

UNIVERSIDADE FEDERAL DO RIO GRANDE
Instituto de Matemática, Estatística e Física
Programa de Pós-Graduação em Modelagem Computacional

Débora Spenassato

**TEORIA DA RESPOSTA AO ITEM: INFERÊNCIA BAYESIANA EM MODELOS
SIMÉTRICO E ASSIMÉTRICO**

Rio Grande/RS

2011

Débora Spenassato

**TEORIA DA RESPOSTA AO ITEM: INFERÊNCIA BAYESIANA EM MODELOS
SIMÉTRICO E ASSIMÉTRICO**

Dissertação apresentada à Banca Examinadora do Programa de Pós-Graduação em Modelagem Computacional da Universidade Federal do Rio Grande, como exigência parcial para obtenção do título de Mestre em Modelagem Computacional.

Orientador: Paul Gehrard Kinas, PhD.

Rio Grande/RS

2011

Agradecimentos

Agradeço primeiramente a Deus pelas conquistas até aqui alcançadas e pela saúde para ir em busca de meus sonhos.

Ao professor e orientador Paul G. Kinas, pelo acompanhamento, disponibilidade e pelos conhecimentos transmitidos durante esse período.

A CAPES (Coordenação de Aperfeiçoamento de Nível Superior) pelo auxílio concedido.

Aos meus pais, pelo apoio e incentivo.

Aos meus tios e primos de Rio Grande que me acolheram nessa cidade.

Às amigas Caroline e Carla, que contribuíram indiretamente com meu trabalho. Ao amigo Adriano pelo companheirismo.

A todos os colegas, pela amizade que fizemos nessa caminhada e a todos aqueles que, direta ou indiretamente, colaboraram para a concretização desse trabalho.

Resumo

No presente estudo discutimos a modelagem de dados dicotômicos através da inferência bayesiana, utilizando o método de Monte Carlo via Cadeias de Markov (MCMC) para modelos da Teoria da Resposta ao Item (TRI) de três parâmetros. Todos os procedimentos de inferência foram implementado no OpenBUGS por intermédio do software R. Consideramos um teste composto por 15 itens, uma amostra de 500 indivíduos e produzimos vinte conjuntos de dados sintéticos para cada um dos seis cenários distintos obtidos a partir de diferentes combinações da distribuição normal ou normal assimétrica das habilidades, com funções de ligação logística ou logística assimétrica. Nosso objetivo foi avaliar a qualidade e eficiência do procedimento de estimação com funções de ligação logística e logística assimétrica e, distribuição normal das habilidades. Concluimos que, havendo algum tipo de assimetria, seja nas habilidades ou na CCI, as estimativas em geral foram melhores utilizando o ajuste assimétrico. O maior erro nas estimativas ocorre quando os dados sintéticos são obtidos a partir de uma forte distribuição normal assimétrica negativa para as habilidades associada a uma função de ligação assimétrica. O melhor ajuste acontece quando a distribuição das habilidades e a função de ligação são simétricas.

Palavras-chave: TRI, INFERÊNCIA BAYESIANA, MCMC, MODELOS LOGÍSTICO SIMÉTRICO E ASSIMÉTRICO, DISTRIBUIÇÃO NORMAL ASSIMÉTRICA.

Abstract

In this study we discussed the modeling of dichotomous data by Bayesian inference, using Markov Chain Monte Carlo (MCMC) method to Item Response Theory (IRT) models of three parameters. All inference procedures were implemented in OpenBUGS through software R. We considered a test consisting of 15 items, a sample of 500 individuals and produced twenty sets of synthetic data for each one of the six distinct scenarios obtained from different combinations of normal or skew-normal distribution of abilities, with logistic or asymmetric-logistic link functions. Our objective was to evaluate the quality and efficiency of the estimation procedure with logistic and asymmetric-logistic link functions and a normal distribution for abilities. We concluded that, having some kind of asymmetry, in the ability or ICC, the estimates in general improved with the asymmetric fit. The largest error in the estimates occurs when synthetic data are obtained from a strong negative skew-normal distribution for the abilities associated to a asymmetric link function. The best fit happens when the distribution of abilities and the link function are symmetric.

Key-words: IRT, BAYESIAN INFERENCE, MCMC, LOGISTIC SYMMETRIC AND ASYMMETRIC MODELS, SKEW-NORMAL DISTRIBUTION.

LISTA DE FIGURAS

Figura 1 – Exemplo de uma CCI de três parâmetros identificando o significado de cada parâmetro.	6
Figura 2 – Gráfico das distribuições utilizadas em nosso estudo para a habilidade de indivíduos (distribuição normal e distribuição normal assimétrica).....	8
Figura 3 – Distribuição acumulada das distribuições da figura acima.	8
Figura 4 – CCI considerando fixos os valores dos parâmetros $a = 2$, $b = 1$, $c = 0,25$, $\theta \sim N(0,1)$ e variando λ . Se $\lambda = 1,8$ e $\lambda = 0,3$ (modelo assimétrico), se $\lambda = 1$ (modelo simétrico)...	11
Figura 5 - Composição dos diferentes modelos da TRI para itens dicotômicos.	11
Figura 6 - Histograma das habilidades verdadeiras utilizadas no estudo.	26
Figura 7 – Composição dos seis conjuntos de dados sintéticos gerados para análise.	27
Figura 8 – Traço das cadeias para dois parâmetros, indicando convergência.	28
Figura 9 – Diagnóstico Gelman e Rubin (<i>Rhat</i>) para alguns parâmetros do ajuste assimétrico para <i>Dados 1</i> , indicando convergência.	29
Figura 10 – Diagnóstico de Geweke para testar estacionariedade. Valores longe do intervalo de -2 a 2 indicam não estacionariedade.	29
Figura 11 – Medidas de dependência entre as amostras (autocorrelação) de cadeias de Markov. Os gráficos indicam autocorrelação muito fraca.	30
Figura 12 – Correlação entre as discriminações dos 15 itens. Como podemos observar a correlação é praticamente nula entre os parâmetros de discriminação.....	30
Figura 13 – Esquema ilustrando como foram calculadas as estatísticas das distribuições posteriores de cada parâmetro dos itens (ETAPA 2) e os sumários estatísticos para as 20 réplicas (ETAPA 3).	31
Figura 14 – Esquema das etapas para <i>Dados 1</i> , da mesma forma como é válido para os outros dados.	31
Figura 15 – Ordem de classificação dos dados em relação ao melhor ajuste para as habilidades (θ_j).	35

LISTA DE TABELAS

Tabela 1 – Valores verdadeiros dos parâmetros dos itens utilizados para gerar a matriz resposta ao teste dos n indivíduos	25
Tabela 2 – Prioris adotadas para o ajuste dos parâmetros de interesse	28
Tabela 3 – Médias sobre as 20 réplicas dos \sqrt{EQM} e \overline{VM} das estimativas das habilidades obtidas com ajuste dos modelos com CCI simétrica e assimétrica para os seis tipos de dados	34
Tabela 4 – Número médio de subestimativas (-), estimativas adequadas (0) e sobrestimativas (+) para ajuste simétrico e assimétrico referente aos seis conjuntos de dados para a média posterior.....	36
Tabela 5 – Número médio de subestimativas (-), estimativas adequadas (0) e sobrestimativas (+) para ajuste simétrico e assimétrico referente aos seis conjuntos de dados para a mediana posterior.....	36
Tabela 6 – Proporção média de cobertura dos ICr 95%	38
Tabela 7 – Amplitude média dos ICr 95%	38
Tabela 8 – Média dos valores DIC e pD utilizados na seleção do modelo mais adequado	39
Tabela 9 – Tempo médio de execução (em minutos) das 20 réplicas para ajuste simétrico e assimétrico.....	40
Tabela A. 1 – Ajuste utilizando modelo simétrico para a_i em <i>Dados 1</i>	47
Tabela A. 2 – Ajuste utilizando modelo assimétrico para a_i em <i>Dados 1</i>	48
Tabela A. 3 – Ajuste utilizando modelo simétrico para b_i em <i>Dados 1</i>	49
Tabela A. 4 – Ajuste utilizando modelo assimétrico para b_i em <i>Dados 1</i>	49
Tabela A. 5 – Ajuste utilizando modelo simétrico para c_i em <i>Dados 1</i>	50
Tabela A. 6 – Ajuste utilizando modelo assimétrico para c_i em <i>Dados 1</i>	51
Tabela A. 7 – Estimativas de λ_i pelo modelo assimétrico para <i>Dados 1</i>	52
Tabela A. 8 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos, utilizando o modelo simétrico para o ajuste de <i>Dados 1</i>	53
Tabela A. 9 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando o modelo assimétrico para o ajuste de <i>Dados 1</i>	53
Tabela A. 10 - DIC's e pD's para <i>Dados 1</i>	53
Tabela A. 11 – Ajuste utilizando modelo simétrico para a_i em <i>Dados 2</i>	54
Tabela A. 12 – Ajuste utilizando modelo assimétrico para a_i em <i>Dados 2</i>	55

Tabela A. 13 – Ajuste utilizando modelo simétrico para b_i em <i>Dados 2</i>	55
Tabela A. 14 – Ajuste utilizando modelo assimétrico para b_i em <i>Dados 2</i>	56
Tabela A. 15 – Ajuste utilizando modelo simétrico para c_i em <i>Dados 2</i>	57
Tabela A. 16 – Ajuste utilizando modelo assimétrico para c_i em <i>Dados 2</i>	58
Tabela A. 17 – Estimativas de λ_i pelo modelo assimétrico para <i>Dados 2</i>	58
Tabela A. 18 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos, utilizando modelo simétrico para ajuste de <i>Dados 2</i>	59
Tabela A. 19 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos, utilizando modelo assimétrico para ajuste de <i>Dados 2</i>	59
Tabela A. 20 - DIC's e pD's para <i>Dados 2</i>	60
Tabela A. 21 – Ajuste utilizando modelo simétrico para a_i em <i>Dados 3</i>	60
Tabela A. 22 – Ajuste utilizando modelo assimétrico para a_i em <i>Dados 3</i>	61
Tabela A. 23 – Ajuste utilizando modelo simétrico para b_i em <i>Dados 3</i>	62
Tabela A. 24 – Ajuste utilizando modelo assimétrico para b_i em <i>Dados 3</i>	63
Tabela A. 25 – Ajuste utilizando modelo simétrico para c_i em <i>Dados 3</i>	63
Tabela A. 26 – Ajuste utilizando modelo assimétrico para c_i em <i>Dados 3</i>	64
Tabela A. 27 – Estimativas de λ_i pelo modelo assimétrico para <i>Dados 3</i>	65
Tabela A. 28 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para <i>Dados 3</i>	66
Tabela A. 29 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para <i>Dados 3</i>	66
Tabela A. 30 – DIC's e pD's para <i>Dados 3</i>	66
Tabela A. 31 – Ajuste utilizando modelo simétrico para a_i em <i>Dados 4</i>	67
Tabela A. 32 – Ajuste utilizando modelo assimétrico para a_i em <i>Dados 4</i>	67
Tabela A. 33 – Ajuste utilizando modelo simétrico para b_i em <i>Dados 4</i>	68
Tabela A. 34 – Ajuste utilizando modelo assimétrico para b_i em <i>Dados 4</i>	69
Tabela A. 35 – Ajuste utilizando modelo simétrico para c_i em <i>Dados 4</i>	70
Tabela A. 36 – Ajuste utilizando modelo assimétrico para c_i em <i>Dados 4</i>	71
Tabela A. 37 – Estimativas de λ_i pelo modelo assimétrico para <i>Dados 4</i>	71
Tabela A. 38 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para <i>Dados 4</i>	72
Tabela A. 39 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para <i>Dados 4</i>	72
Tabela A. 40 – DIC's e pD's para <i>Dados 4</i>	73

Tabela A. 41 – Ajuste utilizando modelo simétrico para a_i em <i>Dados 5</i>	73
Tabela A. 42 – Ajuste utilizando modelo assimétrico para a_i em <i>Dados 5</i>	74
Tabela A. 43 – Ajuste utilizando modelo simétrico para b_i em <i>Dados 5</i>	75
Tabela A. 44 – Ajuste utilizando modelo assimétrico para b_i em <i>Dados 5</i>	76
Tabela A. 45 – Ajuste utilizando modelo simétrico para c_i em <i>Dados 5</i>	76
Tabela A. 46 – Ajuste utilizando modelo simétrico para c_i em <i>Dados 5</i>	77
Tabela A. 47 – Estimativas de λ_i pelo modelo assimétrico para <i>Dados 5</i>	78
Tabela A. 48 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para <i>Dados 5</i>	79
Tabela A. 49 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para <i>Dados 5</i>	79
Tabela A. 50 – DIC's e pD's para <i>Dados 5</i>	79
Tabela A. 51 – Ajuste utilizando modelo simétrico para a_i em <i>Dados 6</i>	80
Tabela A. 52 – Ajuste utilizando modelo assimétrico para a_i em <i>Dados 6</i>	81
Tabela A. 53 – Ajuste utilizando modelo simétrico para b_i em <i>Dados 6</i>	81
Tabela A. 54 – Ajuste utilizando modelo assimétrico para b_i em <i>Dados 6</i>	82
Tabela A. 55 – Ajuste utilizando modelo simétrico para c_i em <i>Dados 6</i>	83
Tabela A. 56 – Ajuste utilizando modelo assimétrico para c_i em <i>Dados 6</i>	84
Tabela A. 57 – Estimativas de λ_i para modelo assimétrico para <i>Dados 6</i>	85
Tabela A. 58 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para <i>Dados 6</i>	85
Tabela A. 59 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para <i>Dados 6</i>	86
Tabela A. 60 – DIC's e pD's para <i>Dados 6</i>	86

LISTA DE SIGLAS

MCMC – Monte Carlo via Cadeias de Markov

TRI – Teoria de Resposta ao Item

TCM – Teoria Clássica de Medidas

CCI – Curva Característica do Item

ENEM – Exame Nacional de Ensino Médio

SAEB – Sistema de Avaliação da Educação Básica

SARESP – Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo

f.d. – Função Distribuição

f.d.p – Função Densidade de Probabilidade

LPE – Expoente Logístico Positivo

DIC – *Deviance Information Criterion* (Critério de Informação da Deviância)

EAIC – Critério de Informação de Akaike Esperado

EBIC – Critério de Informação (Schwarz) Bayesiana Esperado

SQRL – Soma dos quadrados dos resíduos latentes posterior

\overline{EQM} – Média do erro quadrático médio

$\sqrt{\overline{EQM}}$ – Raiz quadrada da média do erro quadrático médio

\overline{VM} – Média do viés médio

VP – Viés Padronizado

ICr 95% – Intervalo de credibilidade de 95%

Sumário

1. INTRODUÇÃO	1
1.1. Modelos TRI dicotômicos	2
1.1.1. Modelos TRI simétricos de três parâmetros	4
1.2. Distribuição Normal Assimétrica	7
1.2.1. Modelos TRI Assimétricos	9
1.3. Justificativa	12
1.4. Objetivos	13
1.5. Estrutura do trabalho	13
2. MÉTODO DE ESTIMAÇÃO	14
2.1. Inferência Bayesiana	14
2.2. Método de Monte Carlo via Cadeias de Markov (MCMC)	15
2.3. Inferência na TRI e escolha de prioris	18
2.4. Técnicas de diagnóstico de convergência das cadeias	20
2.5. Critérios para seleção de modelos	22
3. METODOLOGIA	24
3.1. Geração dos dados sintéticos	24
3.2. Estimação dos parâmetros do modelo	27
3.3. Análise da qualidade das estimações	31
4. RESULTADOS	34
4.1. Estimativas das habilidades θ_j	34
4.2. Estimativas dos parâmetros dos itens	35
4.3. Escolha do melhor modelo	39
4.4. Tempo de simulação	39
5. CONCLUSÕES	41
REFERÊNCIAS BIBLIOGRÁFICAS	43
APÊNDICE	47

1. INTRODUÇÃO

Entre as metodologias utilizadas para quantificar o resultado de um teste, destacam-se a Teoria Clássica de Medidas (TCM) e Teoria da Resposta ao Item (TRI).

Pela teoria clássica de medidas (TCM) as análises dos itens do teste são feitas com a soma dos escores, referentes aqueles com acerto, do conjunto de itens (BAKER, 2001; RABELO, 2009). Sendo assim, a TCM, que ainda é a metodologia mais utilizada em avaliações de conhecimento, não permite a comparabilidade de indivíduos que não foram submetidos ao mesmo teste. Na TRI, o importante não é conhecer apenas quantas questões foram respondidas corretamente, mas sim quais delas. Indivíduos que acertaram o mesmo número de questões poderão ter escores diferentes se o conjunto de itens corretos não forem os mesmos. Essas são diferenças fundamentais entre essas metodologias utilizadas para a análise de testes.

Estudos sobre a Teoria da Resposta ao Item surgiram nos anos 1950 com o objetivo de fornecer informações mais ricas para a análise de testes de inteligência (JUSTINO e ANDRADE, 2007), cujos resultados podem variar em razão dos instrumentos de medida utilizados. Mas, somente após o avanço tecnológico dos anos de 1980, com o desenvolvimento de softwares para algoritmos complexos, que essa teoria começou a ser difundida (RABELO, 2009).

A TRI utiliza-se de modelos estatísticos para fazer inferências sobre o nível de habilidade de um indivíduo com base no conjunto de itens respondidos corretamente no teste. Para isso, admite-se a existência de uma ou mais habilidades que não podem ser diretamente observadas. Porém, é possível estimá-las a partir das respostas desse indivíduo a um conjunto de perguntas especialmente formuladas para esse fim. Uma vez obtidos os parâmetros dos itens e da habilidade, é possível calcular a probabilidade do indivíduo acertar a cada item, dado seu padrão de respostas e, assim, diagnosticar possíveis acertos ao acaso. Na literatura, as habilidades também são chamadas de traços latentes, proficiências e variável latente. Nesse estudo optamos por usar somente o termo “habilidade”.

A metodologia da TRI permite extrair informações de cada item que compõem o teste, como a dificuldade, a discriminação e a probabilidade de um indivíduo com baixa habilidade acertar ao item por simples palpite (acerto ao acaso), além de que, as estatísticas obtidas são independentes do grupo de indivíduos que foram utilizados para obtê-las se alguns pressupostos como, todos os parâmetros dos itens estarem em uma mesma escala e os itens do teste estarem medindo apenas uma habilidade (*unidimensionalidade do teste*) são satisfeitos

(BAKER, 2001). Devido à invariância dos parâmetros, esses itens poderão ser aplicados novamente a outro grupo. Além disso, é possível comparar o desempenho de indivíduos que se submetem a testes diferentes ou parcialmente diferentes, elaborar e interpretar escalas de habilidade garantindo melhor análise da habilidade dos indivíduos.

Essa nova metodologia de avaliação de testes ainda é recente no Brasil. Atualmente, vem sendo aplicada em avaliações da qualidade educacional a nível nacional, como nas provas do ENEM (Exame Nacional de Ensino Médio), no SAEB (Sistema de Avaliação da Educação Básica) e SARESP (Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo). Além da sua utilidade para a análise do desempenho escolar, a TRI também tem aplicabilidade em outras áreas do conhecimento.

Alexandre et al. (2002) apresenta uma proposta do uso de modelos da TRI elaborados para medir a Gestão pela Qualidade Total (GQT), na análise de construtos. Em Tezza (2009) são apresentados vantagens e oportunidades da TRI no contexto de engenharia de produção, mostrando sua aplicação no mercado de *e-business*, criando uma escala padronizada para medir o grau de usabilidade de sites de *e-commerce*, concluindo que a TRI pode ser uma valiosa ferramenta de gestão, capaz de gerar informações quantitativas e precisas. Guewehr (2007) aplica a TRI aos resultados obtidos no uso de três instrumentos de avaliação de qualidade de vida e compara os resultados desse com os provenientes dos métodos tradicionais de análises (TCM e Análise Fatorial).

Ainda há muito a ser explorado na metodologia da TRI, principalmente em relação ao desenvolvimento de ferramentas computacionais mais eficazes que possam auxiliar na análise de testes em larga escala. Cabe ressaltar que os processos de estimação utilizados para inferir sobre os parâmetros do modelo são responsáveis pela qualidade das estimativas.

1.1. Modelos TRI dicotômicos

No contexto da TRI, o modelo a ser escolhido dependerá fundamentalmente de três fatores: do número de populações envolvidas, do número de habilidades que estão sendo medidas e da resposta ser dicotômica ou não (ANDRADE et al., 2000). Outra questão importante que deve ser observada é quais parâmetros dos itens há interesse em investigar.

Quanto às respostas, os modelos podem ser classificados como dicotômicos e não dicotômicos. Segundo Andrade et al. (2000), trata-se de modelos tanto para a análise de itens abertos (de resposta livre) quanto para a análise de itens de múltipla escolha que podem ser

avaliados como: certo ou errado (*dicotômicos*); itens que são elaborados ou corrigidos de modo a ter-se uma ou mais categorias intermediárias **ordenadas** entre as categorias certo e errado (modelo *gradual*), como por exemplo, respostas certa, parcialmente certa ou errada; e, as **não ordenadas** entre si (modelo *nominal*), que trata da análise dos distratores, ou seja, avalia-se se as respostas foram as mais próximas da correta ou não [para mais detalhes dos modelos de resposta gradual e nominal veja Andrade et al. (2000), Pinheiro et al. (2010) e Soares (2005)]. Esses modelos de resposta não dicotômica visam obter mais informações das respostas dos indivíduos do que simplesmente se eles responderam corretamente ou incorretamente aos itens.

Nesse trabalho utilizamos somente modelos da TRI para itens dicotômicos e que avaliam apenas uma habilidade (unidimensional). Modelos que consideram que mais de uma habilidade está sendo medida são chamados de modelos multidimensionais.

Para itens dicotômicos há o modelo de *um parâmetro* que envolve apenas a dificuldade do item, o modelo de *dois parâmetros*, que envolve a dificuldade e a discriminação e, o modelo de *três parâmetros*, que envolve a dificuldade, a discriminação e a probabilidade de acerto ao acaso. Em todos os modelos a habilidade dos indivíduos é representada por θ .

Considerando n indivíduos submetidos a um teste de múltipla escolha, o qual possui I itens e k alternativas cada um. Se um indivíduo j responde corretamente ao item i , tem-se $y_{ij} = 1$ e $y_{ij} = 0$ em caso contrário. Assumindo que $y_{ij} \sim b(1, p_{ij})$, a probabilidade p_{ij} de que o indivíduo j responda corretamente ao item i , é dada pela função de resposta ao item

$$p_{ij} = P(Y_{ij} = 1 / \theta_j) = c_i + (1 - c_i)F(m_{ij}) \quad (1.1)$$

onde $i = 1, \dots, I$; $j = 1, \dots, n$ e $m_{ij} = a_i(\theta_j - b_i) = a_i\theta_j - b_i^*$ é uma função linear da variável latente θ_j associada à habilidade do indivíduo j , $b_i^* = a_i b_i$, a_i corresponde à discriminação, b_i à dificuldade e c_i à probabilidade de acerto ao acaso associados ao item i .

A função $F(\cdot)$ de (1.1) é estritamente não-decrescente, isto é, $\lim_{x \rightarrow -\infty} F(x) = 0$ e $\lim_{x \rightarrow +\infty} F(x) = 1$ e, é chamada de função de resposta ao item ou *curva característica do item (CCI)* e satisfaz alguns pressupostos como da *unidimensionalidade* do teste, ou seja, existe uma habilidade dominante responsável pela realização do conjunto de itens. O modelo

descrito baseia-se no fato de que os indivíduos com maior habilidade possuem maior probabilidade de acertar o item e que esta relação é não-linear (ANDRADE et al., 2000).

Admite-se que de acordo com essa única habilidade, as respostas dos indivíduos aos itens que compõem o teste são independentes entre si (*independência local*), ou seja, fixado uma habilidade θ , a resposta do indivíduo a um item não irá influenciar as respostas aos demais itens do teste (ANDRADE et al., 2000). Assim, de acordo com a independência local, para um indivíduo que possui certa habilidade, a probabilidade conjunta das respostas de um indivíduo a um conjunto de I itens é dado pelo produto das probabilidades de resposta do indivíduo j a cada item, onde $\zeta_i = (a_i, b_i, c_i)$ e $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_i, \dots, \zeta_I)$, conforme mostra a equação a seguir:

$$P(Y_{.j} = y_{.j} / \theta_j, \zeta) = \prod_{i=1}^I P(Y_{ij} = y_{ij} / \theta_j, \zeta_i) \quad (1.2)$$

De acordo com (1.1), reescreve-se (1.2) da seguinte forma:

$$P(Y_{.j} = y_{.j} / \theta_j, \zeta) = \prod_{i=1}^I (p_{ij})^{y_{ij}} (1 - p_{ij})^{1-y_{ij}} \quad (1.3)$$

Possíveis funções para $F(\cdot)$ são a probito e a logística, como mostramos a seguir.

1.1.1. Modelos TRI simétricos de três parâmetros

O maior interesse nesse trabalho é em estudar os modelos de três parâmetros, que significa estarmos interessados em estimar os parâmetros de discriminação, de dificuldade e de acerto ao acaso de cada item (a_i , b_i e c_i) na modelagem de resposta correta.

Em (1.1), considerando $F(\cdot) = \Phi(\cdot)$, com $\Phi(\cdot)$ a função acumulada da distribuição normal padrão (introduzido por Lord na TRI em 1952), a CCI é conhecida como ogiva normal ou probito normal.

O modelo probito com 3 parâmetros é dado por

$$p_{ij} = c_i + (1 - c_i)\Phi(m_{ij}) \quad (1.4)$$

Dos modelos dicotômicos para testes de múltipla-escolha, em particular, os utilizados na TRI, destacamos os logísticos unidimensionais e, principalmente, o modelo de três parâmetros, que é utilizado nesse estudo devido ao fato de ser o mais geral em termos de parâmetros dos itens a serem analisados, além de ser muito usado em análise de testes educacionais.

Considerando em (1.1) $F(.) = G(.)$, com $G(m_{ij}) = 1/(1 + e^{-m_{ij}})$ representando a função de distribuição acumulada da distribuição logística padrão (introduzido por Birnbaum na TRI em 1968) a CCI é conhecida como função logística.

Sua formulação é dada por,

$$p_{ij} = c_i + (1 - c_i) \frac{1}{1 + e^{-D a_i (\theta_j - b_i)}} \quad (1.5)$$

onde D é um fator de escala igual a 1 para a função logística, ou 1,7 para obter uma aproximação para a distribuição normal (ANDRADE et al., 2000).

No contexto de modelos lineares generalizados, a representação dada em (1.1) utiliza uma função de ligação F^{-1} , na TRI temos que $F^{-1}(p_{ij}) = m_{ij}$. Quando essa função de ligação é Φ^{-1} (isto é, inversa da função distribuição acumulada Normal padronizada) ela é denominada ligação probito; quando é G^{-1} chama-se função de ligação logito ou logística (SANTOS et al., 2009). Esses modelos são chamados de modelos simétricos.

Assim como o parâmetro de dificuldade b , a habilidade θ pode assumir qualquer valor entre $-\infty$ e $+\infty$. Portanto, precisa-se estabelecer uma origem e uma unidade de medida para a definição da escala. Esses valores são escolhidos de modo a representar, respectivamente, o valor médio (μ) e o desvio padrão (σ) dos desempenhos dos indivíduos da população em estudo (ANDRADE et al., 2000). A construção das escalas é necessária para estabelecer uma relação entre os parâmetros dos itens e a habilidade do indivíduo. Usualmente, utiliza-se a escala com média igual a zero e desvio padrão igual a um. A escala da habilidade é uma escala arbitrária, onde o importante são as relações de ordem existentes entre seus pontos e não necessariamente sua magnitude (SANTOS et al., 2009). Nos modelos simétricos, consideramos que $\theta_j \sim N(0,1)$. Valores positivos de θ_j indicam habilidade superior à média e valores negativos, habilidade abaixo da média.

1.1.2. Interpretação dos parâmetros e Representação Gráfica

A representação gráfica de $P(Y_{ij} = 1 / \theta_j, a_i, b_i, c_i)$, ou seja, a probabilidade do indivíduo j responder corretamente o item i , considerando os parâmetros desse item e uma habilidade θ_j , é dada pela Curva Característica do Item (CCI), como mostra a Fig. 1 a seguir.

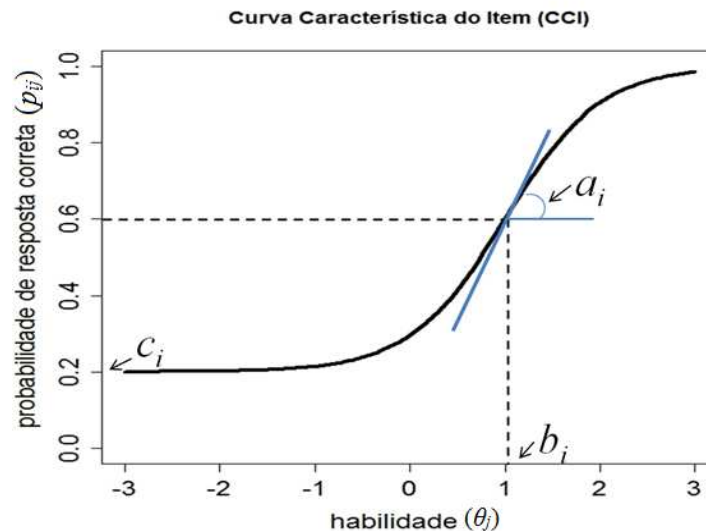


Figura 1 – Exemplo de uma CCI de três parâmetros identificando o significado de cada parâmetro.

É possível fazer várias análises a partir da Fig. 1, tais como:

- (1) o eixo das ordenadas representa a probabilidade de acerto do item enquanto o eixo das abscissas indica a habilidade dos indivíduos;
- (2) à medida que a habilidade do indivíduo aumenta, a probabilidade de resposta correta também aumenta;
- (3) b_i é o parâmetro que representa a dificuldade do item i , medido na mesma escala da habilidade, $b_i \in (-\infty, +\infty)$. Quanto maior o seu valor, maior é a habilidade exigida do indivíduo para responder corretamente determinado item;
- (4) $a_i > 0$ é proporcional à inclinação da CCI no ponto b_i e corresponde a discriminação do item i ; quanto maior o valor deste parâmetro, maior será o poder de discriminação;
- (5) c_i é o parâmetro do item que representa a probabilidade de acerto ao acaso, ou seja, de indivíduos com baixa habilidade responderem corretamente o item i , $c_i \in [0, 1]$.

O parâmetro b representa a habilidade necessária para uma probabilidade de acerto igual a $(1 + c)/2$. Quando $c = 0$, b representa o ponto na escala da habilidade onde a probabilidade de resposta correta ($y_{ij} = 1$) é 0,5 (ANDRADE et al., 2000).

É possível perceber ao analisar o gráfico da CCI que um indivíduo com habilidade superior a 3 tem probabilidade muito alta de acertar determinado item (acima de 80%), já um indivíduo com baixa habilidade (-3) tem probabilidade c (de 20%) de acertar determinado item. Pode-se concluir, também, que um valor alto de a representa discriminação alta, quando uma pequena mudança na habilidade pode resultar em uma mudança acentuada na probabilidade de resposta correta.

Um indivíduo com habilidade 1,3 na escala $N(0,1)$ e que tem uma habilidade de 1,3 desvios padrão acima da média, apresenta um conhecimento consideravelmente maior do conteúdo em estudo do que um indivíduo com habilidade -0,3, ou seja, 0,3 desvios padrão abaixo da média dessa mesma população. Para analisar o que esses valores significam é preciso uma interpretação minuciosa da escala de habilidade dos testes. No gráfico da CCI, espera-se que os valores dos parâmetros θ e b variem entre -3 e 3, isto é, entre 3 desvios padrão acima e abaixo da média.

1.2. Distribuição Normal Assimétrica

De acordo com Santos et al. (2009) e Bazán (2005), uma variável aleatória X tem distribuição normal assimétrica com parâmetro de locação μ , parâmetro de escala σ^2 e parâmetro de assimetria d , se a sua função densidade de probabilidade é

$$f_X(x) = \phi_{SN}(x; \mu, \sigma^2, d) = \frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(d \frac{x-\mu}{\sigma}\right) \quad (1.6)$$

onde $\phi(\cdot)$ é a densidade de uma distribuição normal padrão e $\Phi(\cdot)$ é a função de distribuição acumulada de uma normal padrão. Simbolizamos essa distribuição com $X \sim SN(\mu, \sigma^2, d)$ (SN do termo inglês *skew normal*).

A distribuição acumulada dessa variável X é dada por

$$\Phi_{SN}(x; \mu, \sigma^2, d) = 2\Phi_2\left(\begin{pmatrix} x \\ 0 \end{pmatrix}; \begin{pmatrix} \mu \\ 0 \end{pmatrix} \begin{pmatrix} \sigma^2 & -\lambda_i \\ -\lambda_i & I \end{pmatrix}\right) \quad (1.7)$$

em que $\Phi_2(\cdot)$ é a f.d. da normal padrão bivariada com o coeficiente de correlação $-\lambda_i$, $\lambda_i = \frac{d_i}{\sqrt{1+d_i^2}} \in [-1,1]$, onde d_i é o parâmetro de assimetria e a transformação de d_i para λ_i é de

um para um [para mais detalhes da distribuição normal assimétrica veja Azzalini (1985, 1986) e Pinheiro (2007), que traz algumas propriedades dessa distribuição].

As Fig. 2 e 3 mostram as diferenças entre a forma da distribuição normal simétrica e duas distribuições normais assimétricas; uma assimétrica positiva e outra negativa.

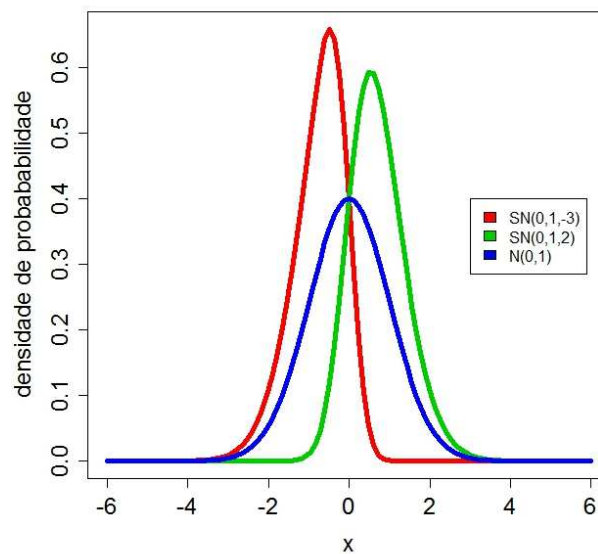


Figura 2 – Gráfico das distribuições utilizadas em nosso estudo para a habilidade de indivíduos (distribuição normal e distribuição normal assimétrica).

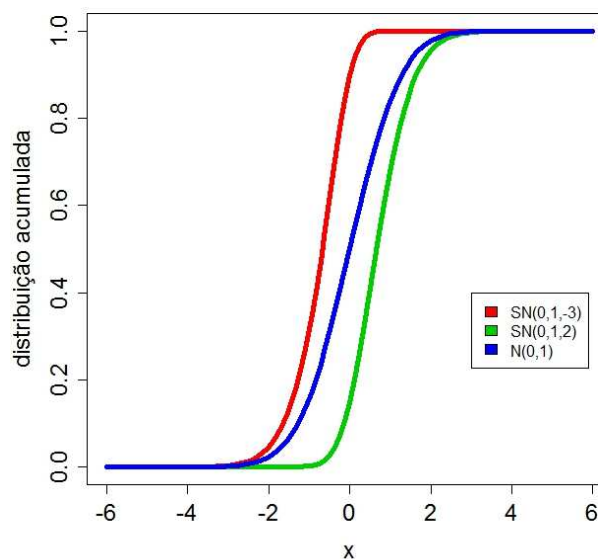


Figura 3 – Distribuição acumulada das distribuições da figura acima.

1.2.1. Modelos TRI Assimétricos

É comum utilizar modelos simétricos para descrever tanto a distribuição das habilidades dos indivíduos como a CCI. Porém, segundo Chen et al. (1999) esses modelos nem sempre fornecem bons ajustes para determinados conjuntos de dados, especialmente quando a probabilidade de uma dada resposta se aproxima de zero de forma diferente de como se aproxima de 1, afirmando que a suposição de simetria é muito restritiva para modelar a conduta humana.

É possível assumir que a habilidade tem distribuição normal assimétrica e/ou que a CCI seja assimétrica. Bazán (2005) apresenta uma família de modelos chamados TRI-NA (Teoria da Resposta ao Item Normal Assimétrica) para modelos TRI de dois parâmetros, que é formada por modelos probito que incluem assimetria na função de ligação e/ou na habilidade dos indivíduos. Pinheiro (2007) além de abordar os modelos apresentados em Bazán (2005) traz o modelo probito t-Assimétrico, que modela a probabilidade do indivíduo acertar um item através da f.d. da t-Student bivariada e usa a distribuição normal assimétrica para a habilidade θ . Em Bolfarine e Bazán (2008) utiliza-se função de ligação expoente logístico positivo (LPE) para o modelo de dois parâmetros.

Quando a assimetria está presente apenas na distribuição da habilidade θ , tem-se que p_{ij} é da forma (1.1) e $F(m_{ij}) = \Phi(m_{ij})$ ou $G(m_{ij})$ e, $\theta_j \sim SN(\mu_\theta, \sigma_\theta^2, d_\theta)$. Quando X tem $\mu_\theta = 0$, $\sigma_\theta^2 = 1$ e $d_\theta \neq 0$, a densidade de X em (1.6) se reduz a densidade da normal assimétrica padrão, dada por Azzalini (1985) e, neste caso, escrevemos que $X \sim SN(d)$ cuja função de densidade é dada por:

$$f_x(x) = 2\phi(x)\Phi(dx)$$

Caso não exista assimetria na habilidade ($d = 0$), temos $\theta_j \sim N(\mu, \sigma^2)$, ou a versão padronizada quando $\mu_\theta = 0$ e $\sigma_\theta^2 = 1$.

Se uma função de ligação assimétrica for considerada na modelagem da probabilidade de resposta correta aos itens, a CCI resultante será

$$p_{ij} = P(Y_{ij} = 1 | \theta_j, a_i, b_i, c_i, \lambda_i) = c_i + (1 - c_i)F(m_{ij}; \lambda_i) \quad (1.8)$$

Em (1.8), quando $F(m_{ij}; \lambda_i) = \Phi_{SN}(m_{ij}; \lambda_i)$, o parâmetro de assimetria $\lambda_i \in R$ e pode ser visto como um parâmetro de penalização: quando $\lambda_i > 0$, a probabilidade de resposta correta é penalizada para baixos valores da habilidade; já quando $\lambda_i < 0$, a probabilidade de resposta correta ao item i é penalizada para altos valores da habilidade. Quando $\lambda_i = 0$ obtém-se o modelo simétrico (SANTOS et al., 2009b).

Samejima (1997, 2000) criou uma família de modelos assimétricos chamados de expoente logístico positivo (FLPE), essa família tem como caso particular o modelo logístico simétrico. Samejima (2000) considera $F(m_{ij}; \lambda_i) = G(m_{ij})^{\lambda_i} = G(m_{ij}; \lambda_i)$, com $G(\cdot)$ uma f.d. da função logística padrão. Na TRI o parâmetro de assimetria λ_i é positivo e se relaciona à complexidade do item i . Um item é dito complexo se requer muitos passos para ser resolvido. Espera-se que a probabilidade de resolver um item complexo diminua quando o número de passos ou etapas aumenta, dada a habilidade do indivíduo. Tal parâmetro permite dar mais crédito a itens mais difíceis respondidos corretamente ($\lambda > 1$ ou assimetria positiva), bem como penalizar itens fáceis que são respondidos incorretamente ($0 < \lambda < 1$ ou assimetria negativa), se $\lambda = 1$ obtemos o modelo logístico (SANTOS et al., 2009b; SAMEJIMA, 2000). Para mais detalhes dessa família de modelos FLPE ver Samejima (1997, 2000).

Desta forma, podemos obter outra CCI para o modelo de TRI considerando a seguinte f.d.,

$$F(x) = (1 + e^{-x})^{-\lambda}, \lambda > 0 \quad (1.9)$$

A ligação correspondente a distribuição dada em (1.9) foi proposta por Prentice (1976) e a f.d. é chamada de distribuição *Burr type II* padrão ou ligação *power logit* denominado por Bolfarine e Bazán (2008).

Quando $F(\cdot)$ é a f.d. (1.9), indexado com λ_i e avaliado em m_{ij} , tem-se o chamado modelo LPE e λ está associado à escolha da função de ligação. Optamos por chamar esse modelo em nosso estudo por *modelo logístico assimétrico*.

A Fig. 4 traz diferentes CCI's com os parâmetros dos itens a , b e c fixos e diferentes valores para a assimetria λ .

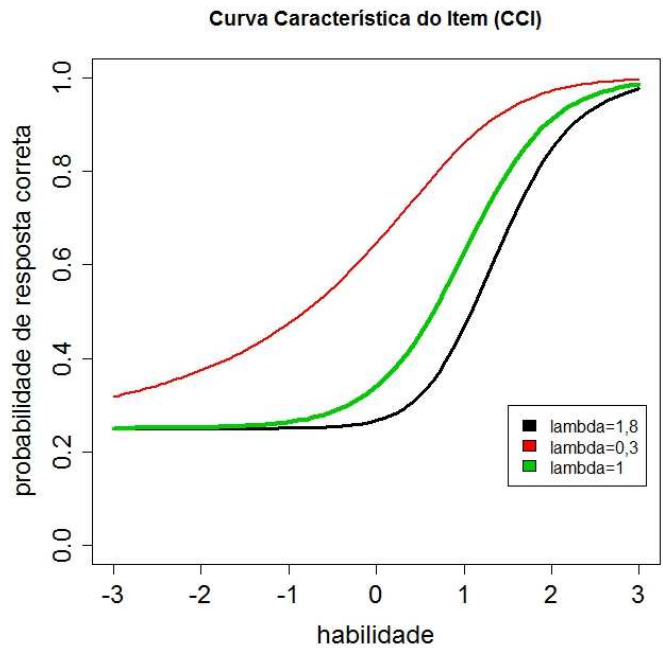


Figura 4 – CCI considerando fixos os valores dos parâmetros $a = 2$, $b = 1$, $c = 0,25$, $\theta \sim N(0,1)$ e variando λ . Se $\lambda = 1,8$ e $\lambda = 0,3$ (modelo assimétrico), se $\lambda = 1$ (modelo simétrico).

Analisando a Fig. 4, podemos verificar grande mudança na probabilidade de resposta correta conforme λ varia, principalmente quando λ se aproxima de zero. Para um indivíduo com habilidade média zero ($\theta = 0$) a probabilidade de acertar o item é aproximadamente 0,65 quando $\lambda = 0,3$. Já para $\lambda = 1,8$, a probabilidade de acertar o item é baixa, cerca de 0,27 e para $\lambda = 1$ é 0,34.

Na Fig. 5, apresentamos os quatro diferentes modelos de três parâmetros da TRI que podemos obter quando considerando simetria ou assimetria na CCI e na distribuição das habilidades.

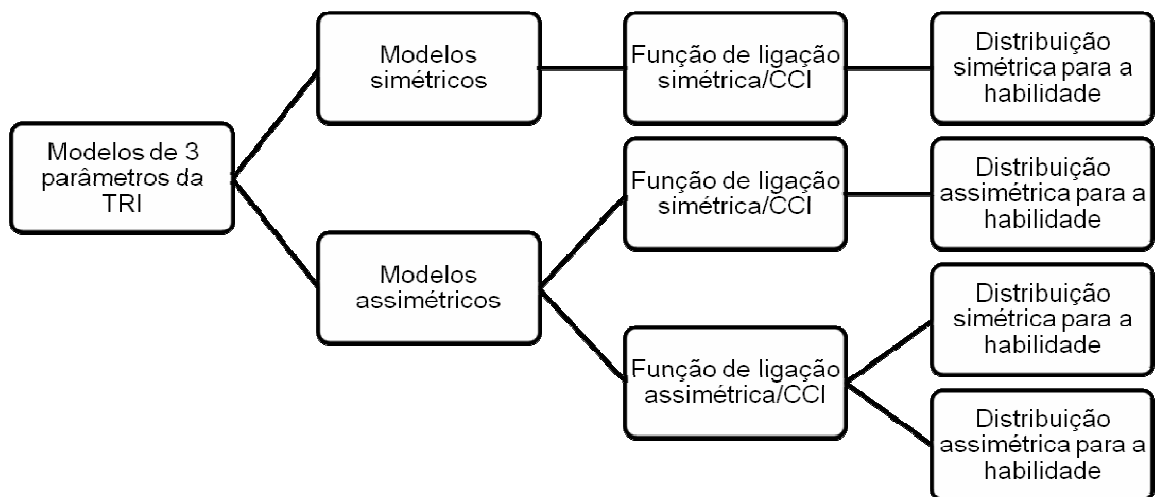


Figura 5 - Composição dos diferentes modelos da TRI para itens dicotômicos.

1.3. Justificativa

É importante utilizar modelos que detectem a assimetria sempre que esteja presente nos dados. Sob essa perspectiva, adotar um modelo simétrico para a modelagem das probabilidades de respostas poderiam não ser úteis para fazer inferência, uma vez que, devido à possível assimetria, as estimativas das proficiências e dos parâmetros dos itens não seriam confiáveis (SANTOS et al., 2009b).

Os modelos assimétricos podem ser úteis na aplicação a populações nas quais haja certa evidência de assimetria como, por exemplo, em populações de baixa renda, classificação por regiões, etc., na busca de se obter vantagens no uso desses modelos em relação aos modelos de TRI simétricos, por serem mais flexíveis.

O que geralmente ocorre em estudos da TRI é supor que os indivíduos são provenientes de uma população com distribuição das habilidades $N(0,1)$. Mas, se essa suposição não for verdadeira? As estimativas serão confiáveis? E, se a distribuição das habilidades é assimétrica, será que utilizando um modelo assimétrico (função de ligação assimétrica) para o ajuste é melhor do que se fosse utilizado um modelo simétrico? Foi pensando nisso que elaboramos esse estudo.

Muitas aplicações descritas na literatura e que lidam com assimetria, utilizam o modelo probito normal de dois parâmetros, ou seja, não incluem o parâmetro de acerto ao acaso (c). Em contraste, utilizamos os modelos de ajuste com CCI logística simétrica e assimétrica de três parâmetros nesse estudo, além de considerar uma distribuição normal assimétrica para a habilidade na geração de dados simulados para análise, visto que, a distribuição da habilidade da população da qual a amostra é retirada em análises de testes pode ter essa estrutura. A CCI logística assimétrica, proposta por Samejima (1997; 2000) vem sendo utilizada em estudos recentes como de Bolfarine e Bazán (2008).

A finalidade desse estudo, portanto, é verificar se o ajuste melhora utilizando uma função de ligação assimétrica (assimetria associada aos itens).

Finalmente, o método TRI de avaliação, uma vez bem compreendido e aplicado, poderá auxiliar os professores na elaboração de instrumentos mais efetivos para avaliar as habilidades dos alunos, buscando metodologias e estratégias adequadas para suprir as dificuldades apresentadas por eles. Também, é de suma importância a expansão de estudos utilizando esta metodologia em outras áreas do conhecimento.

1.4. Objetivos

O objetivo geral é investigar a eficiência de estimações nos modelos simétricos e assimétricos para a TRI de três parâmetros.

Objetivos específicos

- Implementar computacionalmente a geração de dados considerando as funções de ligação logística simétrica e logística assimétrica, utilizando distribuições normal padrão e normal assimétrica para as habilidades;
- Implementar computacionalmente a estimação bayesiana dos parâmetros do modelo da TRI utilizando funções de ligação logística simétrica e logística assimétrica e considerando a distribuição normal padrão para a habilidade;
- Analisar a eficiência da estimação dos parâmetros dos itens e das habilidades em modelos de ajuste com CCI simétrica e assimétrica.

1.5. Estrutura do trabalho

Segue um breve roteiro para os próximos capítulos. No capítulo 2 discutimos o processo inferencial sob a abordagem bayesiana, o método MCMC, técnicas de diagnóstico de convergência e métodos de seleção de modelos. No capítulo 3, detalhamos a metodologia utilizada, bem como as três etapas do processo: (1) geração dos dados sintéticos; (2) estimação dos parâmetros do modelo e (3) análise da qualidade das estimações. No capítulo 4, apresentamos os principais resultados obtidos do estudo. Finalmente, no capítulo 5, as conclusões e sugestões para trabalhos futuros.

Ao final, são apresentados dois apêndices: o Apêndice 1 mostra resultados detalhados do estudo feito, como estimativas dos parâmetros, desvios padrão, intervalos de credibilidade de 95%, etc., o Apêndice 2 traz o algoritmo utilizado para ajuste simétrico e assimétrico dos parâmetros e como foi feita a geração de dados para dois dos seis casos em estudo utilizando o software R.

2. MÉTODO DE ESTIMAÇÃO

O processo de estimação dos parâmetros dos itens (a_i , b_i e c_i) é chamado de *calibração*. No caso de provas distintas, os parâmetros dos itens ou as habilidades de respondentes de diferentes grupos devem estar todos na mesma escala métrica para que os itens ou habilidades possam ser comparáveis, e a isso dá-se o nome de *equalização* (ANDRADE et al., 2000).

A inferência em modelos da TRI pode ser feita por máxima verossimilhança ou bayesiana. Os métodos bayesianos de estimação permitem contornar dificuldades de estimação por máxima verossimilhança. Particularmente, quando há itens respondidos corretamente ou incorretamente por todos os respondentes (ANDRADE et al., 2000) e, também, quando há respondentes com habilidades elevadas que respondem incorretamente itens fáceis ou vice-versa (AZEVEDO, 2003). Implementamos o procedimento bayesiano de estimação através do método de Monte Carlo via Cadeias de Markov, pelo qual os parâmetros dos itens e das habilidades dos indivíduos são estimados.

A quantidade de parâmetros a serem estimados é consideravelmente grande, sendo igual a $n + (3 + 1_s)I$ para o total de n indivíduos, I itens e 1_s a função indicadora que será igual a um se o modelo é assimétrico e igual a zero caso contrário.

2.1. Inferência Bayesiana

Se X denota a totalidade da informação disponível sobre o parâmetro desconhecido θ antes da aquisição de dados, então a incerteza do pesquisador é representada pela distribuição priori $p(\theta|X)$. Os dados (Y) são descritos por uma distribuição de probabilidades, dada por $p(Y|\theta, X)$, denominada função de verossimilhança de θ para um valor fixado de y . Para atualizar a informação disponível sobre θ e obter a sua distribuição posterior, aplica-se o Teorema de Bayes definido por (PAULINO et al., 2003):

$$p(\theta|Y, X) = \frac{p(Y|\theta, X) p(\theta|X)}{p(Y|X)}$$

Para obter a distribuição posterior, recorre-se a soluções analíticas fechadas sempre que possível. No entanto, na grande maioria das aplicações, essas soluções são extremamente

difíceis e/ou indisponíveis. Nesse caso, recorre-se a procedimentos de simulação estocástica para gerar uma amostra da distribuição posterior e utiliza-se aproximação de Monte Carlo para as integrais de interesse (GELMAN et al., 2004).

A estimação bayesiana consiste em utilizar a distribuição posterior como sendo a representação atualizada e completa da incerteza sobre θ . Estimativas pontuais de Bayes consistem de alguma característica particular dessa posterior e dependerá da função de perda escolhida (KINAS e ANDRADE, 2010). Por exemplo, a Moda posterior minimiza a função de perda 0-1; a Média posterior minimiza a função de perda quadrática e a Mediana posterior minimiza a função de perda absoluta.

No contexto Bayesiano, as medidas de incerteza usuais associadas à distribuição posterior são sua variância e os intervalos de credibilidade.

Para qualquer estimador $\delta(x)$ de θ , podemos calcular a variância posterior, dada pela expressão (KINAS e ANDRADE, 2010):

$$V_{\delta}(x) = E^{p(\theta|x)}[\theta - \delta(x)]^2 = V(\theta | x) + [E^p[\theta | x] - \delta(x)]^2$$

O intervalo de credibilidade de 95% (ICr_{95%}) para θ é definido como sendo o intervalo delimitado pelos percentis 2.5% e 97.5% da distribuição posterior $p(\theta|Y)$ (KINAS e ANDRADE, 2010). Quanto menor for o tamanho do intervalo mais concentrada é a distribuição do parâmetro e mais precisa é a estimação (EHLERS, 2007).

2.2. Método de Monte Carlo via Cadeias de Markov (MCMC)

A associação de algoritmos para a simulação de distribuições com o método de Monte Carlo para aproximações de integrais conduz aos chamados “Métodos de Monte Carlo via Cadeias de Markov” (PAULINO et al., 2003). Há alguns métodos que podem ser utilizados para construção de uma cadeia de Markov, como o algoritmo de Metropolis-Hastings e um caso particular desse, o Amostrador de Gibbs.

De acordo com Albert (2007), a amostragem MCMC configura uma cadeia de Markov irreduzível e aperiódica para a qual a distribuição estacionária é igual à distribuição posterior de interesse. Uma cadeia de Markov descreve o movimento probabilístico de transição entre vários estados, ou seja, as probabilidades de mover-se do local atual para outros locais são

especificadas e dependem apenas de sua localização atual e não dos locais anteriormente visitados.

No método MCMC há dois problemas a resolver e que podemos controlar da seguinte forma:

(1) *Dependência do ponto de partida*: um dos problemas é detectar o tamanho do período *burn-in*, ou seja, as iterações iniciais a serem descartadas por ainda não serem distribuídas a partir da distribuição posterior. No entanto, após um certo número de iterações realizadas, o efeito dos valores iniciais desaparece e a distribuição das novas iterações se aproximam da verdadeira distribuição posterior. Uma maneira de estimar o comprimento do período de *burn-in* é examinar o *trace plot* de valores simulados de um componente ou função específica de ψ em relação ao número de iteração (ALBERT, 2007).

(2) *Grau de autocorrelação entre os valores simulados*: uma estatística padrão para medir o grau de dependência entre sucessivas amostras na cadeia é a autocorrelação, que mede a correlação entre os conjuntos $\{\psi^j\}$ e $\{\psi^{j+L}\}$, onde L é o número de iterações que separa os dois conjuntos de valores (ALBERT, 2007). Portanto, para não haver autocorrelação e obter uma amostra independente, define-se o valor de salto (*thinning*), cuja amostra final é obtida saltando-se s valores a partir do ponto inicial.

Suponha que desejamos simular a partir de uma densidade posterior $g(\psi/y)$ ou, $g(\psi)$ por simplificação, onde $g(\psi)$ é o núcleo de transição da distribuição posterior. Um algoritmo de **Metropolis-Hastings** (uma extensão do algoritmo de Metropolis, que foi desenvolvido por Hastings em 1970) começa com um valor inicial ψ^0 e especifica-se uma regra para simular o t -ésimo valor na seqüência ψ^t dado o valor imediatamente anterior ψ^{t-1} . Essa regra consiste em simular um valor candidato ψ^* de uma *densidade proposta* e calcular uma probabilidade de aceitação P que indica a probabilidade de o valor candidato ser aceito para ser o próximo valor na seqüência. Especificamente, esse algoritmo pode ser descrito da seguinte forma (ALBERT, 2007):

- Simular um valor candidato ψ^* da densidade proposta $p(\psi^* | \psi^{t-1})$
- Calcular a taxa

$$R = \frac{g(\psi^*)p(\psi^{t-1} | \psi^*)}{g(\psi^{t-1})p(\psi^* | \psi^{t-1})}$$

- Calcular a probabilidade de aceitação $P = \min\{R, 1\}$
- Gerar um valor ψ^t tal que

$$\psi^t = \begin{cases} \psi^* & \text{com probabilidade } P \\ \psi^{t-1} & \text{com probabilidade } 1 - P. \end{cases}$$

Se as condições de regularidade são satisfeitas na densidade proposta $p(\psi^*/\psi^{t-1})$, então a sequência de amostras simuladas ψ^1, ψ^2, \dots converge para a distribuição posterior $g(\psi)$.

O método de **Amostragem Gibbs** (*Gibbs Sampling*), introduzido inicialmente por Geman e Geman (1984) com grande contribuição de Gelfand e Smith (1990), exige que se decomponha a distribuição posterior conjunta de interesse em distribuições condicionais completas para cada parâmetro das quais se consiga obter amostras. Seja $h(\psi/y)$ a distribuição posterior conjunta, que é determinada pelas distribuições condicionais completas $h(\psi_i / y, \psi_{(-i)})$, $i = 1, \dots, k$, com $\psi = (\psi_1, \psi_2, \dots, \psi_k)$ e $\psi_{(-i)}$ representa o vetor ψ sem a i -ésima componente, ou seja, $\psi_{(-i)} = (\psi_1, \dots, \psi_{i-1}, \psi_{i+1}, \dots, \psi_k)$. O algoritmo é descrito a seguir (PAULINO et al., 2003):

1. Inicialize $t = 1$ e o vetor $\psi^{(0)} = (\psi_1^{(0)}, \dots, \psi_k^{(0)})$ com um valor arbitrário.
2. Obtenha os valores de $\psi^{(t)}$ a partir de $\psi^{(t-1)}$, procedendo iterativamente da seguinte forma:
 - Amostre $\psi_1^{(t)}$ de $h(\psi_1 / y, \psi_2^{(t-1)}, \dots, \psi_k^{(t-1)})$,
 - Amostre $\psi_2^{(t)}$ de $h(\psi_2 / y, \psi_1^{(t)}, \psi_3^{(t-1)}, \dots, \psi_k^{(t-1)})$,
 - ...
 - Amostre $\psi_k^{(t)}$ de $h(\psi_k / y, \psi_1^{(t)}, \dots, \psi_{k-1}^{(t)})$. Completando assim uma iteração, ou seja, de $\psi^{(0)}$ para $\psi^{(t)} = (\psi_1^{(t)}, \dots, \psi_k^{(t)})$.
3. Incremente o contador de modo que $t = t + 1$. Se t é menor que o número de amostras desejadas T , retorne ao passo 2. Caso contrário pare e retorne os valores de $\psi^{(1)}, \psi^{(2)}, \dots, \psi^{(T)}$.

Segundo Azevedo (2003), para utilizar esse algoritmo de Amostragem Gibbs para a construção das cadeias é necessário que as distribuições condicionais completas dos parâmetros desconhecidos sejam conhecidas e que dela seja possível gerar valores simulados; o que nem sempre é possível. Para resolver esse problema, recorre-se ao algoritmo de Metropolis-Hastings. A combinação dessas duas estratégias é chamada de *Metropolis-Hastings within Gibbs*.

Portanto, o algoritmo chamado de *Metropolis-Hastings within Gibbs* amostra iterativamente a partir das condicionais completas de acordo com o algoritmo Gibbs, mas, para completar essas condicionais, cuja forma é conhecida a menos de uma constante de normalização, é necessário o uso do algoritmo de Metropolis-Hastings (PATZ e JUNKER, 1999).

Uma vez obtida a amostra da distribuição posterior, calculam-se estimativas amostrais das características de interesse dessa distribuição como, por exemplo, a sua média ou mediana (EHLERS, 2007; KINAS e ANDRADE, 2010). O uso de técnicas de diagnóstico é necessário para verificar se a convergência foi atingida.

2.3. Inferência na TRI e escolha de prioris

Para fazermos inferências acerca dos parâmetros estudados nos modelos da TRI, necessitamos da especificação da função de verossimilhança e das distribuições de probabilidade a priori para os mesmos.

A função de verossimilhança associada a modelos da TRI para itens dicotômicos, assumindo independência local é dado por

$$L(\theta, a, b, c / y) = \prod_{j=1}^n \prod_{i=1}^I p_{ij}^{y_{ij}} (1 - p_{ij})^{1 - y_{ij}}$$

p_{ij} é dado pelo modelo geral (1.1).

Se tomarmos o modelo dado em (1.1), ou ainda, incluímos também a assimetria λ na CCI (1.9) e, considerarmos prioris independentes para os parâmetros, podemos escrevê-la como

$$\pi(\theta, a, b, c, \lambda) = \prod_j^n \phi\left(\frac{\theta_j - \mu}{\sigma}\right) \prod_i^I \pi_1(a_i) \pi_2(b_i) \pi_3(c_i) \pi_4(\lambda_i)$$

onde $\phi(\cdot)$ é o f.d.p normal padronizada e $\pi_1, \pi_2, \pi_3, \pi_4$ são as f.d.p's para os parâmetros a_i, b_i, c_i e λ_i , respectivamente. Devemos especificar a priori π_4 quando consideramos assimetria na CCI, caso contrário não é preciso.

Vamos considerar que cada parâmetro a_i tem distribuição *Normal truncada à direita*, $N(\mu_a, \sigma_a^2)I(a_i > 0)$ e cada parâmetro b_i tem distribuição *Normal*, $N(\mu_b, \sigma_b^2)$. Como c_i é uma probabilidade entre 0 e 1, utilizamos a distribuição *Beta*, $B(\alpha, \beta)$ e para λ_i uma distribuição *Gama*, $G(\alpha, \beta)$, com média $\frac{\alpha}{\beta}$ e variância $\frac{\alpha}{\beta^2}$. O modelo tem a seguinte estrutura hierárquica:

$$\begin{aligned} y_{ij} / \theta_j, a_i, b_i, c_i, \lambda_i &\sim b(1, p_{ij}) \\ p_{ij} &= c_i + (1 - c_i)F_\lambda(m_{ij}) \\ \theta_j &\sim N(0, 1) \\ a_i &\sim N(1, 0.5)I(a_i > 0) \\ b_i &\sim N(0, 2) \\ c_i &\sim B(2, 8) \\ \lambda_i &\sim G(0.25, 0.25) \end{aligned}$$

As *prioris* para a_i e b_i foram usadas nos estudos de Sahu (2002), Pinheiro (2007), Bazán (2005) e Marques (2008). A *priori* para λ_i foi utilizada em estudos de Carlin e Louis (2000) e Bolfarine e Bazán (2008) quando utilizada a ligação logística assimétrica (1.9).

Marques (2008) apresenta um estudo do efeito do tamanho do *salto* nas estimativas dos parâmetros dos itens e uma análise de sensibilidade à escolha de distribuições a priori para o modelo probito normal de três parâmetros. Foram utilizados dois conjuntos de dados, o primeiro com amostra de 131 indivíduos e o segundo com 277 indivíduos. O estudo utilizou critérios como erro de Monte Carlo e a razão Monte Carlo para analisar o efeito dos saltos, os quais quando muito elevados indicam que muitos dos valores gerados serão descartados, resultando em gasto maior de tempo computacional; e para análise da sensibilidade às *prioris* utilizaram o teste de similaridade baseado no coeficiente de concordância de Kendall (CCK). A conclusão do estudo é que para os dois conjuntos de dados são suficientes considerar um salto igual a 2 e uma amostra igual a 10000. Porém, o tamanho da amostra interfere nas estimativas, pois para a amostra composta por 277 indivíduos produziu intervalos de credibilidade bem menores e, conseqüentemente, estimativas mais precisas, além de apresentar menor sensibilidade a escolha de *prioris* do que com uma amostra menor de 131 indivíduos.

2.4. Técnicas de diagnóstico de convergência das cadeias

Um aspecto muito importante na inferência Bayesiana é verificar a convergência da cadeia para a distribuição estacionária. Salientamos que as técnicas de diagnóstico foram projetadas para verificar uma condição necessária mas não suficiente para a convergência. Isto é, não existem provas conclusivas para se afirmar quando a cadeia de Markov convergiu para a distribuição estacionária (SAS/STAT(R)).

Algumas técnicas são sugeridas por Rodrigues et al. (2009), Paulino et. al (2003) e SAS/STAT(R), que serão utilizadas em nosso estudo. Segundo esses autores, uma abordagem empírica para controlar a convergência é desenhar **figuras (traço) das cadeias** simuladas para detectar comportamentos de não-estacionariedade. Há alguns testes estatísticos de diagnóstico como o teste de **Geweke** (1992) que é baseado em séries temporais e faz uma comparação entre as médias das n primeira iterações (descartando o *burn-in*) e as n últimas iterações; se a diferença padronizada entre as médias for muito grande é indicativo de não convergência. O teste de **autocorrelação**, que serve para verificar a independência entre valores sucessivos, outra exigência para usar os dados simulados como amostra da posterior [mais detalhes em SAS/STAT(R)].

Outro teste muito conhecido é o teste **Gelman e Rubin** (1992). Esse teste é baseado em técnicas de análise da variância, onde se faz uso de várias cadeias em paralelo a partir de diferentes pontos. Para que haja convergência, todas as cadeias devem apresentar o mesmo comportamento, sendo que a variância entre as cadeias deve ser bem menor que a variância dentro das mesmas.

Para calcular essas variâncias definimos m como as múltiplas sequências simuladas, cada uma de tamanho n (depois de descartar a primeira parte das simulações). Para cada escalar estimado t , classificamos a mostra das simulações como t_{ij} ($i = 1, \dots, n; j = 1, \dots, m$). Essas variâncias são calculadas como a seguir (GELMAN et al., 2004).

Variância entre as cadeias:

$$E = \frac{n}{m-1} \sum_{j=1}^m (\bar{t}_{.j} - \bar{t}_{..})^2, \text{ onde } \bar{t}_{.j} = \frac{1}{n} \sum_{i=1}^n t_{ij} \text{ e } \bar{t}_{..} = \frac{1}{m} \sum_{j=1}^m \bar{t}_{.j}$$

Variância média dentro das cadeias:

$$D = \frac{1}{m} \sum_{j=1}^m s_j^2, \text{ onde } s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (t_{ij} - \bar{t}_{.j})^2$$

onde $\bar{t}_{.j}$ é a média das observações da cadeia e $\bar{t}_{..}$ é a média de todas as observações.

É possível estimar a variância posterior marginal do estimando $\hat{V}(t/y)$ da seguinte forma (GELMAN et al., 2004):

$$\hat{V}^+(t/y) = \frac{n-1}{n} D + \frac{1}{n} E$$

De acordo com Gelman et al. (2004), enquanto as cadeias não tiverem convergido ainda haverá influência dos pontos iniciais que foram escolhidos de forma a ter sobredispersão e a variância entre (E) superestimar a verdadeira variância. Em contrapartida, a variância dentro (D) deve ser uma subestimativa da verdadeira variância, visto que cada cadeia não teria coberto adequadamente todo o espaço de estados e resultaria em menor variabilidade.

Outra forma de monitorar a convergência das cadeias é estimando o fator de redução potencial de escala. O fator de redução é dado por:

$$\hat{R} = \sqrt{\frac{\hat{V}^+(t/y)}{D}}$$

Um indicador de convergência da cadeia será dado pela sua proximidade do valor 1, caso \hat{R} seja alto (acima de 1.1), devemos proceder executando mais simulações para melhorar a inferência sobre a distribuição alvo associada ao escalar estimado (GELMAN et al., 2004). Este método de Gelman e Rubin (\hat{R}) faz parte dos métodos de diagnóstico implementados no CODA e no OpenBUGS conhecido como *Rhat*.

2.5. Critérios para seleção de modelos

Para avaliar qual modelo ajusta melhor os dados em estudo são adotados algumas medidas de diagnóstico para comparação e seleção de modelos. Utilizamos os critérios mais aplicados ao contexto bayesiano e que também são utilizados em estudos da TRI.

Bazán (2005) compara diferentes modelos da família TRI-NA utilizando o valor esperado do desvio bayesiano posterior (\bar{D}), o desvio bayesiano das médias posteriores (\hat{D}), o número efetivo de parâmetros pD e o critério de informação da deviância (DIC); essas medidas são apresentadas por Spiegelhalter et al. (2002). Além disso, Bazán utiliza o critério de Informação de Akaike Esperado (EAIC), o Critério de Informação (Schwarz) Bayesiana Esperado (EBIC) e a Soma dos Quadrados dos Resíduos Latentes posterior (SQRL). Pinheiro (2007) além de fazer uso do DIC e SQRL, utiliza o pseudo fator de bayes. Em Paulino et al. (2003) também podemos encontrar detalhes do pseudo fator de bayes, AIC, BIC e DIC.

Critérios que combinam medidas de complexidade do modelo com a qualidade de seu ajuste são baseados na **deviância** (*deviance*) (SPIEGELHALTER et al., 2002), definida como

$$D(\theta, a, b, c, \lambda) = -2 \ln(p(y | \theta, a, b, c, \lambda))$$

A deviância esperada é uma medida de predição do erro, considerado como um sumário padronizado e pode ser usado para comparar o desempenho de diferentes modelos (GELMAN et al., 2004), esse valor pode ser aproximado usando amostras do MCMC (BOLFARINE e BAZÁN, 2008)

$$\mathbf{E}[D(\theta, a, b, c, \lambda)] \approx \frac{1}{G} \sum_{g=1}^G D(\theta^g, a^g, b^g, c^g, \lambda^g) = \bar{D}$$

onde G é o número de amostras MCMC. \bar{D} é uma medida Bayesiana de ajuste ou adequancia, chamado de **média posterior do desvio** (SPIEGELHALTER et al., 2002).

O critério DIC (*Deviance Information Criterion*) apresentado por Spiegelhalter et al. (2002) é utilizado principalmente em problemas de inferência bayesiana para comparação de modelos hierárquicos e amostras da distribuição posterior dos parâmetros foram obtidas através do método MCMC.

O DIC pode ser estimado usando as saídas desse método,

$$DIC = \hat{D} + 2\rho D$$

onde ρD é uma medida de *complexidade* dado pelo número efetivo de parâmetros do modelo,

$$\rho D = E[D(\theta, a, b, c, \lambda)] - D(E[\theta], E[a], E[b], E[c], E[\lambda])$$

em que $D(E[\theta], E[a], E[b], E[c], E[\lambda])$ é o **desvio da média posterior**, ou seja, a média dos valores das médias posteriores geradas dos parâmetros do modelo e, pode ser aproximado pelo método da seguinte forma,

$$\hat{D} = D\left(\frac{1}{G} \sum_g \theta^g, \frac{1}{G} \sum_g a^g, \sum_g b^g, \sum_g c^g, \frac{1}{G} \sum_g \lambda^g\right)$$

No Openbugs, o DIC e ρD são saídas default do algoritmo, que considera $\rho D = \text{var}(\text{deviance})/2$ e $DIC = \bar{D} + \rho D$ (STURTZ et al., 2005). Quanto menor for o valor de DIC, melhor é o ajuste do modelo.

3. METODOLOGIA

O presente estudo refere-se a modelos da TRI de três parâmetros para respostas dicotômicas e foi efetuado em três etapas: (1) geração dos dados sintéticos; (2) estimação bayesiana dos parâmetros do modelo através do método MCMC e (3) análise da qualidade das estimações. Para isso, consideramos uma amostra de $n = 500$ indivíduos e um teste composto por $I = 15$ itens.

Todo processo foi implementado em linguagem R de programação, utilizando o recurso do OpenBUGS (THOMAS et al., 2006) através das bibliotecas R2WinBUGS (STURTZ et al., 2005) e BRUGS (THOMAS et al., 2006). Destacamos que o algoritmo de simulação MCMC encontra-se implementado no software OpenBUGS, sendo necessário especificar apenas as prioris para os parâmetros. Para análises utilizamos o pacote CODA (PLUMMER, 2010) e para gerar dados da distribuição normal assimétrica utilizamos o pacote SN (AZZALINI, 2010).

3.1. Geração dos dados sintéticos

a) Definição dos parâmetros dos itens e das habilidades dos indivíduos

Primeiramente consideramos conhecidos todos os parâmetros do modelo a , b , c , λ e θ para a geração dos vetores de respostas aos I itens dos n indivíduos, com 1 para acerto e 0 para erro.

Os valores de a , b , c , λ e θ foram gerados conforme critérios estabelecidos na literatura para construir um bom teste. Tais como:

- Os itens que fazem parte do teste devem ter boa qualidade, isto é, devem discriminar os indivíduos que estão fazendo o teste. Para Hambleton e Swaminathan (1985) itens com $a \geq 1$ apresentam bom poder de discriminação. Assim, valores para esse parâmetro foram sorteados de uma distribuição uniforme entre 0,5 e 2, buscando itens com níveis de discriminação média a alta;
- O teste deve conter itens fáceis, médios e difíceis. Segundo Azevedo (2003), espera-se que os valores de b variem entre -3 e 3, para Hambleton et al. (1991), entre -2 e 2.

Com base nisso, para não obter valores muito extremos, b 's foram sorteados de uma distribuição uniforme entre -1,5 e 2,3;

- Como c expressa a probabilidade de acerto ao acaso, supondo que o teste seja composto por 5 alternativas por item, tem-se que $c = 1/5 = 0,2$. Porém, considerando que, mesmo que um indivíduo tenha pouco conhecimento sobre o assunto, algumas alternativas fornecem certa informação levando a preferí-las e, assim, acaba excluindo alguma(s) alternativa(s), mudando essa probabilidade. Esses valores foram gerados aleatoriamente com reposição entre 0,15 e 0,3 com variação de 0,05;
- Como os valores de λ (assimetria nos itens) devem ser maior que zero por restrição do modelo, seus valores foram sorteados de uma distribuição uniforme entre 0,1 e 2.

Obtivemos assim, os valores verdadeiros para os parâmetros dos itens, mostrados na Tabela 1.

Tabela 1 – Valores verdadeiros dos parâmetros dos itens utilizados para gerar a matriz resposta ao teste dos n indivíduos

Item	a	b	c	λ
1	1.9193	-1.5084	0.15	1.8768
2	0.5257	0.3147	0.15	1.6329
3	1.9006	2.6490	0.20	0.7448
4	1.2944	0.5945	0.20	1.5051
5	0.6009	0.3470	0.30	1.5222
6	0.9168	0.1289	0.30	0.2896
7	1.0856	1.8745	0.20	0.8612
8	1.3699	0.5679	0.30	1.3059
9	1.0160	-1.0583	0.20	0.1103
10	1.5166	-0.8571	0.20	0.7172
11	1.8393	-0.0527	0.25	0.6546
12	1.1218	-0.0729	0.30	0.5861
13	1.5806	1.8159	0.25	1.6911
14	0.8972	0.1216	0.30	0.3981
15	1.3775	2.3123	0.20	1.0519

As habilidades dos indivíduos foram sorteadas das seguintes distribuições: (1) distribuição normal com média zero e desvio padrão um, $N(0,1)$; (2) distribuição normal assimétrica com média zero, desvio padrão um e assimetria negativa três, $SN(0, 1, -3)$ e (3) distribuição normal assimétrica com média zero, desvio padrão um e assimetria positiva dois, $SN(0, 1, 2)$.

A distribuição das habilidades verdadeiras está ilustrada nos histogramas abaixo (Fig. 6).

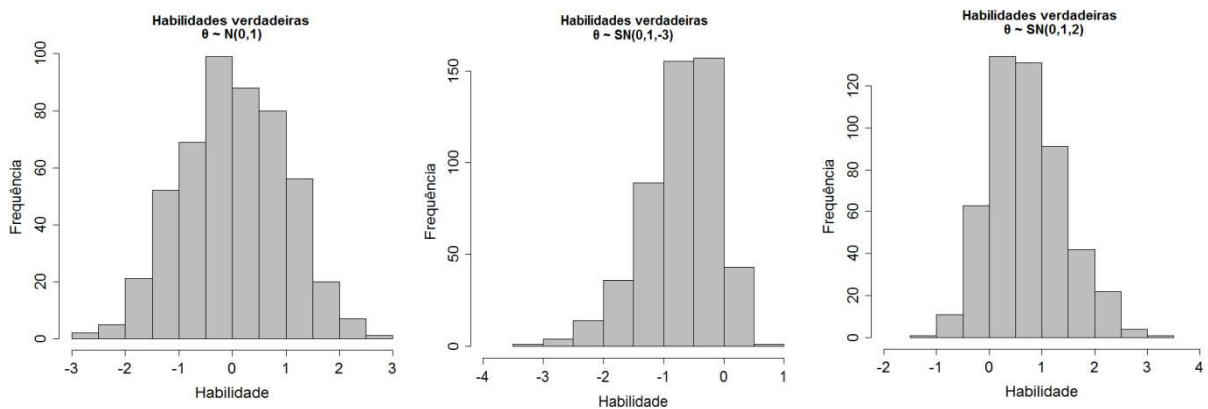


Figura 6 - Histograma das habilidades verdadeiras utilizadas no estudo.

b) Geração das respostas dos n indivíduos aos I itens (denominados dados)

Utilizamos seis conjuntos de dados sintéticos (*Dados 1*, *Dados 2*, ..., *Dados 6*), formados conforme a combinação apresentada na Fig. 7. Uma vez fixado todos os parâmetros para cada situação, geramos a matriz $n \times I$ das probabilidades de acerto da questão i pelo indivíduo j (p_{ij}). Em seguida, obtivemos a matriz $n \times I$ das respostas y_{ij} a partir de distribuições de Bernoulli com parâmetro p_{ij} , composta de zeros e uns.

Em cada uma das seis situações produzimos 20 réplicas. Cada réplica manteve os mesmos valores de a , b , c , λ e habilidades conforme o modelo definido para tal. Ao final, dispõem-se de $6 \times 20 = 120$ matrizes de dados sintéticos $Y_{(n \times I)}$.

Os resultados dessa etapa (*dados*) foram armazenados para servirem de base para a estimação dos parâmetros, assim como os valores verdadeiros dos parâmetros dos itens foram armazenados para futura comparação (análise das estimativas).

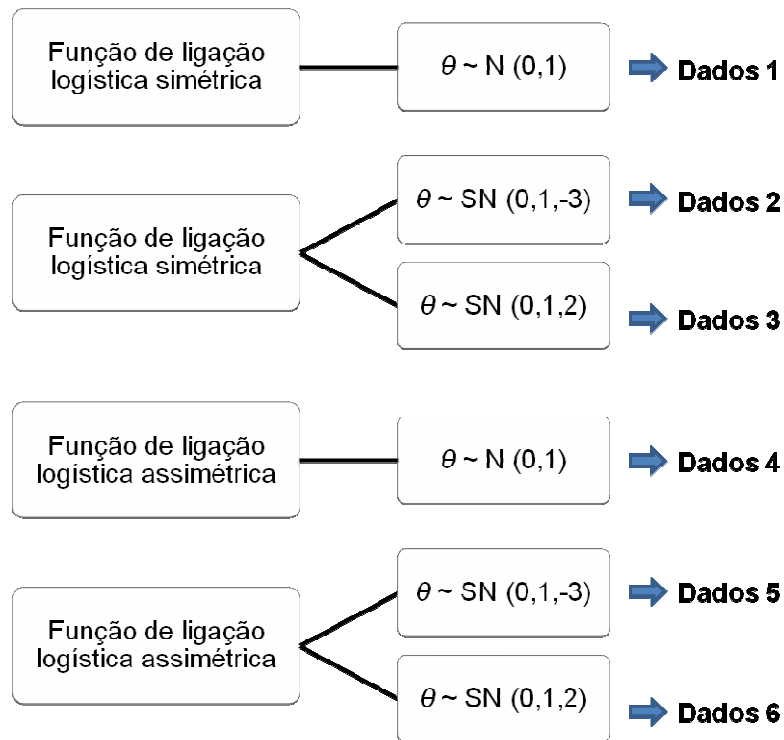


Figura 7 – Composição dos seis conjuntos de dados sintéticos gerados para análise.

3.2. Estimação dos parâmetros do modelo

Nessa etapa, consideramos conhecida apenas a matriz $Y_{(n \times I)}$ das respostas referente ao desempenho dos indivíduos em cada item. Utilizamos então o procedimento de inferência para estimar todos os parâmetros do modelo. Essas inferências podem ser, por exemplo, a respeito da dificuldade do item.

A amostra de 500 indivíduos e um teste com 15 itens implica num total de 545 parâmetros a serem estimados simultaneamente para o ajuste utilizando o modelo simétrico e 560 parâmetros para o ajuste utilizando o modelo assimétrico.

Quando falamos em “ajuste simétrico”, estamos nos referindo ao modelo logístico simétrico (função de ligação simétrica) já quando falamos em “ajuste assimétrico”, utilizamos o modelo logístico assimétrico (função de ligação assimétrica).

De posse desses desempenhos, começamos o processo de inferência para os parâmetros dos itens e das habilidades. Cada matriz $Y_{(n \times I)}$ de dados é ajustada por dois modelos: TRI logístico simétrico de três parâmetros e TRI logístico assimétrico de três

parâmetros, com as distribuições a priori já especificadas na seção 2.3 e mostradas novamente na Tabela 2.

Tabela 2 – Prioris adotadas para o ajuste dos parâmetros de interesse

Parâmetros	a	b	c	θ	λ
Priori	$N(1,0.5)I(0,)$	$N(0,2)$	$B(2,8)$	$N(0,1)$	$G(0,25, 0,25)$

Como destacado anteriormente, temos $6 \times 20 = 120$ matrizes de desempenhos $Y_{(n \times I)}$ a serem analisados em busca de um bom ajuste, ou seja, de obter a estimativa mais próxima possível dos parâmetros verdadeiros. Mas, cada um desses desempenhos foi ajustado através de dois modelos diferentes, o que implica num total de 240 ($120 \times 2 = 240$) ajustes a serem analisados em nosso estudo.

As inferências sobre os parâmetros dos modelos da TRI foram feitas por meio dos algoritmos Metropolis-Hastings dentro da Amostragem Gibbs a partir de amostras geradas da distribuição posterior conjunta dos mesmos, utilizando três cadeias paralelas com 50000 iterações, num período *burn-in* (descartados) de 30000 iterações e um salto *thinning* igual a 20 entre iterações, obtendo-se então, para cada parâmetro, 1000 iterações de cada cadeia, totalizando 3000 valores.

Essas especificações foram definidas a partir de testes (algumas simulações) feitos antes de começar o processo de inferência, os quais foram analisados por meio de critérios apresentados em 2.4, para assegurar a convergência das cadeias de Markov e estacionariedade, que podem ser calculados utilizando o pacote CODA no software R.

A seguir, mostramos os testes utilizados; como exemplo, *Dados 1*.

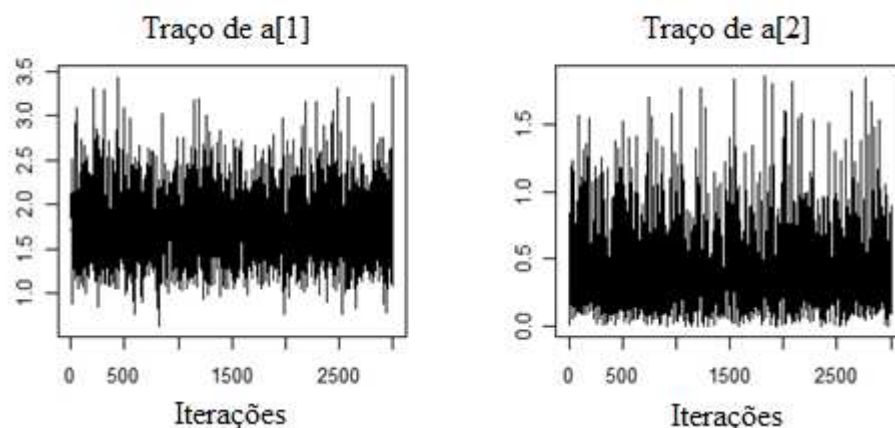


Figura 8 – Traço das cadeias para dois parâmetros, indicando convergência.

	mean	sd	2.5%	25%	50%	75%	97.5%	Rhat
a[1]	1.748	0.395	1.103	1.466	1.704	1.982	2.619	1.001
a[2]	0.417	0.283	0.052	0.221	0.352	0.536	1.146	1.002
a[3]	1.331	0.405	0.665	1.038	1.292	1.579	2.255	1.003
...								
a[15]	1.421	0.479	0.588	1.067	1.391	1.728	2.429	1.010
b[1]	-0.612	0.880	-2.258	-1.209	-0.633	-0.070	1.294	1.013
b[2]	0.716	1.344	-2.026	-0.196	0.596	1.638	3.375	1.034
b[3]	0.850	0.961	-1.044	0.218	0.873	1.484	2.674	1.029
....								
b[15]	1.004	0.999	-0.790	0.299	0.952	1.684	2.926	1.055
p.al[1]	0.193	0.108	0.031	0.110	0.181	0.261	0.437	1.001
p.al[2]	0.259	0.139	0.041	0.144	0.243	0.365	0.526	1.001
p.al[3]	0.161	0.055	0.041	0.125	0.167	0.202	0.253	1.016
...								
p.al[15]	0.233	0.056	0.091	0.206	0.244	0.272	0.317	1.002
lambda[1]	0.873	0.734	0.205	0.425	0.663	1.055	2.764	1.008
lambda[2]	1.852	2.204	0.244	0.561	1.061	2.148	8.276	1.030
lambda[3]	2.625	2.184	0.603	1.202	1.908	3.216	9.039	1.028
...								
lambda[15]	3.188	2.608	0.597	1.420	2.414	4.117	10.505	1.028
EQM	-0.024	0.045	-0.112	-0.054	-0.024	0.006	0.063	1.002
VM	0.707	0.046	0.623	0.675	0.704	0.737	0.802	1.001

Figura 9 – Diagnóstico Gelman e Rubin (*Rhat*) para alguns parâmetros do ajuste assimétrico para *Dados 1*, indicando convergência.

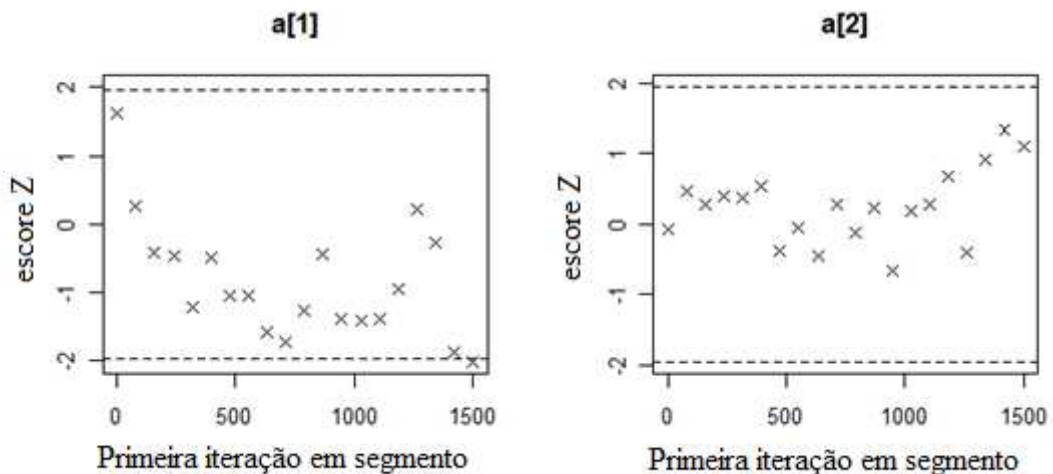


Figura 10 – Diagnóstico de Geweke para testar estacionariedade. Valores longe do intervalo de -2 a 2 indicam não estacionariedade.

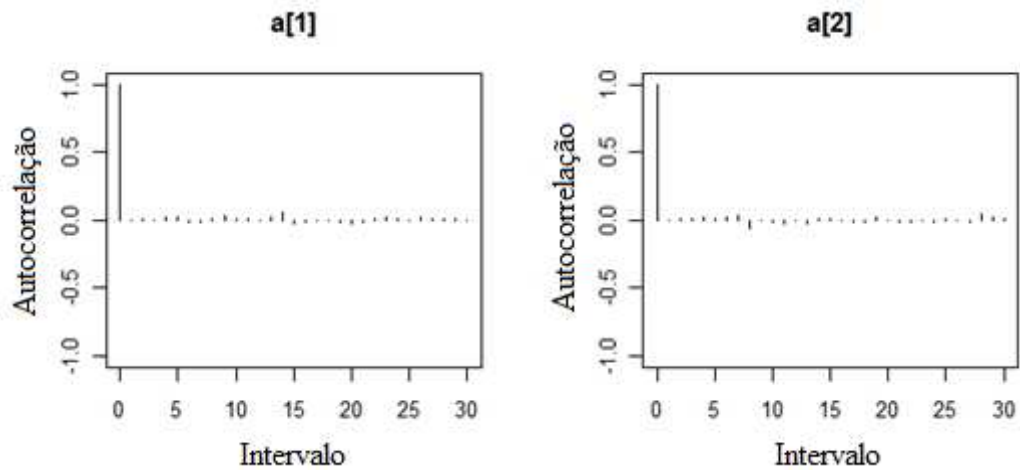


Figura 11 – Medidas de dependência entre as amostras (autocorrelação) de cadeias de Markov. Os gráficos indicam autocorrelação muito fraca.

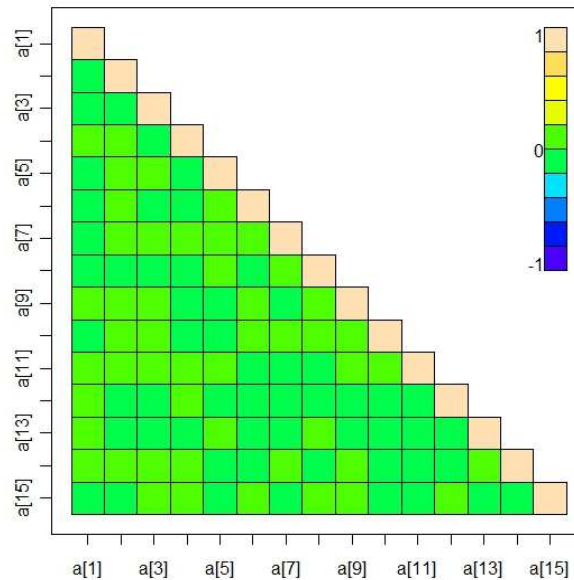


Figura 12 – Correlação entre as discriminações dos 15 itens. Como podemos observar a correlação é praticamente nula entre os parâmetros de discriminação.

Como maneira prática de avaliar a qualidade das inferências, optamos por trabalhar com as *diferenças das posteriores em relação ao valor verdadeiro* para cada parâmetro a_i , b_i , c_i , λ_i , $i = 1, \dots, I$, retendo-se as médias, medianas, desvios padrão, ICr 95% e amplitudes dessas diferenças. A Fig. 13 (ETAPA 2) resume como foram feitas as análises das distribuições posteriores.

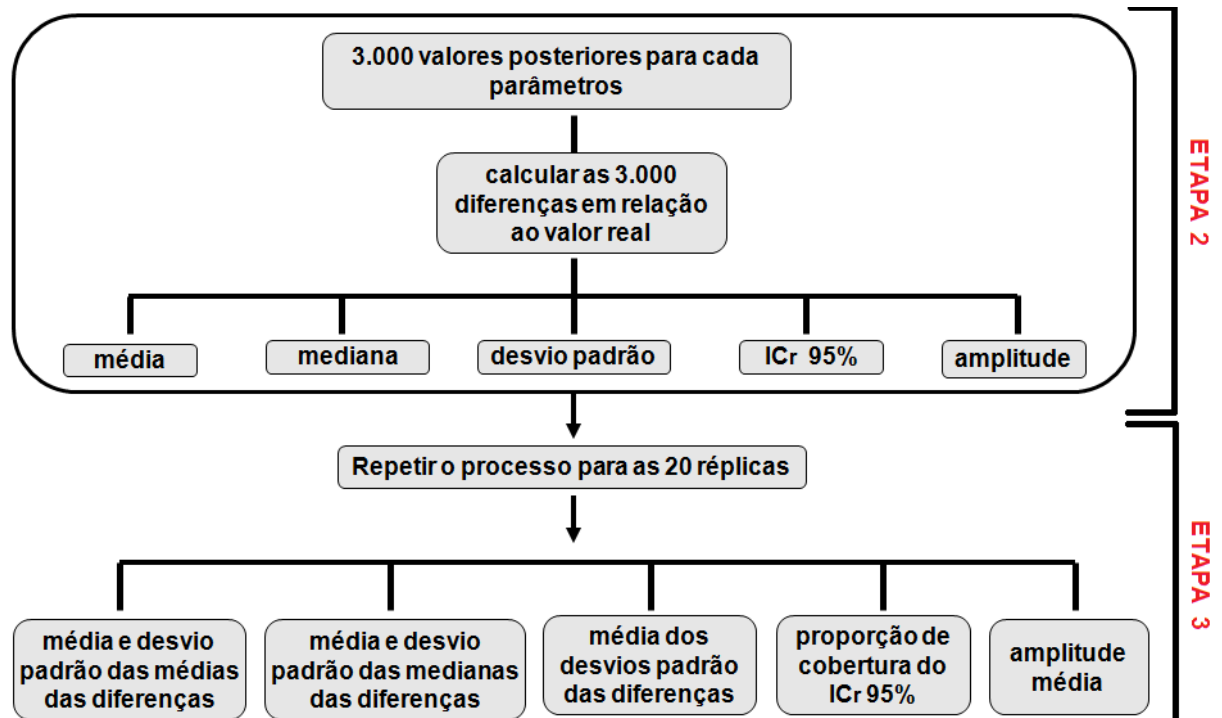


Figura 13 – Esquema ilustrando como foram calculadas as estatísticas das distribuições posteriores de cada parâmetro dos itens (ETAPA 2) e os sumários estatísticos para as 20 réplicas (ETAPA 3).

A Fig. 14 mostra a sequência do processo para validar os estudos e avaliar a qualidade das estimativas.

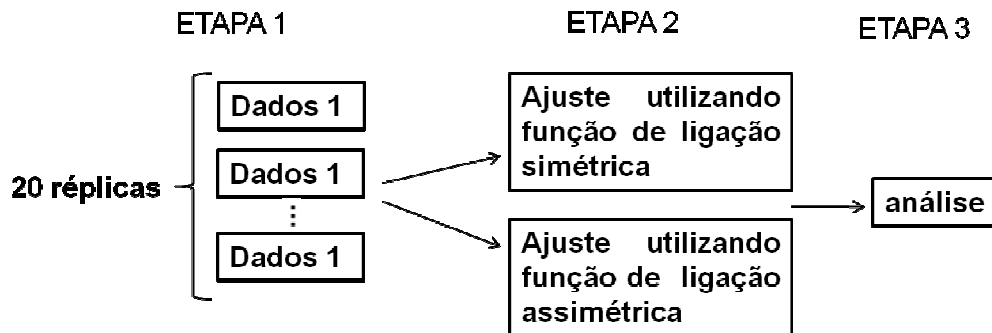


Figura 14 – Esquema das etapas para *Dados 1*, da mesma forma como é válido para os outros dados.

3.3. Análise da qualidade das estimações

Como o maior interesse era analisar o ajuste dos modelos em relação aos parâmetros dos itens, para os parâmetros θ_j guardamos apenas as estatísticas de interesse a seguir, que também nos mostram a qualidade das estimativas das habilidades.

Erro quadrático médio (EQM): É a soma dos quadrados da diferença entre os verdadeiros valores e os valores estimados do parâmetro θ_j , dado por:

$$EQM_r(\theta) = \frac{1}{n} \sum_{j=1}^n (\hat{\theta}_j^{(r)} - \theta_j^{(r)})^2$$

onde θ_j é o valor verdadeiro e $\hat{\theta}_j$ é a média da posterior para θ_j e r é a réplica.

Para as $R = 20$ réplicas, temos:

$$\overline{EQM(\theta)} = \frac{1}{R} \sum_{r=1}^R EQM_r(\theta)$$

Vício médio (VM): É definido como a soma das diferenças entre os verdadeiros valores e os valores estimados do parâmetro θ_j , dado por:

$$VM_r(\theta) = \frac{1}{n} \sum_{j=1}^n (\hat{\theta}_j^{(r)} - \theta_j^{(r)})$$

Assim como temos o viés médio de θ , calculado a partir das réplicas.

$$\overline{VM(\theta)} = \frac{1}{R} \sum_{r=1}^R VM_r(\theta)$$

Para avaliar a qualidade das inferências recorre-se aos sumários estatísticos sobre as 20 réplicas, como ilustrado na Fig. 13 (ETAPA 3).

Uma medida de análise da qualidade das estimativas é o viés padronizado (VP), que nos dá o valor residual em unidades de desvio padrão. Assim, podemos fazer uma comparação dos ajustes por cada modelo de forma mais correta. O VP é o resultado da divisão do resíduo pelo desvio padrão dos resíduos, dado por:

$$VP_a = \frac{(\hat{a}_i - a_i)}{\hat{\sigma}_i} \quad (3.1)$$

onde $(\hat{a}_i - a_i)$ são os resíduos dos parâmetros de discriminação para os I itens e $\hat{\sigma}_i$ é o desvio padrão dos resíduos associado ao item i . Da mesma forma como obtemos o VP para os parâmetros de discriminação a , obtemos para os demais parâmetros em estudo.

A análise do melhor ajuste (simétrico x assimétrico) para cada conjunto de dados foi feita através dos critérios de seleção dos modelos apresentados na seção 2.5; destacamos o DIC além de uma comparação entre os sumários estatísticos. Em cada simulação, guardamos também o tempo de execução de cada ajuste.

4. RESULTADOS

Para analisar a eficácia dos procedimentos de estimação, comparamos as estimativas de cada parâmetro com o seu valor verdadeiro. Para verificar qual ajuste foi mais eficiente, mostramos uma comparação entre o *ajuste simétrico* (com CCI logística simétrica) e o *ajuste assimétrico* (com CCI logística assimétrica) para cada conjunto de dados e o motivo que nos levou a escolher um ou outro modelo. Detalhes sobre todos os parâmetros podem ser encontrados no Apêndice 1. Nesse capítulo concentramos os esforços na análise dos padrões gerais que emergem das simulações.

É importante destacar que o parâmetro de assimetria da distribuição Normal Assimétrica (d) das habilidades não foi estimado. Apenas avaliamos seu efeito sobre o procedimento de estimação com ajustes simétricos e assimétricos.

4.1. Estimativas das habilidades θ_j

A Tabela 3 sintetiza a performance dos modelos de ajuste simétrico e assimétrico para cada conjunto de dados quanto a estimativa das habilidades.

Tabela 3 – Médias sobre as 20 réplicas dos \sqrt{EQM} e \overline{VM} das estimativas das habilidades obtidas com ajuste dos modelos com CCI simétrica e assimétrica para os seis tipos de dados

Dados	Ajuste Simétrico		Ajuste Assimétrico	
	\sqrt{EQM}	\overline{VM}	\sqrt{EQM}	\overline{VM}
<i>Dados 1</i>	0,8105	-0,023	0,8106	-0,023
<i>Dados 2</i>	1,195	0,714	1,195	0,716
<i>Dados 3</i>	1,112	-0,720	1,125	-0,730
<i>Dados 4</i>	0,841	-0,020	0,851	-0,022
<i>Dados 5</i>	1,214	0,717	1,214	0,716
<i>Dados 6</i>	1,13	-0,648	1,141	-0,657

Analisando a raiz quadrada do erro quadrático médio - \sqrt{EQM} - dos θ_j verificamos que para *Dados 1* e *Dados 4*, em que os modelos de ajuste simétrico e assimétrico são, respectivamente, os modelos corretos, se produziram os melhores ajustes.

Nos conjuntos de dados em que as habilidades ou a CCI são assimétricas, ambos modelos produzem ajustes piores (Fig. 15). No entanto, os dois modelos de ajuste dão resultados muito parecidos em termos tanto de \sqrt{EQM} quanto do viés médio \overline{VM} .

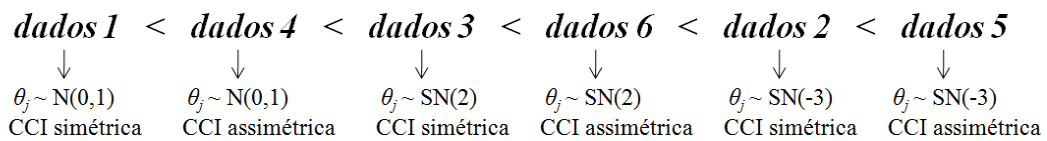


Figura 15 – Ordem de classificação dos dados em relação ao melhor ajuste para as habilidades (θ_j).

A análise dos \overline{VM} indicam que os modelos de ajuste tendem a produzir subestimativas das habilidades (viés médio negativo). A exceção ocorre com os dados gerados de uma distribuição Normal assimétrica fortemente negativa, SN(0, 1, -3).

4.2. Estimativas dos parâmetros dos itens

Para os seis conjuntos de dados, temos um total de 90 (15 x 6) parâmetros de discriminação (a), assim como de dificuldade (b) e de probabilidade de acertar um item ao acaso (c) nos modelos de ajuste simétrico. Para os modelos de ajuste assimétrico acrescentam-se a esses os 15 parâmetros de assimetria (λ).

Analisamos quantas dessas estimativas foram subestimadas e quantas foram sobrestimadas em relação ao uso da média e da mediana posterior como estimativas pontuais dos parâmetros. As Tabelas 4 e 5 mostram esses resultados em razão do viés padronizado (VP) definido em (3.1), que foram classificados da seguinte forma:

- (1) Subestimativas: quando a média ou mediana em questão se afasta mais que 1 desvio padrão para menos ($VP < -1$);
- (2) Sobrestimativas: quando a média ou mediana em questão se afasta mais que 1 desvio padrão para mais ($VP > 1$);
- (3) Adequadas: demais casos, ou seja, $-1 \leq VP \leq 1$.

Tabela 4 – Número médio de subestimativas (-), estimativas adequadas (0) e sobrestimativas (+) para ajuste simétrico e assimétrico referente aos seis conjuntos de dados para a média posterior

Dados	Ajuste Simétrico									Ajuste Assimétrico											
	<i>a</i>			<i>b</i>			<i>c</i>			<i>a</i>			<i>b</i>			<i>c</i>			λ		
	-	0	+	-	0	+	-	0	+	-	0	+	-	0	+	-	0	+	-	0	+
<i>Dados 1</i>	1	14	0	3	12	0	2	13	0	1	14	0	3	12	0	0	15	0	0	15	0
<i>Dados 2</i>	10	5	0	0	8	7	3	12	0	8	7	0	1	11	3	0	14	1	0	14	1
<i>Dados 3</i>	6	9	0	13	2	0	0	15	0	1	14	0	3	12	0	0	15	0	1	14	0
<i>Dados 4</i>	3	12	0	6	7	2	1	14	0	2	13	0	2	13	0	0	15	0	0	15	0
<i>Dados 5</i>	13	2	0	3	5	7	2	13	0	9	6	0	1	12	2	0	15	0	0	14	1
<i>Dados 6</i>	8	7	0	10	4	1	1	14	0	1	14	0	4	11	0	0	15	0	1	14	0

A interpretação das Tabela 4 e 5 é feita da seguinte maneira: por exemplo, no ajuste simétrico para *Dados 1* (Tabela 4), sobre a média das 20 réplicas para cada parâmetro de discriminação obtemos 15 estimativas, das quais uma foi subestimada e 14 consideradas adequadas.

Quanto maior for o número de parâmetros na coluna das estimativas adequadas (0), melhor será a qualidade da inferência quanto ao viés.

Tabela 5 – Número médio de subestimativas (-), estimativas adequadas (0) e sobrestimativas (+) para ajuste simétrico e assimétrico referente aos seis conjuntos de dados para a mediana posterior

Dados	Ajuste Simétrico									Ajuste Assimétrico											
	<i>a</i>			<i>b</i>			<i>c</i>			<i>a</i>			<i>b</i>			<i>c</i>			λ		
	-	0	+	-	0	+	-	0	+	-	0	+	-	0	+	-	0	+	-	0	+
<i>Dados 1</i>	2	13	0	3	12	0	1	14	0	1	14	0	4	11	0	0	15	0	0	15	0
<i>Dados 2</i>	12	3	0	0	9	6	3	12	0	10	5	0	1	11	3	0	14	1	0	15	0
<i>Dados 3</i>	8	7	0	14	1	0	0	15	0	3	12	0	3	12	0	0	15	0	3	12	0
<i>Dados 4</i>	4	11	0	7	5	3	1	14	0	3	12	0	4	11	0	0	15	0	0	15	0
<i>Dados 5</i>	13	2	0	3	6	6	1	14	0	10	5	0	1	13	1	0	14	1	0	15	0
<i>Dados 6</i>	10	5	0	11	3	1	0	15	0	2	13	0	4	11	0	0	15	0	2	13	0

Verificamos que os dois ajustes tem suas melhores performances com *Dados 1* e *Dados 4*. Isso não surpreende pois se trata dos dois casos em que os modelos de ajuste simétrico e assimétrico são idênticos aos modelos que geraram os dados, respectivamente.

Para dados com assimetria somente nas habilidades (*Dados 2* e *Dados 3*) os ajustes por ambos os modelos é levemente superior aos casos em que se acrescenta também assimetria na CCI (*Dados 5* e *Dados 6*).

As estimativas dos parâmetros a e c no ajuste simétrico e a no ajuste assimétrico, quando com viés, sempre se apresentam como subestimativas. Em contraste, os parâmetros b e λ produzem tanto subestimativas quanto sobrestimativas. Além disso, verifica-se que ajustes assimétricos tendem a produzir um número maior de estimativas adequadas que os ajustes simétricos. Isso se destaca mais fortemente nos parâmetros b que nos parâmetros a .

Utilizando ajuste simétrico, as estimativas de b não foram boas nos casos onde os dados verdadeiros tem algum tipo de assimetria (*Dados 2, 3, 4, 5 e 6*). Podemos verificar isso através das Tabelas A.13, A.23, A.33, A.43 e A.53 em Apêndice 1.

Quando os valores verdadeiros dos parâmetros dos itens são extremos das suas respectivas distribuições, as estimativas de alguns deles apresentam problemas quanto à proporção de cobertura dos ICr 95%, VP's, média das diferenças e média dos desvios elevados, como por exemplo, na maioria das estimativas de $a[3]$, $b[3]$ e $c[2]$ nos variados dados. Esses resultados foram obtidos analisando em cada tabela do Apêndice 1 (Tabelas A.1 a A.7; A.11 a A.17; A.21 a A.27; A.31 a A.37; A.41 a A.47 e A.51 a A.57) quais eram os parâmetros que possuíam os maiores valores para essas estatísticas e verificando se esses eram extremos na Tabela 1.

É inconclusivo com base nesse estudo se a média ou a mediana posterior devem servir como estimativas pontuais.

Intervalos de Credibilidade de 95% das estimativas

A Tabela 6 sintetiza as proporções de cobertura dos Intervalos de Credibilidade de 95% (ICr 95%) para as estimativas dos parâmetros dos itens. Esses, são satisfatórios sempre que essa cobertura está próxima do valor anunciado 95%. Para o total de 300 (15 x 20) intervalos gerados para os parâmetros a_i , verificamos a proporção que cobria os parâmetros verdadeiros. Similarmente obtivemos as proporções para os parâmetros b , c e λ .

Tabela 6 – Proporção média de cobertura dos ICr 95%

Dados	Ajuste Simétrico			Ajuste Assimétrico			
	<i>a</i>	<i>b</i>	<i>c</i>	<i>a</i>	<i>b</i>	<i>c</i>	λ
<i>Dados 1</i>	98,33%	97,33%	97,33%	98%	98,66%	100%	99,66%
<i>Dados 2</i>	73,66%	61,66%	98%	90%	99,66%	100%	99,33%
<i>Dados 3</i>	86,66%	63,66%	100%	99%	95,33%	100%	99%
<i>Dados 4</i>	97,33%	66%	97%	99%	98%	100%	99,33%
<i>Dados 5</i>	67,33%	55,33%	96%	86,33%	97,33%	99,33%	96%
<i>Dados 6</i>	82,66%	48%	100%	98,33%	94,66%	100%	98%

Considerando *Dados 1* tanto o ajuste simétrico (modelo correto) quanto o ajuste assimétrico produzem as coberturas anunciadas. Em *Dados 4* somente o modelo de ajuste assimétrico (modelo correto) produz as coberturas anunciadas. Para todos os demais casos os ajustes simétricos produziram coberturas insatisfatórias (inferiores às anunciadas) para *a* e *b*.

A Tabela 7 sintetiza a amplitude média dos intervalos de credibilidade de 95% dos parâmetros dos itens para os seis conjuntos de dados.

Tabela 7 – Amplitude média dos ICr 95%

Dados	Ajuste Simétrico			Ajuste Assimétrico			
	<i>a</i>	<i>b</i>	<i>c</i>	<i>a</i>	<i>b</i>	<i>c</i>	λ
<i>Dados 1</i>	1,271	1,694	0,339	1,437	4,113	0,358	6,140
<i>Dados 2</i>	1,347	2,409	0,335	1,546	4,465	0,357	8,663
<i>Dados 3</i>	1,259	1,752	0,442	1,518	4,556	0,431	4,709
<i>Dados 4</i>	1,294	1,737	0,357	1,510	4,222	0,351	5,723
<i>Dados 5</i>	1,300	2,255	0,346	1,594	4,524	0,350	8,277
<i>Dados 6</i>	1,283	1,769	0,422	1,619	4,536	0,409	4,664

Analisando a Tabela 7, podemos notar que o ajuste simétrico é mais preciso, produzindo amplitudes médias menores do que o ajuste assimétrico para a maioria dos parâmetros dos itens. Essa diferença é relevante principalmente para os parâmetros de dificuldade do item. Por isso, na Tabela 6, a proporção de cobertura dos ICr 95% para esses parâmetros foram significativamente melhores.

4.3. Escolha do melhor modelo

A Tabela 8 sintetiza os resultados do Critério de Informação de Deviância (DIC) utilizado para seleção do melhor ajuste (CCI simétrica ou CCI assimétrica). Seja \overline{DIC} = média DIC das 20 réplicas e \overline{pD} = média dos pD's das 20 réplicas.

Tabela 8 – Média dos valores DIC e pD utilizados na seleção do modelo mais adequado

Dados	Ajuste Simétrico		Ajuste Assimétrico	
	\overline{DIC}	\overline{pD}	\overline{DIC}	\overline{pD}
<i>Dados 1</i>	9061,9	381,275	8691,1	5,51
<i>Dados 2</i>	9638,85	216,325	9118,35	-277,49
<i>Dados 3</i>	8237,05	327,34	7798,65	-149,185
<i>Dados 4</i>	8213,35	365,82	7894,95	49,735
<i>Dados 5</i>	8677,75	204,03	8265,45	-170,99
<i>Dados 6</i>	7535,1	303,165	7165,25	-70,80

Podemos observar na Tabela 8 que em todos os casos o \overline{DIC} foi menor para o ajuste assimétrico. Inclusive para *Dados 1*, em que o modelo de ajuste simétrico é o verdadeiro modelo a partir do qual os dados foram gerados, o critério DIC privilegia o modelo assimétrico. Isso sugere que DIC não é um critério adequado nesse caso para selecionar o melhor modelo. Se, por exemplo, formos analisar somente a média das diferenças para cada parâmetro em *Dados 1*, a diferença do valor verdadeiro é menor utilizando o ajuste simétrico.

Valores negativos de pD também indicam que pode haver problemas no uso de DIC como critério de seleção de modelos nesse caso.

4.4. Tempo de simulação

Para cada conjunto de dados guardamos o tempo de simulação de cada ajuste, conforme Tabela 9. O computador utilizado para fazer as simulações possui a seguinte configuração: processador Core 2 Duo, 2GB de memória, 64 bits, 2,93 GHz e sistema operacional XP.

Tabela 9 – Tempo médio de execução (em minutos) das 20 réplicas para ajuste simétrico e assimétrico

Dados	Tempo (min)	
	Ajuste Simétrico	Ajuste Assimétrico
<i>Dados 1</i>	127,48	318,74
<i>Dados 2</i>	188,29	319,2
<i>Dados 3</i>	142,41	311,52
<i>Dados 4</i>	105,77	355,29
<i>Dados 5</i>	199,16	327,46
<i>Dados 6</i>	128,37	319,91

Verificamos na Tabela 9 que o tempo de simulação para o ajuste assimétrico é bem mais elevado do que para o ajuste simétrico; na maioria dos casos, mais que o dobro do tempo. Nos surpreende tanta diferença no tempo considerando que são 15 parâmetros a mais no ajuste assimétrico. Entretanto, não sabemos como o OpenBUGS é estruturado internamente para resolver essa alteração no modelo, o que pode ser um motivo para esse aumento no tempo.

O maior tempo de simulação para o ajuste simétrico ocorreu para *Dados 5* e, para o assimétrico, em *Dados 4*. Isso implica que o ajuste simétrico foi mais lento quando havia assimetria nos itens (negativa) e nas habilidades e, para o ajuste assimétrico, o curioso é que o ajuste mais lento ocorreu justamente quando se tratava do mesmo modelo que havia gerado os dados.

5. CONCLUSÕES

A suposição de normalidade está presente ao se assumir que as habilidades são “bem comportadas”, ou seja, um comportamento idealizado das variáveis para descrever a conduta humana. Essa suposição de normalidade está presente também no uso de ligações simétricas como o probito e logito, em funções de resposta dos itens que modelam a probabilidade de resposta correta em função da habilidade e de parâmetros associados aos itens (Bazán, 2005). Porém, essa suposição pode não ser a mais adequada para alguns casos, principalmente em relação à inteligência dos indivíduos.

O presente trabalho teve por objetivo estimar os parâmetros dos itens e das habilidades de indivíduos utilizando modelos para itens dicotômicos de três parâmetros da TRI e estudar o comportamento do ajuste quando existe assimetria na distribuição da habilidade de indivíduos e/ou na função de ligação, além de implementar a geração de dados sintéticos e o método de inferência bayesiana, mediante o uso do software livre R.

As conclusões obtidas com esse estudo foram:

- Os modelos de ajuste tendem a produzir subestimativas das habilidades, exceto nos casos em que essas tem distribuição normal assimétrica negativa. Além disso, quando avaliados em termos do erro quadrático ou do viés nas estimativas das habilidades, os dois modelos de ajuste produzem resultados muito próximos;
- As piores estimativas dos parâmetros são obtidas quando temos assimetria negativa na distribuição da habilidade ($\theta \sim SN(0,1,-3)$);
- A maioria das estimativas que apresentaram problemas em relação à proporção de cobertura dos ICr 95%, VP's, média das diferenças e média dos desvios elevados, seus valores verdadeiros eram extremos das respectivas distribuições;
- As estimativas de b_i não foram muito boas quando utilizado o ajuste simétrico para os casos onde os dados verdadeiros tem algum tipo de assimetria;
- O modelo de ajuste assimétrico é superior em termos da cobertura anunciada (95%) dos intervalos de credibilidade. Entretanto, produz amplitude média dos intervalos maior que o ajuste simétrico;
- Através das análises verificamos que a qualidade das estimativas são melhores utilizando o ajuste assimétrico, exceto quando a distribuição das habilidades e CCI são simétricas (*Dados 1*). Isso implica que, quando há alguma evidência de assimetria nos

dados, devemos preferencialmente utilizar o modelo assimétrico para fazer inferências.

Sugestões para trabalhos futuros:

- Testar outras prioris para verificar a influência nos resultados, a fim de reduzir estimativas de pontos discrepantes que prejudicam a qualidade das estimações finais;
- Testar outros métodos de estimação para comparação, como Estimação Bayesiana Marginal proposta por Mislevy em 1986, entre outros métodos existentes;
- Aumentar o tamanho da amostra e de itens, para conferir se há diferença nas estimativas, ou seja, melhoram as diferenças em relação aos seus valores verdadeiros, verificando cuidadosamente a composição de cada item em termos dos valores verdadeiros de seus parâmetros a fim de diminuir os erros;
- Desenvolver o esquema de dados aumentados para o modelo logístico, uma vez que, pelo que se tem conhecimento não há e, segundo a literatura, é utilizado essa estrutura para evitar trabalhar com a verossimilhança de Bernoulli e obter um modelo mais fácil do ponto de vista teórico;
- Considerar os hiperparâmetros desconhecidos e estimá-los, a fim de verificar se as estimativas serão melhores. Ou ainda, implementar a assimetria da habilidade no ajuste dos modelos, ou seja, não considerar a distribuição $N(0,1)$ como priori;
- Efetuar estudos desse tipo para diferentes modelos de TRI não dicotômicos.

REFERÊNCIAS BIBLIOGRÁFICAS

- Albert, J. *Bayesian Computation with R*. Verlag, New York, LLC: Springer, 2007.
- Alexandre, J.W.C. et al. Uma Proposta de Análise de um Construto para Medição dos Fatores Críticos da Gestão. *Gestão & Produção*, v. 9, n. 2, p. 129-141, ago. 2002.
- Andrade, D. F. et al. *Teoria da Resposta ao Item: Conceitos e Aplicações*, Caxambú-MG: 14 SINAPE, 2000.
- Azevedo, C. L. N. *Métodos de estimação na teoria de resposta ao item*, Dissertação de mestrado em Estatística, Instituto de Matemática e Estatística da Universidade de São Paulo, 2003.
- Azzalini, A. A Class of Distributions which Includes the Normal Ones. *Scandinavian Journal of Statistics*, v. 12, n. 2, p. 171-178, 1985.
- Azzalini, A. Further results on a class of distributions which includes the normal ones. *Statistica*, v. 46, p. 199-208, 1986.
- Azzalini, A. R package sn: The skew-normal and skew-t distributions (version 0.4-16). Università di Padova, Italia, 2010. Disponível em: <<http://azzalini.stat.unipd.it/SN>>
- Baker, F. B. *The Basics of Item Response Theory*. 2 ed. University of Wisconsin: ERIC, 2001.
- Bázan, J. L. G. *Uma família de modelos de resposta ao item normal assimétrica*. Tese de doutorado em Estatística. Universidade Federal de São Paulo, 2005.
- Bolfarine, H.; Bazán, J. L. (2008). *Bayesian estimation for Skewed Logistic IRT models*. Disponível em: www.ime.usp.br/~jbazan/download/articleBBSame.pdf. Acesso em: setembro de 2010.
- Carlin, B.P., Louis, T.A. *Bayes and Empirical Bayes Methods for Data Analysis*. London, Boca Raton: Chapman & Hall/CRC, 2000.
- Chen, M. H. et al. A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, v. 94, n. 448, p. 1172-1186, 1999.
- Ehlers, R. S. *Introdução à Inferência Bayesiana*. Laboratório de Estatística e Geoinformação, Universidade Federal do Paraná, 2003-2007.
- Gelfand, A. E.; Smith, A. F. M. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, v. 85, n. 410, p. 398 - 409, 1990.

Gelman, A. e Rubin, D. B. Inference from iterative simulation using multiple sequence. *Statistical Science*, v.7, n. 4, p. 457-511, 1992.

Gelman, A. et al. *Bayesian Data Analysis*. Texts in Statistical Science. 2. ed. London, Boca Raton: Chapman & Hall/CRC, 2004.

Geman, S. e Geman, D. Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, v. 6, n. 6, p. 721-741, 1984.

Geweke, J. Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments. In: Bernardo et al. *Bayesian Statistics*. v. 4, Oxford, UK: Clarendon Press, 1992.

Guewehr, K. *Teoria da Resposta ao Item na Avaliação da Qualidade de Vida de Idosos*. Dissertação de Mestrado, Faculdade de Medicina, Programa de Pós-Graduação em Epidemiologia da Universidade Federal do Rio Grande do Sul, 2007.

Hambleton, R.; Swaminathan, H.; Rogers, H. J. *Fundamentals of item response theory*. Newbury Park, California: Sage Publications, 1991. Measurement methods for the social sciences series. v. 2.

Hastings, W. K. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, v. 57, n. 1, p. 97-109, 1970.

Justino, G.; Andrade, D. F. *Software para avaliação de aprendizagem utilizando a teoria da resposta ao item*. XIII Workshop sobre Informática na Escola. Anais do XXVII Congresso da SBC, Rio de Janeiro – RJ, 2007.

Kinas, P. G.; Andrade, H. A. *Introdução à Análise Bayesiana (com R)*. Porto Alegre: maisQnada, 2010.

Lord, F. M. *A theory of test scores*. New York: Psychometric Society, 1952.

Marques, K. A. Análise bayesiana em modelos TRI de três parâmetros. Dissertação de Mestrado em Ciência. Universidade de São Paulo – SP, 2008.

Patz, J.R.; Junker, B.W. A Straightforward Approach to Markov Chain Monte Carlo Methods for Item Response Models. *Journal of Educational and Behavioral Statistics*, v. 24, n. 2, p. 146-178, 1999.

Paulino, C. D. et al. *Estatística Bayesiana*. Edição da Fundação Calouste Gulbenkian, Lisboa, 2003.

Pinheiro, A. N. C. *Modelos de Resposta ao Item com função de ligação t-Assimétrica*. Dissertação de Mestrado em Estatística. Universidade Federal de São Carlos – SP, 2007.

Pinheiro, I. R. et al. Modelo nominal da teoria de resposta ao item: uma alternativa. [online] *Avaliação Psicológica*, v. 9, n. 3, p. 437-447, 2010. ISSN 1677-0471.

Plummer, M. et al. CODA: Output analysis and diagnostics for MCMC. Version: 0.13-5, 2010.

Prentice, R. L. A Generalization of the probit and logit methods for dose-response curves. *Biometrika*, v. 32, n. 4, p. 761-768, 1976.

Rabelo, M. L. A Teoria de Resposta ao Item no Novo Enem. *Revista Explicando o Enem - Educar para as Competências*. São Paulo: Abril Educação, p. 65-67, 2009. Disponível em: www.ser.com.br/especiais/enem/Artigo_Mauro.pdf. Acesso em: jul. 2010.

Raftery, A. E.; Lewis, S. M. The Number of Iterations, Convergence Diagnostics and Generic Metropolis Algorithms. In: Gilks, W. R. et al. *Markov Chain Monte Carlo in Practice*. London, UK: Chapman & Hall, 1996.

R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2009. ISBN 3-900051-07-0. Disponível em: <http://www.rproject.org>.

Rodrigues, A. S. et al. *Convergência de Cadeias*. Universidade Federal de São Carlos, Centro de Ciências Exatas e de Tecnologia, Departamento de Estatística. São Carlos, Novembro de 2009.

Sahu, S. K. Bayesian Estimation and Model Choice in Item Response Models. *Journal of Statistical Computation and Simulation*, v. 72, p. 217-232, 2002.

Samejima, F. Departure from normal assumptions: A promise for future psychometrics with substantive mathematical modeling. *Psychometrika*, v. 62, n. 4, p. 471-493, 1997.

Samejima, F. Logistic positive exponent family of models: Virtue of asymmetric item characteristic curves. *Psychometrika*, v. 65, n. 3, p. 319-335, 2000.

Santos, V. L. F.; Gamerman, D.; Soares, T. M. *Metodologia Bayesiana para detecção de itens assimétricos em modelos da Teoria de Resposta do Item*. 1º Congresso Brasileiro de Teoria de Resposta ao Item (CONBRATRI), 2009b. **Anais...** Florianópolis, 9 a 11 de dezembro de 2009.

Santos, V. L. F. et al. *Teoria de Resposta ao Item: uma abordagem generalizada das Curvas Características dos Itens*. XI Escola de Modelos de Regressão, Recife, 1 e 4 de março de 2009. Disponível em: www.eventoexpress.com.br/cdemr/trabalhos/T81.PDF. Acesso em: junho de 2010.

SAS/STAT(R) 9.2 User's Guide, Second Edition. Introduction to Bayesian Analysis Procedures/Background in Bayesian Statistics. Disponível em:

http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#introba yes_toc.htm. Acesso em: 21/01/2011.

Soares, T. M. Utilização da teoria da resposta ao item na produção de indicadores sócio-econômicos. *Pesquisa Operacional*, v. 25, n.1, p. 83-112, 2005.

Spiegelhalter, D. J. et al. Bayesian measures of model complexity and fit. *Journal Royal Statistical Society, Series B (Statistical Methodology)*, v. 64, n. 4, p. 583-639, 2002.

Sturtz, S. et al. R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software*, v. 12, n. 3, p. 1-16, 2005.

Tezza, R.; Bornia, A. C. *Teoria da Resposta ao Item: vantagens e oportunidades para a engenharia de produção*. XXIX Encontro Nacional de Engenharia de Produção. A Engenharia de Produção e o Desenvolvimento Sustentável: Integrando Tecnologia e Gestão. Salvador, BA, Brasil, 06 a 09 de outubro de 2009.

Thomas, A. et al. Making BUGS Open. *R News*, v. 6, n. 1, p. 12-17, 2006.

APÊNDICE

APÊNDICE 1

A seguir, mostramos detalhadamente os resultados obtidos das análises das distribuições posteriores das 20 réplicas, para cada parâmetro, em cada situação.

DADOS 1 - habilidade $N(0,1)$ e CCI simétrica

➤ **Analisando as estimativas dos parâmetros a_i**

Tabela A. 1 – Ajuste utilizando modelo simétrico para a_i em *Dados 1*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,026	-0,058	0,346	2,528	-0,075
a[2]	0,141	0,090	0,278	2,157	0,507
a[3]	-0,667	-0,708	0,374	2,621	-1,783
a[4]	0,003	-0,038	0,323	2,352	0,009
a[5]	0,103	0,047	0,294	2,337	0,350
a[6]	0,019	-0,037	0,316	2,428	0,060
a[7]	-0,138	-0,185	0,336	2,393	-0,411
a[8]	-0,210	-0,258	0,312	2,346	-0,673
a[9]	0,087	0,057	0,263	2,193	0,331
a[10]	0,065	0,029	0,319	2,495	0,204
a[11]	-0,184	-0,228	0,360	2,636	-0,511
a[12]	-0,052	-0,108	0,322	2,451	-0,161
a[13]	-0,334	-0,381	0,372	2,551	-0,898
a[14]	0,039	-0,020	0,312	2,348	0,125
a[15]	-0,272	-0,312	0,378	2,554	-0,720

Verificamos na Tabela A.1 que 8 estimativas dos parâmetros de discriminação a_i foram subestimadas e 7 foram sobrestimadas. A maior diferença do valor verdadeiro foi para $a[3] = -0,667$, que também apresentou o maior desvio padrão (0,374) e maior viés padronizado (VP) de -1,783. Os demais parâmetros de discriminação tiveram VP abaixo de |0,9|. A maior amplitude foi encontrada em $a[11] = 2,636$. Para cada réplica calculamos o intervalo de credibilidade de 95% (ICr 95%), a fim de verificar a sua proporção de cobertura

para cada parâmetro. Desses, apenas a[3] teve proporção inferior a 95%, a saber, 85%. Destacamos que o valor verdadeiro de a[3], que apresentou maior problema nessa situação, é um valor extremo da sua verdadeira distribuição. Podemos concluir que as estimativas foram boas.

Tabela A. 2 – Ajuste utilizando modelo assimétrico para a_i em *Dados 1*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	0,056	0,018	0,409	3,038	0,137
a[2]	0,197	0,136	0,332	2,517	0,593
a[3]	-0,644	-0,681	0,386	2,675	-1,668
a[4]	0,053	0,014	0,365	2,548	0,145
a[5]	0,175	0,114	0,333	2,496	0,526
a[6]	0,125	0,077	0,341	2,550	0,367
a[7]	-0,091	-0,143	0,390	2,666	-0,233
a[8]	-0,078	-0,126	0,380	2,647	-0,205
a[9]	0,286	0,237	0,373	2,724	0,767
a[10]	0,123	0,081	0,381	2,715	0,323
a[11]	-0,115	-0,155	0,395	2,748	-0,291
a[12]	0,067	0,018	0,356	2,527	0,188
a[13]	-0,340	-0,385	0,393	2,695	-0,865
a[14]	0,161	0,108	0,350	2,573	0,460
a[15]	-0,256	-0,305	0,403	2,768	-0,635

Ao analisarmos a Tabela A.2, verificamos que 6 estimativas dos parâmetros de discriminação foram subestimadas e 9 sobrestimadas. A maior diferença do valor verdadeiro foi em $a[3] = -0,644$, assim como o maior VP = $-1,668$, as demais estimativas dos parâmetros de discriminação a_i tiveram VP abaixo de $|0,9|$; a[1] obteve o maior desvio padrão e amplitude, cujos valores são 0,409 e 3,038, respectivamente. Em relação à proporção de cobertura dos ICr 95%, apenas a[3] teve problemas; em 5 das 20 réplicas os intervalos não cobriram o valor verdadeiro (75% cobriram). Destacamos que os parâmetros a[3] e a[1] são valores extremos.

Concluimos que as estimativas foram melhores utilizando o ajuste simétrico, pois produziram menores diferenças dos valores verdadeiros, menores VP's, menores desvios padrão e amplitudes.

➤ Analisando as estimativas dos parâmetros b_i

Tabela A. 3 – Ajuste utilizando modelo simétrico para b_i em *Dados 1*

	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	
parâmetro	média	média	média dos desvios	média das amplitudes	VP
b[1]	0,092	0,075	0,297	2,357	0,310
b[2]	0,379	0,261	0,548	3,792	0,692
b[3]	-0,857	-0,908	0,516	3,435	-1,661
b[4]	-0,095	-0,143	0,370	2,634	-0,257
b[5]	0,004	-0,125	0,566	4,059	0,007
b[6]	0,013	-0,078	0,493	3,545	0,026
b[7]	-0,256	-0,334	0,544	3,671	-0,471
b[8]	-0,408	-0,472	0,384	2,705	-1,062
b[9]	0,060	0,005	0,334	2,649	0,180
b[10]	-0,071	-0,105	0,300	2,361	-0,237
b[11]	-0,182	-0,220	0,326	2,434	-0,558
b[12]	-0,104	-0,180	0,433	3,021	-0,240
b[13]	-0,552	-0,604	0,489	3,263	-1,129
b[14]	-0,061	-0,155	0,477	3,267	-0,128
b[15]	-0,409	-0,467	0,584	3,789	-0,700

Verificamos na Tabela A.3 que 10 estimativas dos parâmetros de dificuldade b_i foram subestimadas e 5 sobrestimadas. A maior diferença do valor verdadeiro ocorre em $b[3] = -0,857$, assim como o maior VP (-1,661). O maior desvio padrão foi de $b[15] = 0,584$; a maior amplitude foi encontrada em $b[5] = 4,059$. Em termos de proporção de cobertura dos ICr 95%, apenas $b[3] = 90\%$ e $b[8] = 80\%$ não apresentaram proporção satisfatória. Dos parâmetros comentados aqui, apenas o valor verdadeiro de $b[3]$ é um valor extremo de sua distribuição.

Tabela A. 4 – Ajuste utilizando modelo assimétrico para b_i em *Dados 1*

	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	
parâmetro	média	média	média dos desvios	média das amplitudes	VP
b[1]	0,378	0,394	0,924	5,821	0,409
b[2]	0,357	0,304	1,209	7,300	0,295
b[3]	-1,962	-2,003	0,998	6,451	-1,966

b[4]	-0,419	-0,445	1,013	6,353	-0,414
b[5]	0,165	0,135	1,223	7,428	0,135
b[6]	0,235	0,202	1,099	6,963	0,214
b[7]	-1,053	-1,102	1,067	6,707	-0,987
b[8]	-0,337	-0,356	1,084	6,630	-0,311
b[9]	0,930	0,933	1,098	6,892	0,847
b[10]	0,291	0,279	0,952	6,106	0,306
b[11]	-0,331	-0,331	0,938	5,996	-0,353
b[12]	0,224	0,209	1,116	6,853	0,201
b[13]	-1,273	-1,318	1,042	6,606	-1,222
b[14]	0,225	0,185	1,103	6,880	0,204
b[15]	-1,436	-1,484	1,012	6,626	-1,419

Verificamos na Tabela A.4 que 7 estimativas dos parâmetros de dificuldade foram subestimadas e 8 sobrestimadas. A maior diferença do valor verdadeiro foi em $b[3] = -1,962$, assim como o maior VP = - 1,966; o maior desvio padrão e amplitude são de $b[5]$, cujos valores são 1,223 e 7,428, respectivamente. Em relação à proporção de cobertura dos ICr 95%, apenas para $b[3]$ obteve 80%. Destacamos que $b[3]$ é um valor extremo.

Podemos notar que os desvios padrão são altos, assim como as amplitudes. Nesse caso, as estimativas foram piores que as encontradas pelo ajuste simétrico.

➤ **Analisando as estimativas dos parâmetros c_i**

Tabela A. 5 – Ajuste utilizando modelo simétrico para c_i em *Dados 1*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,031	0,021	0,094	0,518	0,330
c[2]	0,072	0,071	0,104	0,491	0,692
c[3]	-0,056	-0,052	0,050	0,287	-1,120
c[4]	-0,023	-0,024	0,075	0,405	-0,307
c[5]	-0,045	-0,048	0,122	0,568	-0,369
c[6]	-0,051	-0,050	0,115	0,553	-0,443
c[7]	-0,044	-0,040	0,061	0,322	-0,721
c[8]	-0,092	-0,093	0,094	0,476	-0,979
c[9]	0,030	0,016	0,124	0,626	0,242
c[10]	-0,010	-0,018	0,097	0,524	-0,103
c[11]	-0,055	-0,056	0,086	0,461	-0,640

c[12]	-0,048	-0,049	0,119	0,573	-0,403
c[13]	-0,066	-0,061	0,065	0,357	-1,015
c[14]	-0,049	-0,051	0,118	0,560	-0,415
c[15]	-0,040	-0,035	0,053	0,301	-0,755

Verificamos na Tabela A.5 que 12 estimativas dos parâmetros de probabilidade de acerto ao acaso foram subestimadas e 3 sobrestimadas. A maior diferença do valor verdadeiro ocorre em $c[8] = -0,092$; o maior desvio padrão e amplitude foram obtidos em $c[9]$, cujos valores são 0,124 e 0,626, respectivamente; o maior VP foi de $c[3] = -1,12$. Em relação à proporção de cobertura dos ICr 95% temos que, $c[3]$ obteve 85% e $c[15]$ 90%. Destacamos também que, dos parâmetros comentados aqui, apenas $c[8]$ é um valor extremo.

Tabela A. 6 – Ajuste utilizando modelo assimétrico para c_i em *Dados 1*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,028	0,015	0,098	0,543	0,286
c[2]	0,093	0,089	0,119	0,526	0,782
c[3]	-0,024	-0,019	0,051	0,309	-0,471
c[4]	-0,001	0,000	0,084	0,428	-0,012
c[5]	-0,058	-0,070	0,129	0,575	-0,450
c[6]	-0,070	-0,081	0,122	0,575	-0,574
c[7]	0,000	0,011	0,068	0,355	0,000
c[8]	-0,071	-0,070	0,104	0,508	-0,683
c[9]	0,014	-0,005	0,124	0,646	0,113
c[10]	-0,005	-0,017	0,105	0,548	-0,048
c[11]	-0,053	-0,056	0,091	0,478	-0,582
c[12]	-0,084	-0,096	0,117	0,566	-0,718
c[13]	-0,042	-0,035	0,071	0,384	-0,592
c[14]	-0,070	-0,082	0,122	0,577	-0,574
c[15]	-0,005	0,004	0,059	0,328	-0,085

Analisando a Tabela A.6, verificamos que 11 estimativas dos parâmetros de probabilidade de acerto ao acaso foram subestimadas e 3 foram sobrestimadas. A maior diferença do valor verdadeiro foi em $c[2] = -0,093$, que também apresentou maior VP de 0,782; quem obteve maior desvio padrão foi $c[5] = 0,129$ e maior amplitude foi $c[9] = 0,646$.

Todas estimativas apresentaram proporção de cobertura dos ICr 95% igual a 100%. Dos parâmetros comentados aqui, $c[2]$ e $c[5]$ são extremos.

Concluimos que as diferenças foram menores utilizando o ajuste assimétrico, assim como os VP's. Porém, o ajuste simétrico produziu amplitudes e desvios padrão menores.

Tabela A. 7 – Estimativas de λ_i pelo modelo assimétrico para *Dados 1*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
$\lambda[1]$	0,188	-0,176	1,129	11,795	0,167
$\lambda[2]$	0,794	0,151	1,883	19,882	0,422
$\lambda[3]$	2,264	1,510	2,577	25,071	0,879
$\lambda[4]$	1,006	0,413	1,843	18,325	0,546
$\lambda[5]$	0,535	-0,053	1,685	17,895	0,318
$\lambda[6]$	0,204	-0,197	1,209	13,604	0,169
$\lambda[7]$	1,820	1,070	2,410	23,854	0,755
$\lambda[8]$	0,680	0,119	1,675	16,443	0,406
$\lambda[9]$	-0,175	-0,486	0,984	11,614	-0,178
$\lambda[10]$	0,172	-0,176	1,100	11,208	0,156
$\lambda[11]$	0,621	0,146	1,480	15,209	0,420
$\lambda[12]$	0,215	-0,223	1,303	14,938	0,165
$\lambda[13]$	1,619	0,909	2,293	22,352	0,706
$\lambda[14]$	0,196	-0,208	1,234	13,457	0,159
$\lambda[15]$	2,101	1,352	2,509	23,350	0,837

Em *Dados 1*, o modelo que gerou os dados é simétrico, então, o valor verdadeiro de λ_i é 1. Analisando a Tabela A.7, verificamos que apenas uma estimativa dos parâmetros de assimetria foi subestimada e 14 foram sobrestimadas. A maior diferença do valor verdadeiro foi $\lambda[3] = 2,264$, que também obteve o maior desvio padrão (2,577), a maior amplitude (25,071) e o maior VP (0,879). Em relação à proporção de cobertura dos ICr 95%, todos apresentaram proporção satisfatória, ou seja, acima de 95%. Enfatizamos que $\lambda[3]$ não é um valor extremo.

Uma explicação para o modelo ter produzido algumas diferenças dos valores verdadeiros tão significativas pode ser devido à algumas estimativas discrepantes para esses parâmetros e ao fato de que, no modelo original dos dados, a assimetria é fixa e igual a 1.

Tabela A. 8 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos, utilizando o modelo simétrico para o ajuste de *Dados 1*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	-0,023	-0,023	0,044	0,303
EQM	0,657	0,655	0,043	0,306

Tabela A. 9 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando o modelo assimétrico para o ajuste de *Dados 1*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	-0,023	-0,023	0,044	0,313
EQM	0,666	0,665	0,044	0,308

Analisando as Tabelas A.8 e A.9, que resumem as distribuições posteriores dos 500 θ 's, verificamos que tanto o ajuste simétrico quanto o assimétrico produziram um viés negativo. O ajuste pelo modelo **simétrico** produziu menor EQM.

A Tabela A.10 mostra que, segundo os critérios DIC e pD, o ajuste assimétrico é preferido.

Tabela A. 10 - DIC's e pD's para *Dados 1*

Ajuste simétrico			Ajuste assimétrico								
DIC		pD	DIC		pD						
9113	9012	8936	378,1	382,3	397,2	8725	8730	8680	-6,11	37,4	-17,71
9031	9064	9105	392,6	390,1	388,3	8661	8706	8683	40,24	-18,8	-4,654
9010	9102	9049	375,8	371	377	8722	8656	8762	24,04	40,7	-27,46
9149	9096	9137	374,7	368,4	373,1	8683	8674	8599	-2,977	-3,25	-10,03
9052	9007	8952	378	386,7	395	8607	8748	8804	19,58	-2,89	-40,88
9079	9137	9137	385,5	367,2	367,2	8726	8599	8603	-26,4	18,08	45,75
9013	9057		387,9	389,4		8699	8755		17,13	28,34	
Média = 9061,9			Média = 381,275			Média = 8691,1			Média = 5,5047		
DP = 61,4747			DP = 9,4656			DP = 58,354			DP = 25,673		

DADOS 2 - habilidade N(0,1,-3) e CCI simétrica

➤ **Analisando as estimativas dos parâmetros a_i**

Tabela A. 11 – Ajuste utilizando modelo simétrico para a_i em *Dados 2*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,704	-0,753	0,383	2,854	-1,838
a[2]	0,067	-0,002	0,343	2,351	0,195
a[3]	-1,366	-1,430	0,347	2,226	-3,937
a[4]	-0,533	-0,590	0,359	2,501	-1,485
a[5]	-0,141	-0,215	0,319	2,361	-0,442
a[6]	-0,278	-0,342	0,336	2,429	-0,827
a[7]	-0,529	-0,593	0,342	2,273	-1,547
a[8]	-0,704	-0,768	0,351	2,489	-2,006
a[9]	-0,287	-0,335	0,311	2,461	-0,923
a[10]	-0,467	-0,524	0,376	2,837	-1,242
a[11]	-1,005	-1,060	0,356	2,529	-2,823
a[12]	-0,437	-0,495	0,332	2,464	-1,316
a[13]	-1,008	-1,073	0,343	2,261	-2,939
a[14]	-0,233	-0,304	0,360	2,617	-0,647
a[15]	-0,804	-0,867	0,354	2,261	-2,271

A Tabela A.11 mostra que 14 estimativas dos parâmetros de discriminação a_i foram subestimadas e 1 foi sobrestimada. A maior diferença do valor verdadeiro ocorre em $a[3] = -1,366$, que também obteve o maior VP (-3,937). O maior desvio padrão (0,383) e maior amplitude (2,854) foram encontrados em $a[1]$. Muitas estimativas apresentaram problemas em relação aos ICr 95% sobre as 20 réplicas nessa situação. A proporção de cobertura não foi satisfatória para $a[1]$, $a[3]$, $a[4]$, $a[7]$, $a[9]$, $a[10]$, $a[11]$, $a[12]$, $a[13]$ e $a[15]$. Dos parâmetros que apresentaram problemas, apenas $a[1]$, $a[3]$ e $a[11]$ são valores extremos de sua distribuição.

Podemos perceber que utilizando esse ajuste (simétrico) para os parâmetros de discriminação, as estimativas não foram muito boas, resultando em valores elevados para os VP's.

Tabela A. 12 – Ajuste utilizando modelo assimétrico para a_i em *Dados 2*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,529	-0,578	0,423	3,136	-1,251
a[2]	0,181	0,089	0,433	2,731	0,418
a[3]	-1,151	-1,211	0,449	2,702	-2,563
a[4]	-0,531	-0,600	0,399	2,751	-1,331
a[5]	-0,003	-0,084	0,368	2,564	-0,008
a[6]	-0,262	-0,340	0,381	2,673	-0,688
a[7]	-0,364	-0,433	0,434	2,715	-0,839
a[8]	-0,632	-0,717	0,423	2,819	-1,494
a[9]	-0,083	-0,135	0,368	2,662	-0,226
a[10]	-0,460	-0,513	0,390	2,797	-1,179
a[11]	-0,885	-0,948	0,403	2,820	-2,196
a[12]	-0,363	-0,433	0,373	2,797	-0,973
a[13]	-0,874	-0,951	0,444	2,711	-1,968
a[14]	-0,297	-0,374	0,358	2,608	-0,830
a[15]	-0,671	-0,737	0,449	2,573	-1,494

Na Tabela A.12 verificamos que 14 estimativas de a_i foram subestimadas e 1 foi sobrestimada. A maior diferença do valor verdadeiro foi para $a[3] = -1,151$, que também obteve o maior VP (-2,563). O maior desvio padrão foi de $a[15] = 0,449$ e maior amplitude foi encontrada para $a[1] = 3,136$. A proporção de cobertura dos ICr 95% considerados insatisfatórios foram para as estimativas dos parâmetros de $a[3]$, $a[11]$ e $a[13]$. Desses parâmetros, $a[1]$, $a[3]$ e $a[11]$ são valores extremos de sua verdadeira distribuição.

Concluimos que as estimativas foram melhores no ajuste assimétrico. Entretanto, produziram desvios padrão e amplitudes mais elevadas que no ajuste simétrico.

➤ **Analisando as estimativas dos parâmetros b_i**

Tabela A. 13 – Ajuste utilizando modelo simétrico para b_i em *Dados 2*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	1,597	1,518	0,417	3,236	3,830
b[2]	0,948	0,789	0,715	5,076	1,326

b[3]	-0,165	-0,278	0,744	4,970	-0,222
b[4]	0,978	0,856	0,645	4,323	1,516
b[5]	0,515	0,318	0,747	4,996	0,689
b[6]	0,594	0,433	0,644	4,365	0,922
b[7]	0,295	0,171	0,741	5,139	0,398
b[8]	0,578	0,436	0,668	4,544	0,865
b[9]	0,877	0,756	0,491	3,912	1,786
b[10]	1,227	1,124	0,486	3,648	2,525
b[11]	1,106	0,989	0,581	4,163	1,904
b[12]	0,685	0,539	0,593	4,217	1,155
b[13]	0,108	-0,029	0,726	4,653	0,149
b[14]	0,661	0,498	0,672	4,686	0,984
b[15]	0,058	-0,055	0,732	4,761	0,079

Analisando a Tabela A.13 verificamos que 1 estimativa dos parâmetros de dificuldade foi subestimada e 14 foram sobrestimadas. A maior diferença do valor verdadeiro é $b[1] = 1,597$, assim como o maior VP (3,83); o maior desvio padrão foi de $b[5] = 0,747$ e maior amplitude foi de $b[7] = 5,139$. Em termos de proporção de cobertura dos ICr 95% temos que, para $b[1]$, $b[9]$, $b[10]$ e $b[11]$ a proporção foi 0%. Outras estimativas que também apresentaram proporção insatisfatória foram $b[2]$, $b[4]$ e $b[12]$. Destacamos que apenas $b[1]$ é um valor extremo da distribuição.

Tabela A. 14 – Ajuste utilizando modelo assimétrico para b_i em *Dados 2*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	1,410	1,372	1,048	6,418	1,345
b[2]	0,648	0,601	1,138	7,064	0,569
b[3]	-1,275	-1,319	1,090	7,263	-1,170
b[4]	0,272	0,236	1,149	7,267	0,237
b[5]	0,456	0,416	1,220	7,800	0,374
b[6]	0,583	0,559	1,216	7,619	0,479
b[7]	-0,628	-0,666	1,110	7,250	-0,566
b[8]	0,248	0,208	1,178	7,558	0,211
b[9]	1,372	1,349	1,165	7,132	1,178
b[10]	1,263	1,235	1,111	6,854	1,137
b[11]	0,624	0,588	1,145	7,055	0,545
b[12]	0,597	0,542	1,224	7,343	0,488
b[13]	-0,620	-0,646	1,120	7,193	-0,554

b[14]	0,598	0,536	1,237	7,449	0,483
b[15]	-0,916	-0,950	1,101	7,374	-0,832

A Tabela A.14 nos mostra que 4 estimativas dos parâmetros de dificuldade foram subestimadas e 11 sobrestimadas. A maior diferença do valor verdadeiro foi para $b[1] = 1,41$, assim como o maior VP = 1,345; o maior desvio padrão foi de $b[14] = 1,237$ e maior amplitude foi $b[5] = 7,8$. Quanto à proporção de cobertura dos ICr 95%, todas as estimativas apresentaram desempenhos satisfatórios.

As estimativas não foram boas para os parâmetros de dificuldade nos dois ajustes. Porém, no ajuste assimétrico as diferenças e VP's foram menores para a maioria das estimativas.

➤ **Analisando as estimativas dos parâmetros c_i**

Tabela A. 15 – Ajuste utilizando modelo simétrico para c_i em Dados 2

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,050	0,044	0,101	0,513	0,495
c[2]	0,076	0,081	0,097	0,450	0,784
c[3]	-0,068	-0,063	0,049	0,256	-1,388
c[4]	-0,022	-0,019	0,075	0,369	-0,293
c[5]	-0,040	-0,040	0,123	0,538	-0,325
c[6]	-0,059	-0,061	0,115	0,524	-0,513
c[7]	-0,046	0,040	0,059	0,299	-0,780
c[8]	-0,076	-0,071	0,098	0,463	-0,776
c[9]	0,046	0,032	0,133	0,632	0,346
c[10]	0,015	0,011	0,104	0,515	0,144
c[11]	-0,049	-0,048	0,090	0,437	-0,544
c[12]	-0,065	-0,068	0,114	0,534	-0,570
c[13]	-0,080	-0,074	0,068	0,335	-1,176
c[14]	-0,051	-0,050	0,116	0,524	-0,440
c[15]	-0,061	-0,056	0,052	0,268	-1,173

Na Tabela A.15 percebemos que 11 estimativas dos parâmetros de probabilidade de acerto ao acaso foram subestimadas e 4 sobrestimadas. A maior diferença do valor verdadeiro foi encontrada em $c[13] = -0,08$; o maior VP foi de $c[3] = -1,388$; o maior desvio padrão e

amplitude foram de c[9], cujos valores são 0,133 e 0,632, respectivamente. Apenas c[3] teve desempenho insatisfatório em relação à proporção de cobertura dos ICr 95%.

Tabela A. 16 – Ajuste utilizando modelo assimétrico para c_i em *Dados 2*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,034	0,023	0,100	0,525	0,340
c[2]	0,121	0,141	0,112	0,480	1,080
c[3]	-0,020	-0,010	0,045	0,268	-0,444
c[4]	0,009	0,020	0,087	0,399	0,103
c[5]	-0,038	-0,046	0,137	0,561	-0,277
c[6]	-0,048	-0,058	0,132	0,552	-0,364
c[7]	-0,001	0,010	0,056	0,313	-0,018
c[8]	-0,051	-0,041	0,111	0,480	-0,459
c[9]	0,017	-0,003	0,125	0,637	0,136
c[10]	0,011	0,000	0,111	0,531	0,099
c[11]	-0,041	-0,041	0,098	0,459	-0,418
c[12]	-0,066	-0,076	0,125	0,551	-0,528
c[13]	-0,017	0,002	0,072	0,359	-0,236
c[14]	-0,047	-0,060	0,137	0,571	-0,343
c[15]	-0,002	0,010	0,049	0,292	-0,041

Analisando a Tabela A.16, verificamos que 10 estimativas de c_i foram subestimadas e 5 sobrestimadas. A maior diferença do valor verdadeiro foi em c[2] = 0,121, que também apresentou maior VP (1,08); o maior desvio padrão encontramos em c[5] e c[14] = 0,137 e a maior amplitude em c[9] = 0,637. Todas as proporções de cobertura dos ICr 95% foram satisfatórias. Dessas estimativas destacadas aqui, c[2], c[5] e c[14] são extremos.

O ajuste assimétrico gerou diferenças e VP's menores para a maioria das estimativas de c_i .

Tabela A. 17 – Estimativas de λ_i pelo modelo assimétrico para *Dados 2*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
λ [1]	0,550	0,102	1,402	15,117	0,392
λ [2]	1,653	0,790	2,582	26,088	0,640

$\lambda[3]$	3,565	2,556	3,556	32,170	1,003
$\lambda[4]$	1,761	0,933	2,553	25,572	0,690
$\lambda[5]$	0,865	0,156	2,053	22,599	0,421
$\lambda[6]$	0,871	0,143	2,087	22,510	0,417
$\lambda[7]$	3,109	2,163	3,286	30,100	0,946
$\lambda[8]$	1,433	0,615	2,436	24,760	0,588
$\lambda[9]$	0,110	-0,291	1,188	13,194	0,093
$\lambda[10]$	0,517	0,016	1,504	17,041	0,344
$\lambda[11]$	1,247	0,508	2,222	22,287	0,561
$\lambda[12]$	0,799	0,135	1,941	21,112	0,412
$\lambda[13]$	2,807	1,852	3,203	29,583	0,876
$\lambda[14]$	0,778	0,105	1,973	21,489	0,394
$\lambda[15]$	3,604	2,579	3,617	32,896	0,996

A Tabela A.17 nos mostra que as 15 estimativas dos parâmetros de assimetria λ foram sobrestimadas. A maior diferença do valor verdadeiro foi para $\lambda[15] = 3,604$, que também obteve o maior desvio padrão (3,617) e maior amplitude (32,896). O maior VP foi $\lambda[3] = 1,003$. Não tivemos problemas em relação à proporção de cobertura dos ICr 95%. Lembramos que nesse caso, o valor verdadeiro de λ_i é fixo e igual a 1.

Notemos que as amplitudes e desvios padrão são elevados em todas as estimativas.

Tabela A. 18 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos, utilizando modelo simétrico para ajuste de *Dados 2*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	0,714	0,714	0,044	0,309
EQM	1,428	1,426	0,091	0,644

Tabela A. 19 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos, utilizando modelo assimétrico para ajuste de *Dados 2*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	0,716	0,716	0,045	0,326
EQM	1,428	1,426	0,091	0,678

As Tabelas A.18 e A.19 nos mostram que tanto o ajuste simétrico quanto o ajuste assimétrico produziram um viés positivo, no ajuste simétrico foi menor. Notamos que viés e EQM aumentam quando a distribuição verdadeira das habilidades é assimétrica.

A tabela A.20 nos mostra as estimativas de DIC e pD, as quais são menores para o ajuste assimétrico.

Tabela A. 20 - DIC's e pD's para Dados 2

Ajuste simétrico			Ajuste assimétrico								
DIC		pD	DIC		pD						
9614	9639	9563	184,5	258,1	255,3	9106	9036	9063	-312,3	-317,3	-273
9657	9563	9716	225	273,7	153,3	9091	9119	9040	-354,1	-268	-201,4
9601	9698	9675	190,2	219,7	175,8	9267	9144	9106	-198,3	-322,6	-272,5
9592	9690	9665	169,4	204,1	215,7	9110	9084	9223	-232,7	-285,9	-244,9
9578	9680	9700	209,9	214,1	218,9	9172	9048	9125	-246,4	-213,9	-278,2
9643	9699	9582	241,5	223,7	228,7	9106	9144	9138	-267,5	-278,1	-266,1
9675	9547		233,3	231,6		9132	9113		-443,3	-273,4	
Média = 9638,85 DP = 54,2337			Média = 216,325 DP = 30,539			Média = 9118,35 DP = 56,7898			Média = -277,495 DP = 55,6236		

DADOS 3 – habilidade $N(0,1,2)$ e CCI simétrica

➤ **Analisando as estimativas dos parâmetros a_i**

Tabela A. 21 – Ajuste utilizando modelo simétrico para a_i em Dados 3

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,744	-0,723	0,290	2,486	-2,566
a[2]	0,065	0,004	0,279	2,430	0,233
a[3]	-0,457	-0,485	0,369	2,720	-1,238
a[4]	-0,194	-0,230	0,294	2,583	-0,660
a[5]	-0,045	-0,074	0,241	2,424	-0,187
a[6]	-0,149	-0,175	0,270	2,470	-0,552
a[7]	-0,066	-0,098	0,330	2,647	-0,200
a[8]	-0,301	-0,341	0,320	2,759	-0,941
a[9]	-0,227	-0,243	0,233	2,150	-0,974
a[10]	-0,446	-0,442	0,261	2,501	-1,709

a[11]	-0,702	-0,692	0,273	2,556	-2,571
a[12]	-0,199	-0,224	0,261	2,617	-0,762
a[13]	-0,476	-0,486	0,315	2,606	-1,511
a[14]	-0,079	-0,139	0,278	2,504	-0,284
a[15]	-0,377	-0,407	0,327	2,675	-1,153

Analisando a Tabela A.21, verificamos que 14 estimativas de a_i foram subestimadas e 1 foi sobrestimada. A maior diferença do valor verdadeiro foi $a[1] = -0,744$; o maior desvio padrão foi de $a[3] = 0,369$; a maior amplitude encontramos em $a[8] = 2,759$ e o maior VP em $a[11] = -2,571$. Em termos de proporção de cobertura dos ICr 95%, algumas estimativas apresentaram desempenho insatisfatório, como: $a[1]$, $a[9]$, $a[10]$, $a[11]$, $a[12]$ e $a[13]$. Enfatizamos que apenas $a[1]$, $a[5]$ e $a[11]$ são valores extremos.

Tabela A. 22 – Ajuste utilizando modelo assimétrico para a_i em Dados 3

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,460	-0,507	0,467	3,294	-0,985
a[2]	0,167	0,103	0,336	2,628	0,497
a[3]	-0,510	-0,552	0,435	2,994	-1,172
a[4]	-0,115	-0,168	0,376	2,678	-0,306
a[5]	0,002	-0,053	0,311	2,423	0,006
a[6]	-0,035	-0,092	0,361	2,686	-0,097
a[7]	-0,087	-0,150	0,424	2,840	-0,205
a[8]	-0,208	-0,263	0,379	2,733	-0,549
a[9]	0,015	-0,036	0,393	2,806	0,038
a[10]	-0,207	-0,256	0,415	3,007	-0,499
a[11]	-0,382	-0,429	0,422	2,971	-0,905
a[12]	-0,069	-0,123	0,380	2,665	-0,182
a[13]	-0,290	-0,340	0,420	2,929	-0,690
a[14]	-0,040	-0,092	0,342	2,668	-0,117
a[15]	-0,278	-0,330	0,437	2,949	-0,636

Percebemos na Tabela A.22 que 14 estimativas dos parâmetros de a_i foram subestimadas e 1 foi sobrestimada; $a[3]$ apresentou a maior diferença do valor verdadeiro (-0,51) e, também, o maior VP (-1,172). O maior desvio padrão e amplitude foram de $a[1]$, cujos valores são 0,467 e 3,294, respectivamente. Em termos de proporção de cobertura dos

ICr 95%, apenas a[11] não apresentou desempenho satisfatório. Destacamos que os parâmetros a[1], a[3] e a[11] são valores extremos de sua distribuição.

Concluimos que as estimativas de a_i foram melhores utilizando o ajuste assimétrico, pois gerou diferenças e VP's mais baixos que o ajuste simétrico.

➤ **Analisando as estimativas dos parâmetros b_i**

Tabela A. 23 – Ajuste utilizando modelo simétrico para b_i em *Dados 3*

	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	
parâmetro	média	média	média dos desvios	média das amplitudes	VP
b[1]	-1,073	-1,087	0,370	3,336	-2,900
b[2]	0,131	-0,026	0,622	4,426	0,211
b[3]	-1,427	-1,480	0,472	3,130	-3,023
b[4]	-0,722	-0,799	0,419	3,118	-1,723
b[5]	-0,432	-0,581	0,552	4,326	-0,783
b[6]	-0,674	-0,783	0,460	3,456	-1,465
b[7]	-0,639	-0,713	0,547	3,748	-1,168
b[8]	-0,965	-1,062	0,465	3,406	-2,075
b[9]	-0,661	-0,708	0,340	3,099	-1,944
b[10]	-0,928	-0,964	0,343	3,160	-2,706
b[11]	-1,214	-1,267	0,350	2,959	-3,469
b[12]	-0,843	-0,925	0,402	3,472	-2,097
b[13]	-1,259	-1,340	0,460	3,210	-2,737
b[14]	-0,684	-0,791	0,463	3,781	-1,477
b[15]	-0,950	-1,035	0,572	3,742	-1,661

Verificamos na Tabela A.23 que 14 estimativas dos parâmetros de dificuldade foram subestimadas e 1 foi sobrestimada. A maior diferença do valor verdadeiro foi para a estimativa de $b[3] = -1,427$, o maior VP foi de $b[11] = -3,469$. O maior desvio padrão e amplitude foram de b[1], cujos valores são 0,622 e 4,426, respectivamente. Vários parâmetros apresentaram proporção de cobertura dos ICr 95% abaixo de 95%, como b[1], b[3], b[4], b[6], b[8], b[9], b[10], b[11], b[12], b[13] e b[15]. Dos parâmetros citados aqui, apenas b[1], b[3] e b[15] são valores extremos. Com isso podemos perceber que o ajuste não foi bom nesse caso e também gerou VP's elevados.

Tabela A. 24 – Ajuste utilizando modelo assimétrico para b_i em *Dados 3*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	0,878	0,893	1,201	7,996	0,731
b[2]	0,344	0,309	1,210	7,270	0,284
b[3]	-2,167	-2,209	1,006	6,586	-2,154
b[4]	-0,491	-0,522	1,131	6,999	-0,434
b[5]	0,256	0,239	1,271	7,613	0,201
b[6]	0,276	0,286	1,271	7,749	0,217
b[7]	-1,084	-1,133	1,133	7,263	-0,957
b[8]	-0,496	-0,514	1,164	7,148	-0,426
b[9]	1,027	1,017	1,256	8,110	0,818
b[10]	0,520	0,502	1,173	7,765	0,443
b[11]	-0,281	-0,269	1,119	7,141	-0,251
b[12]	0,248	0,252	1,236	7,781	0,201
b[13]	-1,412	-1,435	1,058	6,603	-1,335
b[14]	0,213	0,187	1,248	7,498	0,171
b[15]	-1,561	-1,594	1,081	6,799	-1,444

Analisando a Tabela A.24 verificamos que 7 estimativas de b_i foram subestimadas e 8 sobrestimadas. A maior diferença do valor verdadeiro foi para $b[3] = -2,167$, assim como o maior VP = - 2,154; o maior desvio padrão foi de $b[5]$ e $b[6] = 1,271$ e maior amplitude encontramos em $b[9] = 8,11$. Apenas $b[3]$ apresentou proporção de cobertura dos ICr 95% insatisfatória, sendo esse parâmetro um valor extremo de sua distribuição.

As estimativas foram melhores utilizando o ajuste assimétrico. Porém, os desvios padrão e amplitudes são elevados.

➤ **Analisando as estimativas dos parâmetros c_i**

Tabela A. 25 – Ajuste utilizando modelo simétrico para c_i em *Dados 3*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,098	0,077	0,142	0,773	0,690
c[2]	0,114	0,109	0,128	0,578	0,891
c[3]	-0,031	-0,028	0,061	0,341	-0,508

c[4]	0,036	0,032	0,113	0,559	0,319
c[5]	-0,037	-0,052	0,142	0,648	-0,261
c[6]	-0,034	-0,048	0,143	0,671	-0,238
c[7]	-0,004	0,001	0,076	0,396	-0,053
c[8]	-0,024	-0,029	0,134	0,637	-0,179
c[9]	0,049	0,027	0,144	0,761	0,340
c[10]	0,043	0,024	0,138	0,730	0,312
c[11]	-0,005	-0,020	0,133	0,677	-0,038
c[12]	-0,040	-0,056	0,141	0,696	-0,284
c[13]	-0,043	-0,042	0,091	0,462	-0,473
c[14]	-0,031	-0,045	0,143	0,679	-0,217
c[15]	-0,010	-0,005	0,074	0,380	-0,135

A Tabela A.25 nos mostra que 10 estimativas de c_i foram subestimadas e 5 sobrestimadas. A maior diferença do valor verdadeiro foi encontrada em $c[2] = 0,114$, que também apresentou maior VP de 0,891; o maior desvio padrão foi $c[9] = 0,144$ e maior amplitude $c[1] = 0,773$. Todas as estimativas tiveram proporção de cobertura dos ICr 95% de 100%. Dos parâmetros comentados aqui, $c[1]$ e $c[2]$ são extremos.

Tabela A. 26 – Ajuste utilizando modelo assimétrico para c_i em *Dados 3*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,055	0,035	0,123	0,725	0,447
c[2]	0,093	0,075	0,135	0,600	0,689
c[3]	-0,006	0,000	0,069	0,376	-0,087
c[4]	0,016	0,003	0,117	0,575	0,137
c[5]	-0,063	-0,087	0,141	0,666	-0,447
c[6]	-0,070	-0,093	0,137	0,694	-0,511
c[7]	0,023	0,033	0,086	0,418	0,267
c[8]	-0,081	-0,098	0,124	0,624	-0,653
c[9]	0,009	-0,011	0,125	0,740	0,072
c[10]	0,005	-0,016	0,123	0,723	0,041
c[11]	-0,038	-0,057	0,124	0,675	-0,306
c[12]	-0,082	-0,103	0,130	0,697	-0,631
c[13]	-0,022	-0,017	0,095	0,478	-0,232
c[14]	-0,076	-0,097	0,132	0,687	-0,576
c[15]	0,028	0,040	0,083	0,418	0,337

Na Tabela A.26 temos que 8 estimativas de c_i foram subestimadas e 7 sobrestimadas. A maior diferença ocorre em $c[2] = 0,093$, que também apresentou maior VP (0,689); o maior desvio padrão foi $c[5] = 0,141$ e maior amplitude $c[9] = 0,74$. Não encontramos problemas com a proporção de cobertura dos ICr 95% para nenhum dos parâmetros, tivemos 100% de cobertura. Dos parâmetros comentados aqui, $c[2]$ e $c[5]$ são extremos.

Podemos considerar que as estimativas foram boas nos dois ajustes. As diferenças em relação aos verdadeiros valores dos parâmetros e os VP's foram menores em 8 estimativas utilizando o ajuste assimétrico.

Tabela A. 27 – Estimativas de λ_i pelo modelo assimétrico para *Dados 3*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
$\lambda[1]$	-0,671	-0,814	0,464	6,988	-1,446
$\lambda[2]$	0,347	-0,148	1,468	17,185	0,236
$\lambda[3]$	1,647	0,950	2,275	23,226	0,724
$\lambda[4]$	0,272	-0,176	1,345	15,464	0,202
$\lambda[5]$	0,068	-0,361	1,265	15,199	0,054
$\lambda[6]$	-0,159	-0,524	1,061	13,499	-0,150
$\lambda[7]$	1,591	0,837	2,373	23,961	0,670
$\lambda[8]$	0,060	-0,347	1,215	13,416	0,049
$\lambda[9]$	-0,563	-0,753	0,603	7,404	-0,934
$\lambda[10]$	-0,511	-0,699	0,592	7,760	-0,863
$\lambda[11]$	-0,267	-0,552	0,875	10,367	-0,305
$\lambda[12]$	-0,287	-0,590	0,884	9,802	-0,325
$\lambda[13]$	0,999	0,388	1,892	19,538	0,528
$\lambda[14]$	-0,116	-0,468	1,018	12,600	-0,114
$\lambda[15]$	1,654	0,870	2,432	23,009	0,680

Ao analisarmos a Tabela A.27, verificamos que 7 estimativas dos parâmetros de assimetria foram subestimadas e 8 sobrestimadas. A maior diferença do valor verdadeiro foi para $\lambda[15] = 1,654$, que também obteve o maior desvio padrão (2,432); o maior VP foi $\lambda[1] = -1,446$ e a maior amplitude $\lambda[7]$. Em relação à proporção de cobertura dos ICr 95%, apenas $\lambda[1]$ apresentou desempenho inferior a 95%, teve 85% de cobertura. Destacamos que $\lambda[1]$ é um valor extremo.

Nesse caso, a função de ligação utilizada para gerar os dados é simétrica, então, o valor verdadeiro de $\lambda_i = 1$. Notemos que as amplitudes e os desvios padrão são elevados.

Tabela A. 28 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para *Dados 3*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	-0,720	-0,720	0,044	0,313
EQM	1,236	1,233	0,083	0,590

Tabela A. 29 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para *Dados 3*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	-0,730	-0,730	0,044	0,315
EQM	1,266	1,265	0,084	0,616

Analisando as Tabelas A.28 e A.29, verificamos que tanto o ajuste simétrico quanto o assimétrico produziram um viés negativo. O ajuste pelo modelo **simétrico** produziu menor EQM e viés.

A tabela A.30 mostra que o ajuste assimétrico foi melhor, pois produziu DIC menor.

Tabela A. 30 – DIC's e pD's para *Dados 3*

Ajuste simétrico			Ajuste assimétrico		
DIC	pD		DIC	pD	
8326 8259 8223	331,0 300,6 317,5		7908 7839 7848	-78,28 -59,08 -102,8	
8231 8265 8134	321,3 336,9 326,9		7369 7716 7899	-414,3 -131,9 -179,3	
8105 8308 8194	340,6 326,1 313,8		7899 7783 7669	-179,3 -111,8 -207,9	
8196 8275 8224	327,9 330,2 329,9		7912 7783 7721	-63,46 -176,7 -234,9	
8152 8169 8288	339,6 322,5 327,1		7749 7816 7835	-227,6 -135,8 -72,17	
8282 8267 8326	344,9 335,7 321,4		7789 8008 7794	-132,5 -67,82 -103,7	
8205 8312	348,1 304,8		7794 7842	-190,8 -113,6	
Média = 8237,05 DP = 64,719	Média = 327,34 DP = 12,259		Média = 7798,65 DP = 128,493	Média = -149,1855 DP = 83,4202	

DADOS 4 (habilidade simétrica e CCI assimétrica)

➤ **Analisando as estimativas dos parâmetros a_i**

Tabela A. 31 – Ajuste utilizando modelo simétrico para a_i em *Dados 4*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	0,201	0,168	0,383	2,795	0,525
a[2]	0,286	0,236	0,321	2,291	0,891
a[3]	-0,583	-0,623	0,377	2,602	-1,546
a[4]	-0,046	-0,091	0,353	2,540	-0,130
a[5]	0,147	0,089	0,308	2,356	0,477
a[6]	-0,225	-0,246	0,235	2,180	-0,957
a[7]	-0,123	-0,174	0,343	2,453	-0,359
a[8]	-0,133	-0,184	0,372	2,558	-0,358
a[9]	-0,223	-0,237	0,310	2,136	-0,719
a[10]	0,000	-0,041	0,340	2,626	0,000
a[11]	-0,279	-0,329	0,372	2,776	-0,750
a[12]	-0,193	-0,232	0,272	2,350	-0,710
a[13]	-0,467	-0,510	0,383	2,665	-1,219
a[14]	-0,039	-0,069	0,256	2,247	-0,152
a[15]	-0,384	-0,429	0,380	2,522	-1,011

Na Tabela A.31 notamos que 11 estimativas de discriminação foram subestimadas e 3 sobrestimadas. A maior diferença do valor verdadeiro foi para $a[3] = -0,583$, que também apresentou maior VP (-1,546); o maior desvio padrão e amplitude ocorrem em $a[1]$, tais valores são 0,383 e 2,795, respectivamente. Somente $a[3]$ e $a[6]$ tiveram proporção de cobertura dos ICr 95% inferior a 95%. Destacamos também, que os parâmetros $a[1]$ e $a[3]$ são valores extremos.

Tabela A. 32 – Ajuste utilizando modelo assimétrico para a_i em *Dados 4*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	0,024	-0,013	0,404	2,785	0,059
a[2]	0,288	0,227	0,361	2,503	0,798

a[3]	-0,625	-0,668	0,400	2,778	-1,562
a[4]	0,023	-0,020	0,378	2,670	0,061
a[5]	0,203	0,140	0,368	2,612	0,552
a[6]	-0,013	-0,063	0,365	2,666	-0,036
a[7]	-0,143	-0,199	0,387	2,693	-0,370
a[8]	-0,080	-0,127	0,384	2,761	-0,208
a[9]	0,239	0,197	0,495	3,350	0,483
a[10]	0,088	0,045	0,398	2,833	0,221
a[11]	-0,138	-0,180	0,400	2,766	-0,345
a[12]	0,102	0,052	0,374	2,665	0,273
a[13]	-0,440	-0,484	0,417	2,792	-1,055
a[14]	0,122	0,074	0,361	2,533	0,338
a[15]	-0,358	-0,406	0,372	2,533	-0,962

A Tabela A.32 nos mostra que 7 estimativas de a_i foram subestimadas e 8 sobrestimadas. A maior diferença do valor verdadeiro ocorre em $a[3] = -0,625$, que também apresentou maior VP (-1,562); o maior desvio padrão e amplitude foram em $a[9]$, como 0,495 e 3,35, respectivamente. Quanto à proporção de cobertura dos ICr 95%, somente $a[3]$ teve desempenho inferior a 95%. Destacamos que o parâmetro $a[3]$ é um valor extremo.

Podemos considerar que as estimativas foram boas nos dois ajustes, mas o ajuste assimétrico foi um pouco superior, apresentando diferenças e VP's inferiores para a maioria das estimativas.

➤ **Analisando as estimativas dos parâmetros b_i**

Tabela A. 33 – Ajuste utilizando modelo simétrico para b_i em Dados 4

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	0,791	0,770	0,280	2,208	2,825
b[2]	1,126	1,031	0,569	3,754	1,979
b[3]	-1,122	-1,174	0,490	3,358	-2,290
b[4]	0,517	0,458	0,454	3,167	1,139
b[5]	0,448	0,332	0,566	3,917	0,792
b[6]	-1,690	-1,751	0,361	3,444	-4,681
b[7]	-0,460	-0,538	0,542	3,712	-0,849
b[8]	0,078	0,021	0,477	3,296	0,164
b[9]	-2,264	-2,264	0,394	3,277	-5,746

b[10]	-0,334	-0,375	0,343	2,688	-0,974
b[11]	-0,618	-0,664	0,359	2,631	-1,721
b[12]	-0,713	-0,793	0,391	3,159	-1,824
b[13]	-0,031	-0,087	0,566	3,744	-0,055
b[14]	-1,338	-1,408	0,375	3,291	-3,568
b[15]	-0,366	-0,427	0,616	3,974	-0,594

Através da Tabela A.33 verificamos que 10 estimativas dos parâmetros de dificuldade foram subestimadas e 5 sobrestimadas. A maior diferença do valor verdadeiro foi para $b[9] = -2,264$, assim como o maior VP = $-5,746$; outros parâmetros também apresentaram VP muito alto, tais como $b[6]$ e $b[14]$. O maior desvio padrão e amplitude média ocorrem em $b[15]$, cujos valores são 0,616 e 3,974, respectivamente. Vários parâmetros apresentaram problemas em relação à proporção de cobertura dos ICr 95%, como $b[1]$, $b[2]$, $b[3]$, $b[4]$, $b[8]$, $b[9]$, $b[10]$, $b[11]$, $b[12]$, $b[13]$ e $b[14]$. Dos parâmetros citados aqui, apenas $b[1]$, $b[3]$ e $b[15]$ são valores extremos de sua verdadeira distribuição. Contudo, podemos perceber que o ajuste não foi bom, pois muitas estimativas apresentaram problemas, com VP's elevados.

Tabela A. 34 – Ajuste utilizando modelo assimétrico para b_i em Dados 4

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	0,692	0,693	0,901	5,581	0,768
b[2]	0,462	0,408	1,121	6,893	0,412
b[3]	-2,048	-2,101	0,997	6,270	-2,054
b[4]	-0,200	-0,244	1,009	6,202	-0,198
b[5]	0,394	0,354	1,168	7,142	0,337
b[6]	0,117	0,117	1,255	8,050	0,093
b[7]	-1,082	-1,114	1,094	6,548	-0,989
b[8]	-0,298	-0,332	1,073	6,774	-0,278
b[9]	0,584	0,582	1,289	9,126	0,453
b[10]	0,274	0,279	1,000	6,511	0,274
b[11]	-0,423	-0,416	0,948	6,136	-0,446
b[12]	0,046	0,062	1,144	7,066	0,040
b[13]	-1,007	-1,046	1,036	6,702	-0,972
b[14]	0,107	0,121	1,207	7,597	0,089
b[15]	-1,576	-1,631	1,038	6,663	-1,518

Analisando a Tabela A.34 percebemos que 7 estimativas de b_i foram subestimadas e 8 sobrestimadas. A maior diferença do valor verdadeiro ocorre em $b[3] = -2,048$, assim como o maior VP = - 2,054; o maior desvio padrão e amplitude foram de $b[9]$, cujos valores são 1,289 e 9,126, respectivamente; apenas $b[3]$ teve 70% de proporção de cobertura para os ICr 95%. Destacamos que $b[3]$ é um valor extremo.

As estimativas de b_i não foram boas nos dois ajustes. Porém, foram melhores no ajuste assimétrico, mas produziram desvios padrão e amplitudes elevadas.

➤ **Analisando as estimativas dos parâmetros c_i**

Tabela A. 35 – Ajuste utilizando modelo simétrico para c_i em *Dados 4*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	-0,008	-0,015	0,072	0,411	-0,111
c[2]	0,025	0,028	0,072	0,362	0,347
c[3]	-0,043	-0,041	0,055	0,312	-0,782
c[4]	-0,031	-0,029	0,065	0,355	-0,477
c[5]	-0,073	-0,072	0,102	0,488	-0,716
c[6]	-0,045	-0,068	0,149	0,774	-0,302
c[7]	-0,028	-0,023	0,068	0,350	-0,412
c[8]	-0,070	-0,063	0,088	0,458	-0,795
c[9]	0,048	0,028	0,142	0,805	0,338
c[10]	0,036	0,027	0,119	0,613	0,303
c[11]	-0,011	-0,013	0,109	0,560	-0,101
c[12]	-0,047	-0,061	0,135	0,669	-0,348
c[13]	-0,085	-0,080	0,056	0,315	-1,518
c[14]	-0,043	-0,063	0,144	0,718	-0,299
c[15]	-0,035	-0,029	0,056	0,312	-0,625

Pela Tabela A.35 verificamos que 12 estimativas de c_i foram subestimadas e 3 sobrestimadas. A maior diferença ocorre na estimativa de $c[13] = -0,085$, que também apresentou maior VP de -1,518; o maior desvio padrão foi $c[6] = 0,149$ e maior amplitude foi $c[9] = 0,805$. Somente $c[13]$ teve desempenho inferior a 95% para a proporção de cobertura dos ICr 95%. Dos parâmetros comentados aqui, apenas $c[6]$ é um valor extremo.

Tabela A. 36 – Ajuste utilizando modelo assimétrico para c_i em *Dados 4*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,002	-0,005	0,080	0,448	0,025
c[2]	0,047	0,055	0,081	0,390	0,580
c[3]	-0,014	-0,007	0,061	0,339	-0,230
c[4]	-0,026	-0,021	0,067	0,355	-0,388
c[5]	-0,053	-0,050	0,112	0,495	-0,473
c[6]	-0,088	-0,109	0,127	0,755	-0,693
c[7]	0,011	0,022	0,080	0,392	0,137
c[8]	-0,080	-0,078	0,095	0,467	-0,842
c[9]	0,005	-0,016	0,123	0,748	0,041
c[10]	0,000	-0,017	0,114	0,611	0,000
c[11]	-0,052	-0,065	0,107	0,560	-0,486
c[12]	-0,086	-0,106	0,125	0,654	-0,688
c[13]	-0,044	-0,035	0,060	0,345	-0,733
c[14]	-0,085	-0,106	0,128	0,713	-0,664
c[15]	-0,038	-0,032	0,057	0,312	-0,667

A Tabela A.36 nos mostra que 10 estimativas dos parâmetros de probabilidade de acerto ao acaso foram subestimadas e 5 sobrestimadas. A maior diferença do valor verdadeiro foi para $c[6] = -0,088$, que também apresentou maior amplitude (0,755); o maior VP foi $c[8] = -0,842$; o maior desvio padrão foi $c[14] = 0,128$. A proporção de cobertura dos ICr 95% nesse caso foi de 100% para todas as estimativas. Todos parâmetros comentados são valores extremos da distribuição.

As diferenças e VP's foram menores em 8 estimativas de c_i utilizando o ajuste assimétrico. Podemos considerar que as estimativas foram razoáveis.

Tabela A. 37 – Estimativas de λ_i pelo modelo assimétrico para *Dados 4*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
$\lambda[1]$	-0,242	-0,693	1,419	15,062	-0,171
$\lambda[2]$	0,850	0,137	2,236	22,565	0,380
$\lambda[3]$	2,128	1,439	2,322	23,326	0,916
$\lambda[4]$	0,924	0,291	2,038	19,329	0,453

$\lambda[5]$	0,482	-0,182	1,981	20,178	0,243
$\lambda[6]$	0,112	-0,055	0,512	6,885	0,219
$\lambda[7]$	1,729	0,982	2,337	24,150	0,740
$\lambda[8]$	0,646	0,046	1,838	19,931	0,351
$\lambda[9]$	0,010	-0,046	0,197	3,749	0,051
$\lambda[10]$	0,192	-0,111	0,933	10,587	0,206
$\lambda[11]$	0,487	0,135	1,089	11,859	0,447
$\lambda[12]$	0,241	-0,083	0,967	11,128	0,249
$\lambda[13]$	1,519	0,747	2,603	24,614	0,584
$\lambda[14]$	0,154	-0,073	0,679	7,621	0,227
$\lambda[15]$	1,929	1,204	2,392	22,525	0,806

Percebemos na Tabela A.37 que 1 estimativa dos parâmetros de assimetria foi subestimada e 14 foram sobrestimadas; a maior diferença do valor verdadeiro foi para $\lambda[3] = 2,128$, que também obteve maior VP (0,916). O maior desvio padrão e amplitude foram de $\lambda[13]$. Em relação à proporção de cobertura dos ICr 95%, apenas $\lambda[3]$ teve menos de 95% de cobertura. Lembramos que $\lambda[13]$ é um valor extremo e, nesse caso, o modelo correto é com assimetria na CCI. As estimativas produziram desvios padrão e amplitudes elevadas.

Tabela A. 38 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para *Dados 4*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	-0,020	-0,020	0,044	0,310
EQM	0,707	0,706	0,046	0,337

Tabela A. 39 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para *Dados 4*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	-0,022	-0,023	0,045	0,313
EQM	0,724	0,722	0,048	0,343

Analisando as Tabelas A.38 e A.39 verificamos que tanto o ajuste simétrico quanto o assimétrico produziram um viés negativo. O ajuste pelo modelo **simétrico** produziu menor EQM e viés.

A Tabela A.40 indica que o melhor modelo a ser utilizado é o ajuste assimétrico, com DIC médio igual a 7894,95.

Tabela A. 40 – DIC's e pD's para Dados 4

Ajuste simétrico			Ajuste assimétrico								
DIC			pD								
DIC			pD								
8050	8246	8246	375,1	372,0	364,1	7760	7945	7836	94,44	-12,73	31,85
8211	8229	8339	364,1	355,0	368,2	7817	7899	7864	42,05	49,82	77,45
8218	8318	8251	362,8	367,8	356,5	7922	7921	7915	80,02	120,9	92,58
8193	8194	8243	367,0	366,6	359,0	7848	7868	7868	47,61	63,24	16,31
8126	8189	8183	363,6	376,9	361,0	7907	7896	7911	56,7	-35,56	64,69
8181	8158	8176	360,5	366,4	370,4	7891	8073	7981	-35,34	36,71	75,92
8202	8314		366,1	373,3		7903	7874		16,75	111,3	
Média = 8213,35			Média = 365,82			Média = 7894,95			Média = 49,7355		
DP = 66,419			DP = 5,8846			DP = 63,553			DP = 43,942		

DADOS 5 - habilidade $N(0,1,-3)$ e CCI assimétrica

➤ **Analisando as estimativas dos parâmetros a_i**

Tabela A. 41 – Ajuste utilizando modelo simétrico para a_i em Dados 5

	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	
parâmetro	média	média	média dos desvios	média das amplitudes	VP
a[1]	-0,561	-0,605	0,406	2,874	-1,382
a[2]	0,087	0,027	0,351	2,330	0,248
a[3]	-1,320	-1,385	0,349	2,296	-3,782
a[4]	-0,577	-0,631	0,360	2,476	-1,603
a[5]	-0,043	-0,116	0,345	2,405	-0,125
a[6]	-0,413	-0,444	0,253	2,209	-1,632
a[7]	-0,491	-0,549	0,352	2,342	-1,395
a[8]	-0,755	-0,826	0,362	2,536	-2,086
a[9]	-0,430	-0,459	0,308	2,010	-1,396
a[10]	-0,576	-0,631	0,368	2,875	-1,565

a[11]	-0,952	-1,014	0,378	2,751	-2,519
a[12]	-0,438	-0,487	0,313	2,563	-1,399
a[13]	-1,113	-1,186	0,337	2,213	-3,303
a[14]	-0,461	-0,497	0,248	2,283	-1,859
a[15]	-0,884	-0,950	0,338	2,188	-2,615

Através da Tabela A.41 podemos notar que 14 estimativas de a_i foram subestimadas e 1 foi sobrestimada. Encontramos a maior diferença do valor verdadeiro em $a[3] = -1,32$, que também apresentou maior VP (-3,782), $a[13]$ também obteve VP elevado (-3,303); $a[1]$ teve o maior desvio padrão (0,406) e maior amplitude (2,874). Muitos parâmetros apresentaram problemas em relação à proporção de cobertura dos ICr 95%. São eles: $a[1]$, $a[3]$, $a[4]$, $a[6]$, $a[8]$, $a[9]$, $a[10]$, $a[11]$, $a[12]$, $a[13]$, $a[14]$ e $a[15]$. Destacamos que, de todos esses parâmetros comentados, apenas $a[1]$, $a[3]$ e $a[11]$ são valores extremos da distribuição de a_i . Podemos perceber que as estimativas não foram muito boas para esse ajuste, com VP's elevados.

Tabela A. 42 – Ajuste utilizando modelo assimétrico para a_i em Dados 5

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,620	-0,670	0,436	2,971	-1,422
a[2]	0,223	0,153	0,442	2,690	0,505
a[3]	-1,150	-1,216	0,443	2,672	-2,596
a[4]	-0,502	-0,566	0,443	2,746	-1,133
a[5]	0,042	-0,060	0,429	2,672	0,098
a[6]	-0,109	-0,162	0,377	2,721	-0,289
a[7]	-0,387	-0,459	0,443	2,710	-0,874
a[8]	-0,657	-0,749	0,449	2,749	-1,463
a[9]	-0,148	-0,204	0,479	2,998	-0,309
a[10]	-0,420	-0,474	0,414	3,052	-1,014
a[11]	-0,887	-0,947	0,401	2,864	-2,212
a[12]	-0,420	-0,478	0,350	2,679	-1,200
a[13]	-0,943	-1,019	0,431	2,616	-2,188
a[14]	-0,269	-0,324	0,327	2,408	-0,823
a[15]	-0,588	-0,650	0,461	2,723	-1,275

Constatamos na Tabela A.42 que 13 estimativas dos parâmetros de discriminação foram subestimadas e 2 sobrestimadas. A maior diferença do valor verdadeiro foi para $a[3] = -1,15$, que também apresentou maior VP (-2,596); o maior desvio padrão foi de $a[9] = 0,479$ e

a maior amplitude foi encontrada em $a[10] = 3,052$. Em termos de proporção de cobertura dos ICr 95%, essa proporção não foi satisfatória para $a[1]$, $a[3]$, $a[10]$, $a[11]$, $a[12]$, $a[13]$ e $a[14]$. Desses, $a[1]$, $a[3]$ e $a[11]$ são valores extremos.

Percebemos que para alguns parâmetros as estimativas não foram muito boas. Porém, o ajuste assimétrico produziu resultados melhores que o ajuste simétrico.

➤ **Analisando as estimativas dos parâmetros b_i**

Tabela A. 43 – Ajuste utilizando modelo simétrico para b_i em *Dados 5*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	2,341	2,267	0,419	3,124	5,587
b[2]	1,604	1,475	0,704	4,727	2,278
b[3]	-0,419	-0,541	0,726	4,610	-0,577
b[4]	1,427	1,315	0,672	4,559	2,124
b[5]	0,998	0,827	0,751	4,952	1,329
b[6]	-1,151	-1,229	0,386	4,118	-2,982
b[7]	0,183	0,058	0,719	4,863	0,255
b[8]	0,945	0,796	0,732	4,769	1,291
b[9]	-1,700	-1,709	0,348	3,189	-4,885
b[10]	0,734	0,623	0,477	3,774	1,539
b[11]	0,610	0,478	0,560	3,872	1,089
b[12]	-0,072	-0,195	0,495	4,033	-0,145
b[13]	0,454	0,322	0,780	5,114	0,582
b[14]	-0,775	-0,882	0,450	4,445	-1,722
b[15]	0,057	-0,069	0,754	5,025	0,076

Analisando a Tabela A.43, verificamos que 5 estimativas de b_i foram subestimadas e 10 sobrestimadas. A maior diferença do valor verdadeiro ocorre em $b[1] = 2,341$, assim como o maior VP (5,587); o parâmetro $b[9]$ também apresentou VP elevado (-4,885). O maior desvio padrão e amplitude foram de $b[13]$, cujos valores são 0,78 e 5,114, respectivamente. Vários parâmetros apresentaram proporção de cobertura dos ICr 95% inferior a 95%, como: $b[1]$, $b[2]$, $b[4]$, $b[5]$, $b[6]$, $b[8]$, $b[9]$, $b[10]$ e $b[11]$. Dos parâmetros comentados aqui, apenas $b[1]$ é um valor extremo, indicando que o ajuste nesse caso não foi bom.

Tabela A. 44 – Ajuste utilizando modelo assimétrico para b_i em *Dados 5*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	1,714	1,652	1,027	6,795	1,669
b[2]	0,831	0,790	1,115	7,424	0,745
b[3]	-1,395	-1,421	1,112	7,282	-1,254
b[4]	0,524	0,467	1,097	7,157	0,478
b[5]	0,646	0,615	1,183	7,512	0,546
b[6]	0,208	0,201	1,253	7,626	0,166
b[7]	-0,611	-0,643	1,102	7,214	-0,554
b[8]	0,479	0,452	1,159	7,560	0,413
b[9]	1,067	1,086	1,393	9,188	0,766
b[10]	1,158	1,126	1,143	7,124	1,013
b[11]	0,529	0,466	1,143	7,039	0,463
b[12]	0,529	0,473	1,232	7,495	0,429
b[13]	-0,429	-0,460	1,125	7,477	-0,381
b[14]	0,282	0,258	1,273	7,930	0,222
b[15]	-0,965	-1,001	1,087	7,214	-0,888

Analisando a Tabela A.44, verificamos que 4 estimativas dos parâmetros de dificuldade foram subestimadas e 11 sobrestimadas. A maior diferença do valor verdadeiro foi em $b[1] = 1,714$, assim como o maior VP (1,669); $b[9]$ apresentou o maior desvio padrão (1,393) e maior amplitude (9,188). Em relação à proporção de cobertura dos ICr 95%, apenas $b[1]$ apresentou desempenho insatisfatório. Destacamos que $b[1]$ é um valor extremo.

As estimativas foram melhores utilizando o ajuste assimétrico, mas gerou desvios padrão e amplitudes elevadas.

➤ **Analisando as estimativas dos parâmetros c_i**

Tabela A. 45 – Ajuste utilizando modelo simétrico para c_i em *Dados 5*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	-0,004	-0,009	0,071	0,385	-0,056
c[2]	0,015	0,021	0,066	0,324	0,227
c[3]	-0,055	-0,049	0,056	0,283	-0,982

c[4]	-0,051	-0,047	0,059	0,303	-0,864
c[5]	-0,074	-0,068	0,097	0,448	-0,763
c[6]	-0,054	-0,075	0,144	0,756	-0,375
c[7]	-0,042	-0,037	0,062	0,311	-0,677
c[8]	-0,088	-0,081	0,088	0,410	-1,000
c[9]	0,042	0,021	0,140	0,793	0,300
c[10]	0,039	0,029	0,125	0,609	0,312
c[11]	-0,020	-0,024	0,109	0,514	-0,183
c[12]	-0,054	-0,068	0,133	0,628	-0,406
c[13]	-0,092	-0,085	0,059	0,296	-1,559
c[14]	-0,047	-0,070	0,146	0,729	-0,322
c[15]	-0,059	-0,053	0,053	0,272	-1,113

Percebemos, através da Tabela A.45, que 12 estimativas de c_i foram subestimadas e 3 sobrestimadas. A maior diferença do valor verdadeiro ocorre em $c[13] = -0,092$, que também apresentou maior VP (-1,559); o maior desvio padrão foi de $c[14] = 0,146$ e maior amplitude $c[9] = 0,793$; $c[8]$ obteve desempenho inferior a 95% em relação à proporção de cobertura dos ICr 95%. Destacamos que apenas o parâmetro $c[14]$ é um valor extremo.

Tabela A. 46 – Ajuste utilizando modelo simétrico para c_i em Dados 5

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,022	0,020	0,082	0,416	0,268
c[2]	0,063	0,079	0,071	0,352	0,887
c[3]	-0,011	0,001	0,053	0,294	-0,208
c[4]	-0,004	0,009	0,062	0,319	-0,065
c[5]	-0,025	0,002	0,113	0,472	-0,221
c[6]	-0,084	-0,106	0,131	0,736	-0,641
c[7]	0,019	0,036	0,063	0,332	0,302
c[8]	-0,032	-0,006	0,099	0,450	-0,323
c[9]	0,006	-0,014	0,123	0,741	0,049
c[10]	0,024	0,007	0,126	0,611	0,190
c[11]	-0,023	-0,033	0,119	0,552	-0,193
c[12]	-0,069	-0,093	0,138	0,672	-0,500
c[13]	-0,038	-0,024	0,054	0,315	-0,704
c[14]	-0,080	-0,104	0,134	0,729	-0,597
c[15]	-0,013	-0,002	0,047	0,284	-0,277

Ao analisarmos a Tabela A.46, notamos que 10 estimativas de c_i foram subestimadas e 5 sobrestimadas. Encontramos a maior diferença do valor verdadeiro em $c[6] = -0,084$ e o maior VP em $c[2] = 0,887$; o maior desvio padrão foi para $c[12] = 0,138$ e maior amplitude em $c[9] = 0,741$. Nenhum dos parâmetros apresentaram problemas quanto à proporção de cobertura dos ICr 95%. Destacamos que, dos parâmetro citados aqui, apenas $c[9]$ não é um valor extremo e que as estimativas utilizando esse ajuste foram boas.

Tabela A. 47 – Estimativas de λ_i pelo modelo assimétrico para *Dados 5*

	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	
parâmetro	média	média	média dos desvios	média das amplitudes	VP
$\lambda[1]$	0,460	-0,154	1,982	18,774	0,232
$\lambda[2]$	1,907	0,965	3,061	28,649	0,623
$\lambda[3]$	3,528	2,532	3,418	30,332	1,032
$\lambda[4]$	2,282	1,355	3,129	29,232	0,729
$\lambda[5]$	1,442	0,478	2,884	27,246	0,500
$\lambda[6]$	0,293	0,053	0,718	9,227	0,408
$\lambda[7]$	3,241	2,237	3,384	31,161	0,958
$\lambda[8]$	1,913	0,957	2,995	28,369	0,639
$\lambda[9]$	0,059	-0,029	0,293	4,589	0,201
$\lambda[10]$	0,461	0,055	1,226	13,733	0,376
$\lambda[11]$	0,936	0,418	1,609	17,536	0,582
$\lambda[12]$	0,456	0,065	1,188	14,603	0,384
$\lambda[13]$	2,875	1,837	3,620	32,464	0,794
$\lambda[14]$	0,373	0,053	0,953	12,031	0,391
$\lambda[15]$	3,509	2,503	3,547	30,081	0,989

Verificamos na Tabela A.47 que todas as estimativas dos parâmetros de assimetria foram sobrestimados e que a maior diferença do valor verdadeiro foi para $\lambda[3] = 3,528$, que também obteve o maior VP (1,03). O maior desvio padrão (3,62) e maior amplitude (32,46) foram de $\lambda[13]$. Quanto à proporção de cobertura dos ICr 95% somente $\lambda[3]$ teve proporção inferior a 95%. Podemos notar que as diferenças foram elevadas para alguns parâmetros, mas os VP's foram baixos. As amplitudes elevadas se deve a pontos discrepantes nas estimativas.

Tabela A. 48 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para *Dados 5*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	0,717	0,717	0,045	0,317
EQM	1,474	1,471	0,093	0,647

Tabela A. 49 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para *Dados 5*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	0,716	0,716	0,045	0,321
EQM	1,474	1,472	0,093	0,659

Analisando as Tabelas A.48 e A.49, podemos considerar que o ajuste pelo modelo **assimétrico** produziu menor viés e igual EQM ao ajuste simétrico; tanto o ajuste simétrico quanto o assimétrico produziram um viés positivo.

A Tabela A.50 mostra as estimativas de DIC e pD para a seleção do melhor ajuste, as quais são menores para o ajuste assimétrico.

Tabela A. 50 – DIC's e pD's para *Dados 5*

Ajuste simétrico			Ajuste assimétrico		
DIC	pD		DIC	pD	
8585 8683 8576	253,5	221,7 187	8282 8300 8338	-38,14	-23,94 -99,4
8643 8657 8588	194,5	190 227,2	8069 8252 8234	-220,5	-175,1 -235,2
8755 8652 8755	189,9	195,1 187,2	8263 8173 8324	-59,93	-174,2 -270,6
8746 8691 8697	227,4	174,5 200	8352 8235 8270	-180,4	-108,6 -297,8
8826 8774 8599	174,1	175,7 252,6	8369 8055 8325	-110,5	-360,7 -118,7
8682 8760 8540	166	202,5 219	8322 8391 8339	-245,3	-145,7 -153,7
8637 8709	197	245,7	8218 8198	-199,4	-202
Média = 8677,75	Média = 204,03		Média = 8265,45	Média = -170,991	
DP = 76,886	DP = 26,544		DP = 90,798	DP = 87,036	

DADOS 6 - habilidade N(0,1,2) e CCI assimétrica

➤ **Analisando as estimativas dos parâmetros a_i**

Tabela A. 51 – Ajuste utilizando modelo simétrico para a_i em *Dados 6*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,579	-0,615	0,346	2,694	-1,673
a[2]	0,156	0,090	0,362	2,483	0,431
a[3]	-0,628	-0,675	0,380	2,731	-1,653
a[4]	-0,033	-0,088	0,383	2,778	-0,086
a[5]	-0,011	-0,075	0,311	2,430	-0,035
a[6]	-0,352	-0,371	0,243	1,902	-1,449
a[7]	-0,077	-0,137	0,390	2,659	-0,197
a[8]	-0,324	-0,377	0,341	2,673	-0,950
a[9]	-0,387	-0,413	0,325	2,024	-1,191
a[10]	-0,516	-0,540	0,301	2,482	-1,714
a[11]	-0,778	-0,807	0,304	2,490	-2,559
a[12]	-0,425	-0,448	0,255	2,169	-1,667
a[13]	-0,366	-0,415	0,389	2,727	-0,941
a[14]	-0,272	-0,292	0,244	2,039	-1,115
a[15]	-0,254	-0,305	0,386	2,608	-0,658

A Tabela A.51 nos mostra que 14 estimativas dos parâmetros de discriminação foram subestimadas e 1 foi sobrestimada. Em a[11] encontramos a maior diferença do valor verdadeiro (-0,778) e, também, o maior VP (-2,559); o maior desvio padrão foi de a[7] = 0,39 e a maior amplitude de a[4] = 2,778. Vários parâmetros apresentaram problemas em relação à proporção de cobertura dos ICr 95%, como a[1], a[3], a[6], a[9], a[10], a[11], a[12], a[13] e a[14]. Dos parâmetros citados aqui, apenas a[1], a[3] e a[11] são valores extremos. Notamos que os VP's são razoáveis.

Tabela A. 52 – Ajuste utilizando modelo assimétrico para a_i em *Dados 6*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
a[1]	-0,337	-0,382	0,442	3,198	-0,762
a[2]	0,320	0,245	0,406	2,793	0,788
a[3]	-0,656	-0,706	0,402	2,863	-1,632
a[4]	-0,060	-0,109	0,407	2,877	-0,147
a[5]	0,087	0,006	0,387	2,666	0,225
a[6]	0,027	-0,025	0,417	2,882	0,065
a[7]	-0,160	-0,221	0,385	2,815	-0,416
a[8]	-0,211	-0,262	0,388	2,759	-0,544
a[9]	0,032	-0,022	0,546	3,363	0,059
a[10]	-0,269	-0,321	0,440	3,146	-0,611
a[11]	-0,411	-0,457	0,431	3,110	-0,954
a[12]	-0,001	-0,056	0,412	2,864	-0,002
a[13]	-0,332	-0,384	0,445	2,910	-0,746
a[14]	-0,072	-0,126	0,382	2,707	-0,188
a[15]	-0,280	-0,334	0,413	2,941	-0,678

Ao analisarmos a Tabela A.52 percebemos que 11 estimativas de a_i foram subestimadas e 4 sobrestimadas. A maior diferença do valor verdadeiro foi em $a[3] = -0,656$, que também apresentou maior VP (-1,632); o maior desvio padrão e amplitude foram encontrados em $a[9]$, os quais são 0,546 e 3,363, respectivamente. Quanto à proporção de cobertura dos ICr 95%, apenas $a[3]$ não obteve valor acima de 95%. Destacamos que o parâmetro $a[3]$ é um valor extremo de sua distribuição verdadeira.

Concluimos que as estimativas de a_i foram melhores utilizando o ajuste assimétrico.

➤ **Analisando as estimativas dos parâmetros b_i**

Tabela A. 53 – Ajuste utilizando modelo simétrico para b_i em *Dados 6*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	-0,289	-0,319	0,368	3,298	-0,785
b[2]	0,987	0,862	0,697	4,661	1,416
b[3]	-1,763	-1,831	0,470	3,354	-3,751

b[4]	-0,086	-0,155	0,458	3,246	-0,188
b[5]	0,257	0,100	0,639	4,353	0,402
b[6]	-2,100	-2,140	0,324	3,100	-6,481
b[7]	-0,666	-0,756	0,593	3,822	-1,123
b[8]	-0,651	-0,740	0,447	3,296	-1,456
b[9]	-2,695	-2,690	0,428	3,540	-6,297
b[10]	-1,181	-1,211	0,350	3,175	-3,374
b[11]	-1,656	-1,697	0,359	3,237	-4,613
b[12]	-1,423	-1,483	0,359	3,380	-3,964
b[13]	-0,474	-0,541	0,531	3,556	-0,893
b[14]	-1,802	-1,852	0,336	3,426	-5,363
b[15]	-0,825	-0,901	0,556	3,651	-1,484

Através da Tabela A.53 notamos que 13 estimativas dos parâmetros de dificuldade foram subestimadas e 2 sobrestimadas; b[9] apresentou a maior diferença do valor verdadeiro (-2,695) e b[2] o maior desvio padrão e amplitude, cujos valores são 0,697 e 4,661, respectivamente; b[6] teve o maior VP (-6,481). Outros parâmetros também apresentaram um VP bastante alto, como podemos ver na tabela. Além disso, vários parâmetros apresentaram problemas em relação à proporção de cobertura dos ICr 95%, como b[1], b[2], b[3], b[6], b[8], b[9], b[10], b[11], b[12] e b[14.] Dos parâmetros citados aqui, apenas b[1] e b[3] são valores extremos da verdadeira distribuição. Como podemos perceber, o ajuste não foi bom.

Tabela A. 54 – Ajuste utilizando modelo assimétrico para b_i em Dados 6

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
b[1]	0,758	0,761	1,107	7,228	0,685
b[2]	0,436	0,376	1,118	7,124	0,390
b[3]	-2,300	-2,332	1,041	6,510	-2,209
b[4]	-0,246	-0,298	1,087	6,738	-0,226
b[5]	0,401	0,374	1,208	7,315	0,332
b[6]	0,030	0,050	1,311	8,571	0,023
b[7]	-1,131	-1,174	1,110	6,941	-1,019
b[8]	-0,285	-0,306	1,119	6,925	-0,255
b[9]	0,885	0,887	1,402	9,544	0,631
b[10]	0,608	0,605	1,210	8,129	0,502
b[11]	-0,375	-0,363	1,136	7,230	-0,330
b[12]	0,085	0,090	1,212	7,667	0,070
b[13]	-1,094	-1,128	1,009	6,557	-1,084

b[14]	0,148	0,169	1,341	8,394	0,110
b[15]	-1,602	-1,639	1,037	6,637	-1,545

Verificando a Tabela A.54 percebemos que 7 estimativas de b_i foram subestimadas e 8 sobrestimadas. A maior diferença do valor verdadeiro encontramos em $b[3] = -2,3$, assim como o maior VP = -2,209; $b[9]$ obteve o maior desvio padrão e amplitude, cujos valores são 1,402 e 9,544, respectivamente. Apenas $b[3]$ teve proporção de cobertura dos ICr 95% inferior a 95%. Destacamos que $b[1]$ é um valor extremo.

As estimativas foram melhores utilizando o ajuste assimétrico. Porém, os desvios padrão e amplitudes foram elevados.

➤ **Analisando as estimativas dos parâmetros c_i**

Tabela A. 55 – Ajuste utilizando modelo simétrico para c_i em Dados 6

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,090	0,073	0,134	0,710	0,672
c[2]	0,083	0,092	0,095	0,448	0,874
c[3]	-0,010	-0,008	0,076	0,407	-0,132
c[4]	0,011	0,013	0,090	0,467	0,122
c[5]	-0,046	-0,048	0,122	0,550	-0,377
c[6]	-0,055	-0,077	0,143	0,800	-0,385
c[7]	0,014	0,021	0,083	0,423	0,169
c[8]	-0,060	-0,065	0,117	0,568	-0,513
c[9]	0,051	0,030	0,144	0,802	0,354
c[10]	0,048	0,028	0,142	0,756	0,338
c[11]	-0,006	-0,025	0,138	0,728	-0,043
c[12]	-0,050	-0,071	0,145	0,770	-0,345
c[13]	-0,069	-0,065	0,067	0,363	-1,030
c[14]	-0,054	-0,078	0,144	0,796	-0,375
c[15]	-0,026	-0,022	0,065	0,346	-0,400

Pela Tabela A.55 percebemos que 9 estimativas dos parâmetros de probabilidade de acerto ao acaso foram subestimadas e 6 sobrestimadas. A maior diferença do valor verdadeiro foi para $c[1] = 0,09$; $c[13]$ apresentou o maior VP (-1,03); $c[12]$ o maior desvio padrão (0,145) e $c[9]$ a maior amplitude (0,802). Nenhum dos parâmetros apresentaram problemas

em relação à proporção de cobertura dos ICr 95%, todos tiveram 100% de cobertura. Destacamos que, dos parâmetros citados aqui, c[1] e c[12] são valores extremos.

Tabela A. 56 – Ajuste utilizando modelo assimétrico para c_i em *Dados 6*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
c[1]	0,049	0,030	0,117	0,675	0,419
c[2]	0,084	0,093	0,102	0,452	0,824
c[3]	0,003	0,005	0,086	0,430	0,035
c[4]	0,024	0,024	0,103	0,496	0,233
c[5]	-0,031	-0,037	0,140	0,580	-0,221
c[6]	-0,090	-0,112	0,126	0,764	-0,714
c[7]	0,022	0,025	0,100	0,463	0,220
c[8]	-0,070	-0,081	0,121	0,579	-0,579
c[9]	0,006	-0,015	0,124	0,740	0,048
c[10]	0,007	-0,013	0,124	0,725	0,056
c[11]	-0,040	-0,060	0,125	0,707	-0,320
c[12]	-0,088	-0,108	0,127	0,727	-0,693
c[13]	-0,032	-0,022	0,074	0,393	-0,432
c[14]	-0,089	-0,109	0,126	0,752	-0,706
c[15]	0,002	0,009	0,072	0,380	0,028

Ao analisarmos a Tabela A.56 verificamos que 7 estimativas de c_i foram subestimadas e 8 sobrestimadas; c[6] apresentou a maior diferença do valor verdadeiro (-0,09) e, também, a maior amplitude (0,764); o maior VP foi c[2] = 0,824 e o maior desvio padrão foi c[5] = 0,14. Nenhum dos parâmetros apresentaram problemas em relação à proporção de cobertura dos ICr 95%, todos tiveram 100% de cobertura. Todos os parâmetro comentados aqui são valores extremos.

As estimativas foram boas nos dois ajustes, mas classificamos como melhor o ajuste assimétrico, que gerou diferenças e VP's menores.

Tabela A. 57 – Estimativas de λ_i para modelo assimétrico para *Dados 6*

parâmetro	Médias posteriores	Medianas posteriores	Desvios posteriores	Amplitudes posteriores	VP
	média	média	média dos desvios	média das amplitudes	
$\lambda[1]$	-1,237	-1,469	0,722	8,935	-1,713
$\lambda[2]$	0,673	-0,031	2,189	23,770	0,307
$\lambda[3]$	1,495	0,875	1,964	18,829	0,761
$\lambda[4]$	0,316	-0,241	1,703	17,176	0,186
$\lambda[5]$	0,195	-0,479	1,930	21,644	0,101
$\lambda[6]$	0,025	-0,123	0,479	7,259	0,052
$\lambda[7]$	1,185	0,564	1,904	19,860	0,622
$\lambda[8]$	0,000	-0,467	1,397	16,030	0,000
$\lambda[9]$	-0,051	-0,082	0,115	2,450	-0,443
$\lambda[10]$	-0,350	-0,500	0,497	8,051	-0,704
$\lambda[11]$	-0,059	-0,292	0,714	8,788	-0,083
$\lambda[12]$	-0,118	-0,312	0,586	7,721	-0,201
$\lambda[13]$	1,074	0,354	2,352	22,257	0,457
$\lambda[14]$	-0,005	-0,187	0,558	7,320	-0,009
$\lambda[15]$	1,780	1,036	2,414	23,632	0,737

Analisando a Tabela A.57, notamos que 6 estimativas de λ_i foram subestimadas e 8 sobrestimadas. A maior diferença ocorre em $\lambda[15] = 1,78$, que também obteve o maior desvio padrão (2,414); o maior VP foi $\lambda[1] = -1,713$. Em relação à proporção de cobertura dos ICr 95%, apenas $\lambda[1]$ teve menos de 100% de cobertura (70%), os demais parâmetros não apresentaram problemas. Destacamos que $\lambda[1]$ é um valor extremo. Observamos que as estimativas dos parâmetros produziram amplitudes elevadas.

Tabela A. 58 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo simétrico para *Dados 6*

parâmetro	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
	média	média	média dos desvios	média das amplitudes
VM	-0,648	-0,649	0,044	0,317
EQM	1,277	1,276	0,085	0,603

Tabela A. 59 – Resumo das estimativas das habilidades (θ) dos 500 indivíduos utilizando modelo assimétrico para *Dados 6*

	Média posterior	Mediana posterior	Desvio posterior	Amplitude posterior
parâmetro	média	média	média dos desvios	média das amplitudes
VM	-0,657	-0,657	0,045	0,312
EQM	1,303	1,302	0,086	0,613

Analisando as Tabelas A.58 e A.59, verificamos que o ajuste **simétrico** produziu menor viés e EQM e, tanto o ajuste simétrico quanto o ajuste assimétrico produziram um viés negativo.

A Tabela A.60 traz os critérios para seleção dos modelos. Notamos que o menor DIC foi encontrado no ajuste assimétrico.

Tabela A. 60 – DIC's e pD's para *Dados 6*

Ajuste simétrico			Ajuste assimétrico		
DIC	pD		DIC	pD	
8582 7521 7492	239,2	292,6 281,3	8188 7130 7186	-148,3	-118,2 -54,8
7434 7580 7465	317,7	297,8 289,5	7143 7220 7129	-24,13	-73,17 -110,5
7519 7245 7538	314,8	343,6 288,5	7061 7157 7121	-104,7	-113,8 -63,17
7468 7333 7520	313,4	323,4 328,1	7097 7018 7136	-48,11	-4,935 -83,17
7416 7526 7596	311,2	302,4 264,6	7039 7108 7113	-35,51	-49,24 -61,36
7401 7584 7501	329,8	320,4 297,6	7085 7018 7119	-58,9	-127 -0,17
7473 7508	328,7	278,7	7090 7147	-66,09	-70,72
Média = 7535,1 DP = 260,4967	Média = 303,165 DP = 25,204		Média = 7165,25 DP = 246,1056	Média = -70,7986 DP = 40,0275	

A partir dessas tabelas apresentadas, podemos tirar algumas conclusões mais gerais sobre as estimativas:

- 1) Quando inserimos assimetria na habilidade dos indivíduos, as estimativas dos parâmetros dos itens pioram, isto é, crescem as diferenças em relação ao valor verdadeiro do parâmetro e seus desvios. Além disso, surgem problemas com os ICr 95% (alguns intervalos de credibilidade não cobrem o verdadeiro valor do parâmetro, diminuindo a média final da proporção de cobertura), principalmente quando usamos o ajuste simétrico;

- 2) Quando as habilidades são simétricas (*Dados 1* e *Dados 4*) as estimativas dos itens são boas tanto pelo ajuste simétrico quanto no ajuste assimétrico. Somente b_i (pelo ajuste simétrico) em *Dados 4* apresentou estimativas não tão boas;
- 3) Quando as habilidades são assimétricas podemos dividir em dois grupos:
 - **Apenas habilidades assimétricas** (*Dados 2* e *Dados 3*): como já apontado anteriormente, as estimativas são melhores utilizando ajuste assimétrico. No caso simétrico, apenas estimativas de c_i foram razoáveis. O que nos leva a crer que o modelo assimétrico é melhor mesmo quando somente a distribuição das habilidades são assimétricas.
 - **Habilidades e CCI são assimétricas** (*Dados 5* e *Dados 6*): consideramos as estimativas dos itens pelo ajuste assimétrico como boas; já as estimativas pelo ajuste simétrico, novamente são razoáveis apenas para c_i . Comparando agora o ajuste assimétrico para os parâmetros dos itens entre *Dados 5* e *Dados 6*, percebemos que foram melhores quando a assimetria é positiva para as habilidades, ou seja, em *Dados 6*;
- 4) A média dos desvios padrão e as amplitudes são elevadas para as estimativas de b_i utilizando ajuste assimétrico em todos os conjuntos de dados;
- 5) As amplitudes são elevadas para os valores de λ_i em todas as estimativas, nos vários ajustes efetuados durante o estudo. Isso porque algumas estimativas apresentaram pontos discrepantes;
- 6) O melhor ajuste para λ_i ocorre em *Dados 6* (Tabela A.57).

APÊNDICE 2

Programa para ajuste simétrico:

```

model
{
  for(j in 1:I)
  {
    a[j]~ dnorm(1,2)I(0,)
    b[j]~ dnorm(0,0.5)
    c[j]~ dbeta(2,8)
  }
  for (i in 1:n)
  {  teta[i] ~ dnorm(0,1)
  }
  for(i in 1:n)
  {  for(j in 1:I)
    {  p[i,j] <- c[j] + (1-c[j])/(1+exp(-a[j]*teta[i]+b[j]))
      y[i,j] ~ dbern(p[i,j])
    }
  }
  for (i in 1:n)
  {  dif[i]<-(teta[i]-reais[i])
    dif2[i]<-pow(dif[i],2)
  }
  VM<-(sum(dif[]))/n
  EQM<- sum(dif2[])/n
}

```

Programa para ajuste assimétrico:

```

model
{

```

```

for(j in 1:I)
{ a[j]~ dnorm(1,2)I(0,)
  b[j]~ dnorm(0,0.5)
  c[j]~ dbeta(2,8)
  lambda[j] ~ dgamma(.25,.25)
}
for (i in 1:n)
{ teta[i] ~ dnorm(0,1)
}
for(i in 1:n)
{ for(j in 1:I)
  {
    po[i,j] <- pow(1+exp(-a[j]*teta[i]+b[j]),-lambda[j])
    p[i,j] <- c[j] + (1-c[j])*po[i,j]
    y[i,j] ~ dbern(p[i,j])
  }
}

for (i in 1:n)
{ dif[i]<-(teta[i]-reais[i])
  dif2[i]<-pow(dif[i],2)
}
VM<-(sum(dif[]))/n
EQM<- sum(dif2[])/n
}

```

Programa para geração de *Dados 1*:

```

n <- 500; I <- 15
a <- runif(J,.5,2)
b.star <- runif(J,-1.5,2.3)
b <- b.star*a
teta <- rnorm(n)

```

```

c <- sample(seq(from=0.15,to=0.3,by=0.05),I,replace=T)
#geracao de dados (matriz 'y')
y <- matrix(rep(0,n*I),ncol=I)
for(i in 1:n)
{ for(j in 1:I)
  { p <- c[j] + (1-c[j])/(1+exp(-a[j]*teta[i]+b[j]))
    y[i,j] <- rbinom(1,1,p)
  }
}
dsim1 = list(y=y,teta=teta,a=a,b=b,c=c)

```

Programa para geração de *Dados 5*:

```

n; I; a; b; c;
teta <- rsn(n,0,1,-3)
lambda<-runif(J,0.1,2)
#geracao de dados (matriz 'y')
y <- matrix(rep(0,n*I),ncol=I)
for(i in 1:n)
{ for(j in 1:I)
  { po <- ((1+exp(-a[j]*teta[i]+b[j])))^(-lambda[j])
    p <- c[j] + (1-c[j])*po
    y[i,j] <- rbinom(1,1,p)
  }
}
dsim5 = list(y=y,teta=teta,a=a,b=b,c=c,lambda=lambda)

```