
| 1     | 11   | 0,65                  | -0,50       | 2-3   | 61   | 0,86                  | 0,40        |
| 1     | 12   | 0,87                  | -1,30       | 2-3   | 62   | 0,91                  | 0,79        |
| 1     | 13   | 1,14                  | 0,91        | 2-3   | 63   | 1,20                  | 0,25        |
| 1     | 14   | 1,32                  | 0,33        | 2-3   | 64   | 1,10                  | -1,30       |
| 1     | 15   | 0,66                  | 0,20        | 2-3   | 65   | 0,74                  | -0,50       |
| 1     | 16   | 0,78                  | 1,50        | 2-3   | 66   | 1,10                  | 0,20        |
| 1     | 17   | 1,20                  | 1,00        | 2-3   | 67   | 1,30                  | 0,89        |
| 1     | 18   | 1,10                  | 1,30        | 2-3   | 68   | 1,35                  | 0,92        |
| 1     | 19   | 0,87                  | 1,77        | 2-3   | 69   | 1,20                  | 0,88        |
| 1     | 20   | 1,20                  | 2,30        | 2-3   | 70   | 0,90                  | 0,70        |
| 1     | 21   | 1,20                  | 2,40        | 3     | 71   | 1,32                  | 0,33        |
| 1     | 22   | 1,35                  | 2,00        | 3     | 72   | 0,66                  | 0,20        |
| 1     | 23   | 1,10                  | 1,20        | 3     | 73   | 0,78                  | 1,50        |
| 1     | 24   | 0,89                  | 1,60        | 3     | 74   | 1,10                  | 1,64        |
| 1     | 25   | 1,10                  | 1,00        | 3     | 75   | 1,23                  | 2,30        |
| 1     | 26   | 1,21                  | 1,30        | 3     | 76   | 1,35                  | 2,70        |
| 1     | 27   | 1,30                  | 1,60        | 3     | 77   | 0,98                  | 1,65        |
| 1     | 28   | 1,14                  | 2,30        | 3     | 78   | 0,89                  | 0,30        |
| 1     | 29   | 1,32                  | 1,33        | 3     | 79   | 0,92                  | 0,10        |
| 1     | 30   | 0,66                  | 0,90        | 3     | 80   | 1,30                  | 2,70        |
| 1-2   | 31   | 0,72                  | 0,50        | 3     | 81   | 1,03                  | 2,56        |
| 1-2   | 32   | 0,89                  | -0,30       | 3     | 82   | 1,20                  | 2,45        |
| 1-2   | 33   | 1,10                  | -1,30       | 3     | 83   | 1,10                  | 2,10        |
| 1-2   | 34   | 0,90                  | -2,10       | 3     | 84   | 1,20                  | 3,00        |
| 1-2   | 35   | 1,30                  | -1,30       | 3     | 85   | 1,10                  | 2,00        |
| 1-2   | 36   | 1,30                  | 0,90        | 3     | 86   | 1,20                  | 3,00        |
| 1-2   | 37   | 1,10                  | 0,82        | 3     | 87   | 1,20                  | 3,30        |
| 1-2   | 38   | 1,35                  | 0,88        | 3     | 88   | 0,95                  | 2,30        |
| 1-2   | 39   | 1,10                  | 0,90        | 3     | 89   | 1,10                  | 2,54        |
| 1-2   | 40   | 0,96                  | -1,30       | 3     | 90   | 1,14                  | 1,80        |
| 2     | 41   | 1,20                  | 0,65        | 3     | 91   | 1,10                  | 1,30        |
| 2     | 42   | 1,25                  | 0,70        | 3     | 92   | 1,10                  | 1,50        |
| 2     | 43   | 1,13                  | 0,80        | 3     | 93   | 1,00                  | 2,50        |
| 2     | 44   | 1,40                  | 0,90        | 3     | 94   | 1,03                  | 2,56        |
| 2     | 45   | 1,10                  | -1,10       | 3     | 95   | 0,90                  | 2,45        |
| 2     | 46   | 0,65                  | -0,50       | 3     | 96   | 1,10                  | 2,10        |
| 2     | 47   | 0,87                  | -1,30       | 3     | 97   | 1,20                  | 3,00        |
| 2     | 48   | 1,14                  | 0,91        | 3     | 98   | 0,90                  | 2,20        |
| 2     | 49   | 1,32                  | 0,33        | 3     | 99   | 1,10                  | 2,50        |
| 2     | 50   | 0,66                  | 0,20        | 3     | 100  | 1,10                  | 2,30        |

Tabela C.4: Parâmetros dos itens escolhidos quando NI=40 e NIC=50%: Estudo de simulação

| Teste | Item | Parâmetros escolhidos |             | Teste | Item | Parâmetros escolhidos |             |
|-------|------|-----------------------|-------------|-------|------|-----------------------|-------------|
|       |      | Discriminação         | Dificuldade |       |      | Discriminação         | Dificuldade |
| 1     | 1    | 0,70                  | -1,30       | 2-3   | 41   | 0,85                  | -1,02       |
| 1     | 2    | 0,80                  | -1,20       | 2-3   | 42   | 0,75                  | -1,20       |
| 1     | 3    | 0,90                  | -1,10       | 2-3   | 43   | 0,70                  | -2,00       |
| 1     | 4    | 1,30                  | -0,50       | 2-3   | 44   | 1,10                  | 0,32        |
| 1     | 5    | 0,89                  | -0,30       | 2-3   | 45   | 1,10                  | -1,50       |
| 1     | 6    | 1,10                  | -1,30       | 2-3   | 46   | 0,78                  | 0,33        |
| 1     | 7    | 0,90                  | -1,15       | 2-3   | 47   | 1,12                  | 0,25        |
| 1     | 8    | 1,30                  | -1,30       | 2-3   | 48   | 1,04                  | 0,54        |
| 1     | 9    | 1,30                  | 0,90        | 2-3   | 49   | 1,06                  | 0,80        |
| 1     | 10   | 1,10                  | -1,10       | 2-3   | 50   | 0,87                  | 0,65        |
| 1     | 11   | 0,65                  | -0,50       | 2-3   | 51   | 0,86                  | 0,40        |
| 1     | 12   | 0,87                  | -1,30       | 2-3   | 52   | 0,91                  | 0,79        |
| 1     | 13   | 1,14                  | 0,91        | 2-3   | 53   | 1,20                  | 0,25        |
| 1     | 14   | 1,32                  | 0,33        | 2-3   | 54   | 1,10                  | -1,30       |
| 1     | 15   | 0,66                  | 0,20        | 2-3   | 55   | 0,74                  | -0,50       |
| 1     | 16   | 0,78                  | 1,50        | 2-3   | 56   | 1,10                  | 0,20        |
| 1     | 17   | 1,20                  | 1,00        | 2-3   | 57   | 1,30                  | 0,89        |
| 1     | 18   | 1,10                  | 1,30        | 2-3   | 58   | 1,35                  | 0,92        |
| 1     | 19   | 0,87                  | 1,77        | 2-3   | 59   | 1,20                  | 0,88        |
| 1     | 20   | 1,20                  | 1,10        | 2-3   | 60   | 0,90                  | 0,70        |
| 1-2   | 21   | 0,72                  | 0,50        | 3     | 61   | 1,32                  | 0,33        |
| 1-2   | 22   | 0,89                  | -0,30       | 3     | 62   | 0,66                  | 0,20        |
| 1-2   | 23   | 1,10                  | -1,30       | 3     | 63   | 0,78                  | 1,50        |
| 1-2   | 24   | 0,90                  | -2,10       | 3     | 64   | 1,10                  | 1,64        |
| 1-2   | 25   | 1,30                  | -1,30       | 3     | 65   | 1,23                  | 2,30        |
| 1-2   | 26   | 1,30                  | 0,90        | 3     | 66   | 1,35                  | 2,70        |
| 1-2   | 27   | 1,10                  | 0,82        | 3     | 67   | 0,98                  | 1,65        |
| 1-2   | 28   | 1,35                  | 0,88        | 3     | 68   | 0,89                  | 0,30        |
| 1-2   | 29   | 1,10                  | 0,90        | 3     | 69   | 0,92                  | 0,50        |
| 1-2   | 30   | 0,96                  | -1,30       | 3     | 70   | 1,40                  | 3,20        |
| 1-2   | 31   | 1,20                  | 0,65        | 3     | 71   | 1,20                  | 3,30        |
| 1-2   | 32   | 1,25                  | 0,70        | 3     | 72   | 0,95                  | 2,30        |
| 1-2   | 33   | 1,13                  | 0,80        | 3     | 73   | 1,10                  | 2,54        |
| 1-2   | 34   | 1,40                  | 0,90        | 3     | 74   | 1,14                  | 1,80        |
| 1-2   | 35   | 1,10                  | -1,10       | 3     | 75   | 1,20                  | 3,10        |
| 1-2   | 36   | 0,65                  | -0,50       | 3     | 76   | 1,22                  | 3,58        |
| 1-2   | 37   | 0,87                  | -1,30       | 3     | 77   | 1,34                  | 3,65        |
| 1-2   | 38   | 1,14                  | 0,91        | 3     | 78   | 1,10                  | 2,13        |
| 1-2   | 39   | 1,32                  | 0,33        | 3     | 79   | 1,20                  | 2,50        |
| 1-2   | 40   | 0,66                  | 0,20        | 3     | 80   | 1,30                  | 3,10        |

Tabela C.5: ANOVA para os traços latentes considerando todas interações: Estudo de simulação

| Fonte de variação       | gl     | SQ        | QM       | $\omega^2$ |
|-------------------------|--------|-----------|----------|------------|
| NE                      | 1      | 43,100    | 43,100   | 0,002      |
| NI                      | 1      | 4781,160  | 4781,160 | 0,197      |
| NIC                     | 1      | 334,860   | 334,860  | 0,014      |
| AGR                     | 2      | 0,920     | 0,460    | <0,001     |
| P                       | 3      | 2,770     | 0,920    | <0,001     |
| NE x NI                 | 1      | 1,230     | 1,230    | <0,001     |
| NE x NIC                | 1      | 8,460     | 8,460    | <0,001     |
| NI x NIC                | 1      | 0,040     | 0,040    | <0,001     |
| NE x AGR                | 2      | 0,480     | 0,240    | <0,001     |
| NI x AGR                | 2      | 0,320     | 0,160    | <0,001     |
| NIC x AGR               | 2      | 3,870     | 1,930    | <0,001     |
| NE x P                  | 3      | 1,690     | 0,560    | <0,001     |
| NI x P                  | 3      | 1,050     | 0,350    | <0,001     |
| NIC x P                 | 3      | 2,000     | 0,670    | <0,001     |
| AGR x P                 | 6      | 1,350     | 0,220    | <0,001     |
| NE x NI x NIC           | 1      | 0,240     | 0,240    | <0,001     |
| NE x NI x AGR           | 2      | 0,570     | 0,290    | <0,001     |
| NE x NIC x AGR          | 2      | 2,970     | 1,480    | <0,001     |
| NI x NIC x AGR          | 2      | 0,040     | 0,020    | <0,001     |
| NE x NI x P             | 3      | 0,420     | 0,140    | <0,001     |
| NE x NIC x P            | 3      | 0,270     | 0,090    | <0,001     |
| NI x NIC x P            | 3      | 0,230     | 0,080    | <0,001     |
| NE x AGR x P            | 6      | 2,210     | 0,370    | <0,001     |
| NI x AGR x P            | 6      | 2,910     | 0,490    | <0,001     |
| NIC x AGR x P           | 6      | 4,690     | 0,780    | <0,001     |
| NE x NI x NIC x AGR     | 2      | 0,040     | 0,020    | <0,001     |
| NE x NI x NIC x P       | 3      | 0,700     | 0,230    | <0,001     |
| NE x NI x AGR x P       | 6      | 0,670     | 0,110    | <0,001     |
| NE x NIC x AGR x P      | 6      | 2,830     | 0,470    | <0,001     |
| NI x NIC x AGR x P      | 6      | 2,440     | 0,410    | <0,001     |
| NE x NI x NIC x AGR x P | 6      | 2,650     | 0,440    | <0,001     |
| Resíduos                | 215904 | 19078,440 | 0,090    |            |

Tabela C.6: ANOVA para os parâmetros de discriminação considerando todas interações: Estudo de simulação

| Fonte de variação       | gl   | SQ       | QM      | $\omega^2$ |
|-------------------------|------|----------|---------|------------|
| NE                      | 1    | 201,613  | 201,613 | 0,117      |
| NI                      | 1    | 51,738   | 51,739  | 0,030      |
| NIC                     | 1    | 161,045  | 161,045 | 0,093      |
| AGR                     | 2    | 0,847    | 0,423   | <0,001     |
| P                       | 3    | 4,715    | 1,572   | 0,002      |
| NE x NI                 | 1    | 1,845    | 1,845   | 0,001      |
| NE x NIC                | 1    | 0,010    | 0,010   | <0,001     |
| NI x NIC                | 1    | 0,052    | 0,052   | <0,001     |
| NE x AGR                | 2    | 1,817    | 0,908   | 0,001      |
| NI x AGR                | 2    | 0,037    | 0,018   | <0,001     |
| NIC x AGR               | 2    | 2,001    | 1,001   | 0,001      |
| NE x P                  | 3    | 2,011    | 0,670   | 0,001      |
| NI x P                  | 3    | 0,959    | 0,320   | <0,001     |
| NIC x P                 | 3    | 0,642    | 0,214   | <0,001     |
| AGR x P                 | 6    | 1,133    | 0,189   | <0,001     |
| NE x NI x NIC           | 1    | <0,001   | <0,001  | <0,001     |
| NE x NI x AGR           | 2    | 0,066    | 0,033   | <0,001     |
| NE x NIC x AGR          | 2    | 0,839    | 0,419   | <0,001     |
| NI x NIC x AGR          | 2    | 0,199    | 0,099   | <0,001     |
| NE x NI x P             | 3    | 0,950    | 0,317   | <0,001     |
| NE x NIC x P            | 3    | 0,128    | 0,043   | <0,001     |
| NI x NIC x P            | 3    | 0,425    | 0,142   | <0,001     |
| NE x AGR x P            | 6    | 2,524    | 0,421   | <0,001     |
| NI x AGR x P            | 6    | 2,158    | 0,360   | 0,001      |
| NIC x AGR x P           | 6    | 0,456    | 0,076   | <0,001     |
| NE x NI x NIC x AGR     | 2    | 0,338    | 0,169   | <0,001     |
| NE x NI x NIC x P       | 3    | 0,117    | 0,039   | <0,001     |
| NE x NI x AGR x P       | 6    | 0,320    | 0,053   | <0,001     |
| NE x NIC x AGR x P      | 6    | 0,299    | 0,050   | <0,001     |
| NI x NIC x AGR x P      | 6    | 0,360    | 0,060   | <0,001     |
| NE x NI x NIC x AGR x P | 6    | 0,226    | 0,038   | <0,001     |
| Resíduos                | 6384 | 1286,158 | 0,202   |            |

Tabela C.7: ANOVA para os parâmetros de dificuldade considerando todas interações: Estudo de simulação

| Fonte de variação       | gl   | SQ       | QM      | $\omega^2$ |
|-------------------------|------|----------|---------|------------|
| NE                      | 1    | 191,632  | 191,632 | 0,068      |
| NI                      | 1    | 50,454   | 50,454  | 0,018      |
| NIC                     | 1    | 273,307  | 273,307 | 0,098      |
| AGR                     | 2    | 4,313    | 2,156   | 0,001      |
| P                       | 3    | 0,308    | 0,103   | <0,001     |
| NE x NI                 | 1    | 0,930    | 0,930   | <0,001     |
| NE x NIC                | 1    | 0,009    | 0,009   | <0,001     |
| NI x NIC                | 1    | 0,005    | 0,005   | <0,001     |
| NE x AGR                | 2    | 1,398    | 0,699   | <0,001     |
| NI x AGR                | 2    | 0,456    | 0,228   | <0,001     |
| NIC x AGR               | 2    | 4,112    | 2,056   | 0,001      |
| NE x P                  | 3    | 1,425    | 0,475   | <0,001     |
| NI x P                  | 3    | 0,194    | 0,065   | <0,001     |
| NIC x P                 | 3    | 1,055    | 0,352   | <0,001     |
| AGR x P                 | 6    | 4,676    | 0,779   | 0,001      |
| NE x NI x NIC           | 1    | 0,004    | 0,004   | <0,001     |
| NE x NI x AGR           | 2    | 0,399    | 0,200   | <0,001     |
| NE x NIC x AGR          | 2    | 1,100    | 0,550   | <0,001     |
| NI x NIC x AGR          | 2    | 0,079    | 0,040   | <0,001     |
| NE x NI x P             | 3    | 0,149    | 0,050   | <0,001     |
| NE x NIC x P            | 3    | 0,402    | 0,134   | <0,001     |
| NI x NIC x P            | 3    | 0,082    | 0,027   | <0,001     |
| NE x AGR x P            | 6    | 1,230    | 0,205   | <0,001     |
| NI x AGR x P            | 6    | 2,766    | 0,461   | <0,001     |
| NIC x AGR x P           | 6    | 5,420    | 0,903   | 0,001      |
| NE x NI x NIC x AGR     | 2    | 0,876    | 0,438   | <0,001     |
| NE x NI x NIC x P       | 3    | 0,219    | 0,073   | <0,001     |
| NE x NI x AGR x P       | 6    | 0,344    | 0,057   | <0,001     |
| NE x NIC x AGR x P      | 6    | 0,997    | 0,166   | <0,001     |
| NI x NIC x AGR x P      | 6    | 1,303    | 0,217   | <0,001     |
| NE x NI x NIC x AGR x P | 6    | 1,268    | 0,211   | <0,001     |
| Resíduos                | 6384 | 2246,615 | 0,352   |            |

Tabela C.8: Estimativas dos parâmetros de discriminação e dificuldade: análise de dados reais

| Testes | Item | Discriminação |       |          |       |           |       | Dificuldade |       |          |        |           |        |
|--------|------|---------------|-------|----------|-------|-----------|-------|-------------|-------|----------|--------|-----------|--------|
|        |      | Média         | EP    | IC (95%) |       | HPD (95%) |       | Média       | EP    | IC (95%) |        | HPD (95%) |        |
| 1-2    | 1    | 0,486         | 0,056 | 0,379    | 0,593 | 0,377     | 0,592 | -2,453      | 0,308 | -3,129   | -1,951 | -3,077    | -1,908 |
| 1-2    | 2    | 0,268         | 0,043 | 0,187    | 0,357 | 0,186     | 0,352 | -3,155      | 0,634 | -4,601   | -2,216 | -4,314    | -2,068 |
| 1-2    | 3    | 0,488         | 0,057 | 0,380    | 0,600 | 0,382     | 0,600 | -2,673      | 0,342 | -3,405   | -2,111 | -3,303    | -2,064 |
| 1-2    | 4    | 0,937         | 0,070 | 0,796    | 1,078 | 0,805     | 1,083 | -0,653      | 0,077 | -0,810   | -0,513 | -0,810    | -0,512 |
| 1-2    | 5    | 0,549         | 0,058 | 0,438    | 0,663 | 0,442     | 0,664 | -2,108      | 0,250 | -2,626   | -1,689 | -2,596    | -1,666 |
| 1-2    | 6    | 1,116         | 0,075 | 0,977    | 1,262 | 0,977     | 1,262 | 0,169       | 0,046 | 0,079    | 0,258  | 0,089     | 0,265  |
| 1-2    | 7    | 0,907         | 0,065 | 0,788    | 1,041 | 0,784     | 1,032 | -0,406      | 0,071 | -0,550   | -0,278 | -0,546    | -0,277 |
| 1-2    | 8    | 0,987         | 0,066 | 0,858    | 1,119 | 0,857     | 1,118 | 0,516       | 0,049 | 0,424    | 0,608  | 0,424     | 0,608  |
| 1-2    | 9    | 0,761         | 0,056 | 0,652    | 0,876 | 0,649     | 0,872 | 0,635       | 0,056 | 0,523    | 0,742  | 0,521     | 0,741  |
| 1-2    | 10   | 1,028         | 0,076 | 0,888    | 1,183 | 0,870     | 1,163 | 1,190       | 0,056 | 1,078    | 1,301  | 1,083     | 1,305  |
| 1-2    | 11   | 0,500         | 0,045 | 0,415    | 0,591 | 0,418     | 0,592 | -0,461      | 0,114 | -0,699   | -0,250 | -0,679    | -0,241 |
| 1-2    | 12   | 0,732         | 0,054 | 0,629    | 0,840 | 0,629     | 0,837 | 0,834       | 0,062 | 0,716    | 0,956  | 0,721     | 0,960  |
| 1-2    | 13   | 0,718         | 0,057 | 0,612    | 0,835 | 0,612     | 0,836 | -0,546      | 0,088 | -0,722   | -0,391 | -0,713    | -0,383 |
| 1-2    | 14   | 0,837         | 0,065 | 0,720    | 0,970 | 0,723     | 0,971 | -0,740      | 0,088 | -0,924   | -0,582 | -0,924    | -0,582 |
| 1-2    | 15   | 0,796         | 0,063 | 0,671    | 0,920 | 0,670     | 0,920 | -0,733      | 0,091 | -0,932   | -0,576 | -0,916    | -0,563 |
| 1-2    | 16   | 0,989         | 0,068 | 0,862    | 1,130 | 0,858     | 1,119 | -0,001      | 0,058 | -0,116   | 0,105  | -0,107    | 0,111  |
| 1-2    | 17   | 0,877         | 0,063 | 0,760    | 1,003 | 0,753     | 0,993 | 0,116       | 0,060 | -0,007   | 0,232  | -0,009    | 0,225  |
| 1-2    | 18   | 0,959         | 0,067 | 0,841    | 1,095 | 0,839     | 1,091 | 0,404       | 0,049 | 0,308    | 0,497  | 0,309     | 0,498  |
| 1-2    | 19   | 1,069         | 0,074 | 0,936    | 1,221 | 0,932     | 1,217 | 0,808       | 0,047 | 0,716    | 0,898  | 0,716     | 0,898  |
| 1-2    | 20   | 1,113         | 0,082 | 0,967    | 1,286 | 0,975     | 1,294 | 0,823       | 0,045 | 0,736    | 0,914  | 0,734     | 0,910  |
| 2-3    | 21   | 0,868         | 0,090 | 0,691    | 1,042 | 0,679     | 1,026 | -0,775      | 0,178 | -1,164   | -0,472 | -1,166    | -0,472 |
| 2-3    | 22   | 0,699         | 0,079 | 0,548    | 0,863 | 0,555     | 0,867 | -0,699      | 0,202 | -1,155   | -0,366 | -1,134    | -0,348 |
| 2-3    | 23   | 0,897         | 0,083 | 0,740    | 1,070 | 0,730     | 1,050 | 0,168       | 0,099 | -0,045   | 0,350  | 0,004     | 0,377  |
| 2-3    | 24   | 0,786         | 0,076 | 0,645    | 0,939 | 0,647     | 0,939 | -0,118      | 0,124 | -0,405   | 0,093  | -0,378    | 0,113  |
| 2-3    | 25   | 0,795         | 0,078 | 0,645    | 0,949 | 0,638     | 0,941 | 0,863       | 0,072 | 0,719    | 0,999  | 0,733     | 1,011  |
| 2-3    | 26   | 1,457         | 0,118 | 1,245    | 1,690 | 1,238     | 1,673 | 0,995       | 0,052 | 0,897    | 1,094  | 0,897     | 1,092  |
| 2-3    | 27   | 0,871         | 0,089 | 0,700    | 1,052 | 0,707     | 1,059 | 1,584       | 0,076 | 1,434    | 1,738  | 1,446     | 1,746  |
| 2-3    | 28   | 0,832         | 0,086 | 0,676    | 1,003 | 0,676     | 1,003 | -0,600      | 0,166 | -0,951   | -0,317 | -0,951    | -0,318 |
| 2-3    | 29   | 1,405         | 0,119 | 1,189    | 1,646 | 1,186     | 1,642 | 0,461       | 0,064 | 0,337    | 0,581  | 0,337     | 0,581  |
| 2-3    | 30   | 0,979         | 0,087 | 0,809    | 1,160 | 0,820     | 1,168 | 0,822       | 0,066 | 0,690    | 0,949  | 0,692     | 0,951  |
| 2-3    | 31   | 1,070         | 0,092 | 0,896    | 1,254 | 0,887     | 1,238 | 0,350       | 0,080 | 0,179    | 0,495  | 0,201     | 0,506  |
| 2-3    | 32   | 1,277         | 0,109 | 1,065    | 1,494 | 1,074     | 1,503 | 0,857       | 0,058 | 0,746    | 0,972  | 0,752     | 0,977  |
| 2-3    | 33   | 1,038         | 0,096 | 0,856    | 1,229 | 0,861     | 1,231 | 1,572       | 0,067 | 1,447    | 1,698  | 1,447     | 1,699  |
| 2-3    | 34   | 0,891         | 0,149 | 0,615    | 1,201 | 0,593     | 1,166 | 3,183       | 0,289 | 2,720    | 3,857  | 2,660     | 3,759  |
| 2-3    | 35   | 1,086         | 0,098 | 0,914    | 1,284 | 0,913     | 1,284 | 1,321       | 0,060 | 1,211    | 1,453  | 1,210     | 1,452  |
| 2-3    | 36   | 1,070         | 0,096 | 0,898    | 1,258 | 0,896     | 1,254 | 0,906       | 0,061 | 0,782    | 1,020  | 0,789     | 1,024  |
| 2-3    | 37   | 1,174         | 0,109 | 0,976    | 1,393 | 0,964     | 1,373 | 1,350       | 0,059 | 1,235    | 1,467  | 1,232     | 1,463  |
| 2-3    | 38   | 1,115         | 0,098 | 0,941    | 1,319 | 0,916     | 1,289 | 0,452       | 0,071 | 0,310    | 0,585  | 0,317     | 0,589  |
| 2-3    | 39   | 0,739         | 0,081 | 0,589    | 0,904 | 0,574     | 0,883 | 1,610       | 0,085 | 1,455    | 1,792  | 1,452     | 1,781  |
| 2-3    | 40   | 1,416         | 0,148 | 1,139    | 1,714 | 1,134     | 1,705 | 2,011       | 0,087 | 1,845    | 2,190  | 1,844     | 2,190  |
| 3-4    | 41   | 1,266         | 0,130 | 1,035    | 1,536 | 1,039     | 1,539 | 1,286       | 0,070 | 1,149    | 1,423  | 1,151     | 1,425  |
| 3-4    | 42   | 1,202         | 0,127 | 0,964    | 1,472 | 0,956     | 1,461 | 2,399       | 0,107 | 2,196    | 2,618  | 2,177     | 2,591  |
| 3-4    | 43   | 1,214         | 0,124 | 0,976    | 1,452 | 0,970     | 1,443 | 2,250       | 0,098 | 2,069    | 2,453  | 2,055     | 2,435  |
| 3-4    | 44   | 1,231         | 0,125 | 0,988    | 1,484 | 0,977     | 1,468 | 2,458       | 0,113 | 2,241    | 2,685  | 2,253     | 2,693  |
| 3-4    | 45   | 1,669         | 0,175 | 1,349    | 2,048 | 1,329     | 2,013 | 0,928       | 0,070 | 0,783    | 1,062  | 0,780     | 1,062  |
| 3-4    | 46   | 1,069         | 0,113 | 0,853    | 1,298 | 0,834     | 1,264 | 0,997       | 0,088 | 0,821    | 1,158  | 0,833     | 1,169  |
| 3-4    | 47   | 0,814         | 0,097 | 0,634    | 1,018 | 0,622     | 0,997 | 1,293       | 0,087 | 1,123    | 1,457  | 1,137     | 1,468  |
| 3-4    | 48   | 0,757         | 0,141 | 0,493    | 1,037 | 0,502     | 1,040 | -0,763      | 0,431 | -1,828   | -0,160 | -1,587    | -0,039 |
| 3-4    | 49   | 1,041         | 0,118 | 0,825    | 1,281 | 0,825     | 1,283 | 0,904       | 0,091 | 0,716    | 1,078  | 0,722     | 1,082  |
| 3-4    | 50   | 1,133         | 0,116 | 0,916    | 1,371 | 0,922     | 1,374 | 1,515       | 0,072 | 1,381    | 1,660  | 1,366     | 1,643  |
| 3-4    | 51   | 1,019         | 0,104 | 0,828    | 1,233 | 0,808     | 1,204 | 2,158       | 0,098 | 1,976    | 2,353  | 1,976     | 2,355  |
| 3-4    | 52   | 0,536         | 0,083 | 0,382    | 0,701 | 0,377     | 0,694 | 1,141       | 0,132 | 0,843    | 1,374  | 0,870     | 1,385  |
| 3-4    | 53   | 1,082         | 0,139 | 0,828    | 1,364 | 0,833     | 1,367 | 0,468       | 0,130 | 0,201    | 0,693  | 0,213     | 0,701  |
| 3-4    | 54   | 1,336         | 0,137 | 1,085    | 1,624 | 1,080     | 1,616 | 2,545       | 0,116 | 2,330    | 2,785  | 2,313     | 2,758  |
| 3-4    | 55   | 1,340         | 0,134 | 1,098    | 1,617 | 1,098     | 1,617 | 1,341       | 0,069 | 1,205    | 1,474  | 1,196     | 1,460  |
| 3-4    | 56   | 1,039         | 0,113 | 0,838    | 1,270 | 0,836     | 1,267 | 1,147       | 0,084 | 0,974    | 1,309  | 0,982     | 1,317  |
| 3-4    | 57   | 0,587         | 0,097 | 0,409    | 0,782 | 0,399     | 0,768 | 0,390       | 0,223 | -0,130   | 0,742  | -0,070    | 0,765  |
| 3-4    | 58   | 0,391         | 0,079 | 0,242    | 0,556 | 0,234     | 0,543 | 0,469       | 0,297 | -0,229   | 0,923  | -0,095    | 0,986  |
| 3-4    | 59   | 0,791         | 0,095 | 0,622    | 0,972 | 0,621     | 0,972 | 1,460       | 0,088 | 1,292    | 1,625  | 1,308     | 1,632  |
| 3-4    | 60   | 1,335         | 0,134 | 1,081    | 1,615 | 1,062     | 1,587 | 2,462       | 0,106 | 2,260    | 2,667  | 2,260     | 2,666  |
| 4      | 61   | 0,763         | 0,162 | 0,454    | 1,078 | 0,458     | 1,081 | 0,619       | 0,348 | -0,257   | 1,097  | -0,025    | 1,159  |
| 4      | 62   | 1,495         | 0,196 | 1,129    | 1,894 | 1,117     | 1,878 | 2,387       | 0,111 | 2,183    | 2,606  | 2,183     | 2,607  |
| 4      | 63   | 1,058         | 0,165 | 0,765    | 1,416 | 0,725     | 1,372 | 1,576       | 0,118 | 1,333    | 1,802  | 1,334     | 1,802  |
| 4      | 64   | 1,137         | 0,182 | 0,796    | 1,518 | 0,779     | 1,491 | 1,311       | 0,137 | 1,015    | 1,547  | 1,048     | 1,577  |
| 4      | 65   | 1,452         | 0,197 | 1,093    | 1,847 | 1,089     | 1,838 | 1,618       | 0,091 | 1,433    | 1,791  | 1,471     | 1,818  |
| 4      | 66   | 1,788         | 0,223 | 1,361    | 2,256 | 1,332     | 2,216 | 1,935       | 0,094 | 1,761    | 2,123  | 1,757     | 2,118  |
| 4      | 67   | 1,891         | 0,238 | 1,472    | 2,381 | 1,445     | 2,352 | 1,879       | 0,091 | 1,712    | 2,052  | 1,706     | 2,042  |
| 4      | 68   | 0,773         | 0,130 | 0,525    | 1,028 | 0,543     | 1,044 | 2,672       | 0,174 | 2,365    | 3,053  | 2,339     | 3,001  |
| 4      | 69   | 1,697         | 0,248 | 1,235    | 2,212 | 1,202     | 2,182 | 2,914       | 0,147 | 2,644    | 3,220  | 2,637     | 3,205  |
| 4      | 70   | 1,645         | 0,224 | 1,220    | 2,083 | 1,194     | 2,052 | 2,708       | 0,130 | 2,473    | 2,969  | 2,466     | 2,958  |
| 4      | 71   | 2,102         | 0,256 | 1,616    | 2,631 | 1,582     | 2,583 | 2,012       | 0,089 | 1,843    | 2,189  | 1,858     | 2,202  |
| 4      | 72   | 1,943         | 0,236 | 1,504    | 2,441 | 1,490     | 2,402 | 1,918       | 0,090 | 1,745    | 2,102  | 1,745     | 2,102  |
| 4      | 73   | 1,603         | 0,200 | 1,246    | 2,010 | 1,266     | 2,027 | 2,241       | 0,102 | 2,046    | 2,462  | 2,038     | 2,437  |
| 4      | 74   | 1,091         | 0,167 | 0,784    | 1,432 | 0,806     | 1,444 | 1,540       | 0,122 | 1,302    | 1,765  | 1,307     | 1,772  |
| 4      | 75   | 0,738         | 0,191 | 0,392    | 1,140 | 0,364     | 1,094 | 0,113       | 0,540 | -1,214   | 0,808  | -0,970    | 0,861  |
| 4      | 76   | 1,730         | 0,213 | 1,346    | 2,175 | 1,346     | 2,180 | 2,362       | 0,106 | 2,173    | 2,573  | 2,163     | 2,558  |
| 4      | 77   | 1,384         | 0,194 | 1,015    | 1,786 | 0,996     | 1,756 | 2,846       | 0,142 | 2,592    | 3,133  | 2,592     | 3,135  |
| 4      | 78   | 1,332         | 0,187 | 0,984    | 1,723 | 0,959     | 1,686 | 2,783       | 0,139 | 2,524    | 3,060  | 2,516     | 3,051  |
| 4      | 79   | 1,287         | 0,177 | 0,974    | 1,639 | 0,974     | 1,639 | 1,712       | 0,102 | 1,501    | 1,919  | 1,516     | 1,926  |
| 4      | 80   | 1,320         | 0,176 | 0,985    | 1,677 | 0,964     | 1,646 | 2,506       | 0,122 | 2,282    | 2,761  | 2,269     | 2,742  |

# Referências Bibliográficas

- [1] Albert, J. (1992), *Bayesian estimation of normal ogive item response curves using Gibbs sampling*, Journal of educational statistics. 17, pp. 251–269.
- [2] Alexandre, J.W.C., Andrade,D.F., Vasconcelos,A.P. e Araújo, A.M.S.(2002). *Uma proposta de análise de um construto para a medição dos fatores críticos da gestão pela qualidade através da teoria da resposta ao item*. Gestão & Produção, v.9, n.2, p.129-141.
- [3] Andrade, D. F., Tavares, H. R. and Valle, R. C. (2000),*Teoria de Resposta ao Item: Conceitos e Aplicações*, 14<sup>o</sup> SINAPE, ABE.
- [4] Andrade, D. F. (1999). *Comparando o Desempenho de Grupos (Populações) de Respondentes Através da Teoria da Resposta ao Item*. Tese apresentada ao Departamento de Estatística e Matemática Aplicada da UFC para o concurso de professor titular.
- [5] Arellano-Valle, R.B. and Azzalini, A. (2008), *The centred parametrization for the multivariate skew-normal distribution*, Journal of Multivariate Analysis, 99, 7 , 1362-1382.
- [6] Azevedo, C. L. N. and Andrade, D. F. (2010), *An estimation method for latent traits and population parameter for the nominal response model*, Brazilian journal of Probability and Statistics, 24, 415-433.
- [7] Azevedo, C. L. N., Bolfarine, H. and Andrade, D. F. (2011a), *Bayesian inference for a skew-normal IRT model under the centred parameterization*, Computational Statistical & Data analysis, 55, 1, 353-365.
- [8] Azevedo, C. L. N., Bolfarine, H. and Andrade, D. F. (2011b), *Parameter recovery for a skew normal IRT model under a Bayesian approach: hierarchical framework, prior and kernel sensitivity and sample size*, Jornal of Statistical Computation & Simulation 1-23.
- [9] Azevedo, C.L.N., Andrade, D.F., Fox, J.-P., (2009a). *Based Gibbs Sampling estimation procedure and model-fit for multiple groups IRT model*. Relatório técnico.
- [10] Azevedo, C.L.N., Fox, J.-P., Andrade, D. F., (2009b). *Gibbs sampling for Bayesian estimation and model fit assessment for a longitudinal IRT model*. Relatório técnico.
- [11] Azzalini, A. (1985), *A class of distribution which includes the normal ones*, Scandinavian journal of statistics. 12, pp. 171-178.

- [12] Azzalini, A. and Capitanio, A. (1999), *Statistical applications of the multivariate skew normal distribution*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 61: 579–602. doi: 10.1111/1467 – 9868.00194
- [13] Baker, F. B. and Kim, Seock-Ho (2004), *Item response theory: parameter estimation techniques*, second edition, Marcel Dekker, Inca, New York.
- [14] Bayes, C. L. and Branco, M. D. (2007), *Bayesian inference for the skewness parameter of the scalar skew-normal distribution*, Brazilian journal of probability and statistics, 21, 141-163.
- [15] Bazán, J. L., Branco, M. D. and Bolfarine, H. (2006), *A Skew Item Response Model*, Bayesian analysis, pp. 861-892.
- [16] Béguin, A. A. B. and Glas, C. A. W. (2001), *MCMC estimation and some model-fit analysis of multidimensional IRT models*, Psychometrika, 66, 4, 541-561.
- [17] Bock, D. R. and Zimowski, M. F. (1997), *The Multiple Groups IRT*, in *Handbook of modern Item Response Theory*, Wim J. van der Linden and Ronald K. Hambleton eds, Springer-Verlag
- [18] Bock, R. D. and Aitkin, M. (1981), *Marginal maximum likelihood estimation of item parameters: An application of an EM algorithm*, Psychometrika. 46, pp. 433–459.
- [19] Branco, M. D' E., Arellano-Valle, R. *Distribuições elípticas assimétricas*. São Paulo: Associação Brasileira de Estatística, 16<sup>o</sup> Sinape 26 a 30.7.2004.
- [20] Chen, M.-H., Shao, Q.-M. and Ibrahim, J. G. (2000), *Monte carlo methods for Bayesian computation, first edition*, Springer-Verlag, New York.
- [21] Chiogna, M. (2005), *A note on the asymptotic distribution of the maximum likelihood estimator for the scalar skew-normal distribution*, Statistical Methods & Applications, 14, 331-341.
- [22] da-Silva, C. Q. and Gomes, A. E. (2011), *Bayesian inference for an item response model for modeling test anxiety*, Computational Statistica & Data Analysis, 55, pp. 3165–23182.
- [23] De Ayala, R. J and Sava - Bolesta, M. (1999), *Item Parameter Recovery for the Nominal Response Model*, Applied psychological measurement, 23, 1, 3-19.
- [24] DeMars, C. E., (2003). *Sample size and the recovery of nominal response model*. Applied Psychological Measurement 23 (1), 3-19.
- [25] Fox, J. -P. and Glas, C. A. W. (2001), *Bayesian Estimation of a Multilevel IRT Model using Gibbs Sampling*, Psychometrika. 66, pp. 269–286.
- [26] Fox, J.-P. (2001), *Multilevel IRT : A bayesian perspective on estimating parameters and testing statistical hypothesis*, Tese de doutorado, OMD-University of Twente.

- [27] Fox, J.-P. (2010), *Bayesian Item Response Modeling: Theory and Applications*, first edition, Springer-Verlag, New York.
- [28] Fox, J.-P. (2004), *Multilevel IRT assessment*, in New Developments in Categorical Data Analysis for the Social and Behavioral Sciences, van der Ark, Croon and Sijtsma eds, Lawrence Erlbaum Associates, Inc, London.
- [29] Fox, J.-P. and Glas, C. A. W. (2005), *Bayesian modification indices for IRT models*, Statistica neerlandica, 59, 95-106.
- [30] Gamerman, D. and Lopes, H. F. (2006), Markov Chain Monte Carlo : Stochastic simulation for bayesian inference, second edition, *Chapman & Hall, London*.
- [31] Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2004), *Bayesian Data Analysis, Second Edition*, Chapman & Hall/CRC, London.
- [32] Genton, M. (2004), *Skew-elliptical distributions and their applications*, Chapman & Hall, London.
- [33] Geweke, J. (1991), *Efficient simulation from the multivariate normal and Student-t distributions subject to linear constraints and the evaluation of constraint probabilities*, relatório técnico.
- [34] Gifford, J. A. and Swaminathan, H. (1990), *Bias and the effect of priors in bayesian estimation of item response models*, Applied psychological measurement, 14, 33, 33-43.
- [35] Gonzalez, R. L. (2004), *Data augmentation in the Bayesian multivariate probit model*, Sheffield Economic Research Paper Series.
- [36] Ghosh, M., Ghosh, A., Chen, Ming-Hui and Agresti, A. (2000), *Noninformative priors for one-parameter item response models*, Journal of Statistical Planning and Inference. 88, pp. 269–286.
- [37] Glas, C. A. W. (1998), *Detection of differential item functioning using Lagrange multiplier tests*, Statistica Sinica, 8, 647-667.
- [38] Harwell, M., Stone, C. A., Hsu, Tse-Chi & Kirisci, L. (1996), *Monte Carlo Studies in Item Response Theory*, Applied psychological measurement, 20, 2, 101–125.
- [39] Harwell, M. R. and Janosky, J. E. (1991), *An empirical study of the effects of small datasets and varying prior variances on item parameter estimation in Bilog*, Applied psychological measurement, 15, 3, 279-291.
- [40] Henze, N. (1986), *A probabilistic representation of the skew-normal distribution*, Scandinavian journal of statistics. 13, pp. 271–275.
- [41] Horn, R. A. and Johnson, C. R. (1991), *Topics in Matrix Analysis*, Cambridge university press, United Kingdom, first edition.

- [42] Kim, Seock-Ho, Cohen, A. S. and Park, Tae-Hak (2005), *Detection of Differential Item Functioning in Multiple Groups*, Journal of Educational Measurement, 32, 3, 261-276.
- [43] Kirisci, L., Hsu, Tse-Chi, Yu, L., (2001). *Robustness of item parameter estimation programs to assumptions of unidimensionality and normality*. Applied Psychological Measurement 25 (2), 146-162.
- [44] Kolen, M. J. and Brennan, R. L. (2004), *Test Equating - Methods and Practices, second edition*, Springer-Verlag, New York.
- [45] Liseo, B. and Loperfido, N. (2006) *A note on reference priors for the scalar skew-normal distribution*, Journal of statistical and planning inference, 136, 373-389.
- [46] Little, R. J. and Rubin, D. B. (2002), *Statistical analysis with missing data, second edition*, John Wiley & Sons, New Jersey.
- [47] Lord, F. M., (1980). *Applications of Item Response Theory to Practical Testing Problems, first edition*. Lawrence Erlbaum Associates, Hillsdale.
- [48] McCullagh, P. and Nelder, J. A. (1989), *Generalized Linear Models, second edition, Chapman & Hall, London*
- [49] Mislevy, R. J. (1984), *Estimating latent distributions*, Psychometrika. 49, pp. 359-381.
- [50] Mislevy, R. J. (1986), *Bayes modal estimation in item response models*, Psychometrika. 51, pp. 177-195.
- [51] Montgomery, D. C. (2004), *Design and Analysis of Experiments, sixth edition*, Chapman & Hall, London.
- [52] Múthen, B. and Lehman, J. (1985), *Multiple Group IRT Modeling: Applications to Item Bias Analysis*, Journal of Educational and Behavioral Statistics, 10, 2, 365-376.
- [53] Patz, R. J. and Junker, B. W. (1999a), *A Straightforward Approach to Markov Chain Monte Carlo Methods for Item Response Models*, Journal of educational and behavioral Statistics. 24, 2, pp. 146-178.
- [54] Patz, R. J. and Junker, B. W. (1999b), *The Applications and Extensions of MCMC in IRT: Multiple Item Types, Missing Data, and Rated Responses*, Journal of educational and behavioral Statistics. 24, 4, pp. 342-366.
- [55] Penfield, R. D. (2001), *Assessing Differential Item Functioning Among Multiple Groups: A Comparison of Three Mantel-Haenszel Procedures*, 14, 3, 235-259.
- [56] Pewsey, A. (2000) *Problems of inference for Azzalini's skew-normal distribution*, Journal of applied statistics. 27, pp. 859-870.

- [57] Poli, E. C. (2007), *Longitudinal study in Mathematics: possibilities and reading of a reality of elementary school*. Doctoral Thesis in Education. State University of Campinas. (In Portuguese)
- [58] Sahu, S. K., Dey, D. K. and Branco, M. D. (2003), *A new class of multivariate skew distributions with applications to Bayesian regression models*, The Canadian Journal of Statistics. 31, pp. 129–150.
- [59] Samejima, F., (1997). *Departure from normal assumptions: a promise for future psychometrics with substantive mathematical modeling*. Psychometrika 62 (4), 471-493.
- [60] Seong, Tae-Je, (1990). *Sensitivity of marginal maximum likelihood estimation of item and ability parameters to the characteristics of the prior ability distributions*. Applied Psychological Measurement 14 (3), 229-311.
- [61] Stern, H. S. and Sinharay, S. (2005), Bayesian model checking and Model Diagnostics, *in Handbook of Statistics, 25, Bayesian Modelling, Thinking and Computation*, Dipak K. Dey and C.R. Rao eds, Elsevier, The Netherlands.
- [62] Sinharay, S. (2004), *Experiences With Markov Chain Monte Carlo Convergence Assessment in Two Psychometric Examples*, Journal of educational and behavioral statistics. 29, 4, pp. 461–488.
- [63] Sinharay, S. (2006), A bayesian item fit analysis for unidimensional item response theory models, *British journal of mathematical and statistical psychology*, 59, 429-449.
- [64] Sinharay, S., Johnson. M. S. and Stern, H. (2006), Posterior Predictive Assessment of Item Response Theory Models, *Applied Psychological Measurement*, 30, 4, 298-321.
- [65] Soares, T. M., Gonçalves, F. B. and Gamerman, D. (2009), *An integrated Bayesian Model for DIF analysis*, Journal of Educational and Behaviourial statistics, 34,3, 348–377.
- [66] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linden, A. (2002), *Bayesian measures of model complexity and fit*, Journal Royal Statistical Society, 64, 3 583 – 639
- [67] Swaminathan, H., Hambleton, R. K. , Sireci, S. G. , Xing, D. and Rizavi, S. M. (2003), *Small ample size estimation in dichotomous item respjnse models: effect of prioris based on judgmental information on the accuracy of item parameter estimates*, Applied psychological measurement, 27, 1, 27-51.
- [68] Tanner, M. A. and Wong, W. H. (1987), *The calculation of posterior distribution by data augmentation*, Journal of the american statistical association. 82, 398, pp. 528–540.
- [69] Tavares, H. R. and Andrade, D. F. (2006), *Latent distribution parameter estimation and approximation*, manuscrito.
- [70] Tavares, H. R. (2001). *Modelos da Teoria da Resposta ao Item para Dados Longitudinais*. Tese de Doutorado. IME/USP.

- [71] Thissen, D., Chen, Wen-Hung & Bock, D. (2003), BILOG-MG 3, *In IRT from SSI*, Du Toiti eds, Scientific Software International.
- [72] Wollack, J. A., Bolt, D. M., Cohen, A. S., and Lee, Young-Sun. (2002), *Recovery of Item Parameters in the Nominal Response Model: A Comparison of Marginal Maximum Likelihood Estimation and Markov Chain Monte Carlo Estimation*, Applied psychological measurement, 26, 3, 339-352.
- [73] Woods, C. M. (2007), *Empirical Histograms in Item Response Theory With Ordinal Data*, Educational and Psychological Measurement. 67, pp. 73-87.
- [74] Woods, C. M. (2009), *Item Response Theory With Estimation of the Latent Density Using Davidian Curves*, Applied Psychological Measurement. 12, pp. 102-117.
- [75] Woods, C. M. (2006), *Ramsay-Curve Item Response Theory (RC-IRT) to Detect and Correct for Nonnormal Latent Variables*, Educational and Psychological Measurement. 11, pp. 253-270.
- [76] Woods, C. M. (2008), *Ramsay-Curve Item Response Theory for the Three-Parameter Logistic Item Response Model*, Applied Psychological Measurement. 32, pp. 447-465.
- [77] Zimowski, M. F., Muraki, E., Mislevy, R. J. and Bock, R. D. (1996), *Bilog-MG : Multiple-group IRT analysis and test maintenance for binary items*, SSI : Scientific Software, Chicago.