

**INFERÊNCIA BAYESIANA
EM
MODELOS DA
TRI**

César Emanuel Cintra Rios

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Área de Concentração: **Estatística**

Orientadora: **Profa. Dra. Márcia D'Elia Branco**

-São Paulo, fevereiro de 2007-

Ao meu pai (in memoriam)

Agradecimentos

- ♦ A minha orientadora Profª. Dra. Márcia D'Elia Branco pela enorme paciência que teve comigo ao longo desses anos, pela compreensão das minhas faltas e, sobretudo pelo incentivo dado, desde quando a conheci na graduação.
- ♦ Aos meus pais, Manoel e Eva, pelo carinho e incentivo dado em toda minha caminhada.
- ♦ A minha querida Zenaide por compreender minha dificuldade, dar todo o apoio necessário e incentivo para que tudo tornasse realidade.
- ♦ A minha família de uma forma geral.
- ♦ Aos membros da banca examinadora.
- ♦ A Ana Paula e Jorge, meus eternos amigos, por toda ajuda dada.
- ♦ Ao meu amigo Holanda, por não ter colaborado em nada, pois sendo professor de história, não teria o que acrescentar.
- ♦ Ao Fábio Rendelucci e seus gráficos,
- ♦ A direção do curso Anglo Vestibulares, aqui representada pelo senhor Marcos Antonio Gabriádes, pelo fornecimento dos dados.
- ♦ Ao IME e a USP pela oportunidade de fazer meu mestrado em tão renomada instituição
- ♦ A todos aqueles que de alguma forma colaboraram para que esse trabalho se tornasse realidade.

Resumo

Neste trabalho apresentamos uma aplicação da Teoria da Resposta ao Item, utilizando um conjunto de dados de um cursinho pré-vestibular da cidade de São Paulo. Foram implementados os modelos com CCI proibido e proibido-assimétrico, bem como supondo assimetria para as habilidades dos indivíduos. A estimação dos parâmetros é obtida através da abordagem bayesiana.

Os programas foram implementados no software WINBUGS de uso livre e já bastante utilizado nos meios acadêmicos, o qual se mostrou uma boa alternativa aos programas clássicos usualmente considerados para a TRI.

Sumário

1	Introdução	
1.1	Introdução à TRI.....	1
1.2	Modelo.....	5
1.3	Modelo logístico de dois parâmetros para itens dicotômicos.....	7
1.4	Interpretação do modelo.....	10
1.5	Estimação dos parâmetros dos itens e das habilidades.....	12
1.6	Objetivos dessa dissertação.....	13
2	Modelos probabilísticos na TRI	
2.1	Introdução.....	14
2.2	Modelo probito normal.....	15
2.3	Modelo probito assimétrico normal.....	19
2.4	Modelo probito – assimétrico normal.....	22
2.5	Especificação das distribuições a priori para os parâmetros dos itens.....	22
3	Inferência bayesiana nos modelos TRI	
3.1	Inferência bayesiana via mcmc.....	24
3.2	Modelo de dados aumentados.....	25
3.3	Implementação no winbugs.....	31
3.4	Programas para implementação no winbugs.....	32
4	Aplicações dos modelos	
4.1	Introdução.....	36
4.2	Valores iniciais.....	37
4.3	Aplicação dos dados no modelo probito normal.....	37
4.4	Modelo probito-assimétrico normal.....	46
4.5	Modelo probito normal-assimétrico.....	61
	Apêndice A	

Apêndice B

Referências Bibliográficas

Capítulo 1

Introdução

1.1 Introdução à TRI

Hoje todas as pessoas que pensam na questão do desenvolvimento e da melhoria nas condições de vida de uma população sabem que uma das condições para isso é o investimento na área da educação, sobretudo a educação de qualidade. Este é um tema obrigatório para qualquer governo que esteja comprometido com a melhoria nas condições de vida da sua população.

Historicamente, todos os países desenvolvidos e aqueles que recentemente saíram da faixa do subdesenvolvimento desenvolveram grandes projetos, com investimento público ou privado, na educação de seus cidadãos, seja na educação de base ou no nível superior.

Vários exemplos poderiam ser citados, mas pode-se destacar mais recentemente a Coreia do Norte e a Índia, cujos governos adotaram políticas focadas nesta área. No Brasil a situação não é diferente. Nossos dirigentes sabem através de vários estudos do IBGE (Instituto Brasileiro de Geografia e Estatística) e do Ministério da Educação que não sairemos da atual condição de país em desenvolvimento se não houver uma política educacional decente para que possamos desenvolver novas tecnologias e poder competir de forma mais equilibrada em vários ramos que envolvem conhecimentos e domínio de tecnologias de ponta. Esta necessidade sempre existiu, mas se acentuou com o processo de globalização das últimas décadas.

Aqui, abre-se uma lacuna para o seguinte questionamento: “O que é uma educação com qualidade?” Obviamente isso pode ser respondido com vários argumentos como: bons professores com boa formação acadêmica, boas instalações físicas, acesso a bibliotecas e centros

de ensino, acesso à Internet, que trás uma universalização da informação, e metodologias educacionais desenvolvidas por educadores comprometidos com essa qualidade.

Porém, toda forma de aprendizado carece de um instrumento de mensuração, ou seja, uma ferramenta que permita ao profissional saber se a informação passada a um indivíduo está efetivamente sendo compreendida e se o instrumento utilizado está adequado para este fim, utilizando, no entanto, uma análise mais científica.

Existem várias formas de se aferir se tal fato está acontecendo, mas não há dúvida de que o instrumento mais utilizado é a aplicação de provas (testes), em que o respondente terá que escolher entre algumas opções de respostas, exemplo de teste de múltipla escolha, ou, no caso de provas objetivas, dar sua resposta a um questionamento aberto, permitindo, assim, não só uma análise de que o indivíduo acertou ou errou, mas sim o seu grau de conhecimento.

Seja num caso ou noutro, como saber se o teste aplicado está também adequado, ou seja, se os testes escolhidos como ferramentas conseguem, através de seus itens (questões), efetivamente mensurar se aquele respondente domina certo conhecimento?

Dessa forma é preciso estabelecer de claramente, sobretudo com o auxílio de um especialista, o que se quer medir efetivamente para se definir qual o melhor instrumento.

Se o interesse fosse somente concluir se ele respondeu certo ou errado ao item, ou seja, ao final sabermos qual o total de acertos no teste, estaríamos abrindo mão de informações muito importantes e fugindo do propósito inicial que seria ter um teste com qualidade e que pudesse ser analisado à luz de uma proposta educacional sem que isso seja feito de forma subjetiva.

Assim, cria-se a necessidade de se entender a forma com que essas informações contidas num teste, através de seus itens, serão tratadas e analisadas por quem se interessar, subtraindo das mesmas, dados que permitam verificar a adequação do item em medir certo conhecimento e com qual habilidade o item é respondido.

Ora, primeiramente pode-se observar que tanto no caso de testes de múltiplas escolhas ou no caso de uma prova objetiva, cuja resposta é aberta, a resposta dada por um respondente poderá ser tratada como certa ou errada. Esta característica dicotômica nos permite classificar esses dados como sendo do tipo binário.

Dados binários são aqueles em que as respostas são do tipo certo ou errado, sucesso ou fracasso, ou seja, aqueles cujos resultados apresentam ou não uma determinada característica.

Nos exemplos que seguem encontramos essas características:

- 1) Uma moeda é lançada e o resultado ou é coroa ou não;
- 2) Uma peça é escolhida ao acaso de um lote sendo com ou sem defeito.

Nos mais diferentes ramos de atividade, encontramos situações em que os dados envolvidos numa experiência são do tipo binário.

Se associarmos aos experimentos acima uma variável aleatória X que assume somente valores 1 e 0 com probabilidade p e $(1 - p)$, respectivamente, então ela será chamada de variável aleatória de Bernoulli.

Em particular, esse trabalho visa explorar dados na área da psicometria, ramo da psicologia cujo interesse é a análise das repostas dadas por um ou mais respondentes a um teste psicológico.

Sendo que da mesma forma que foi destacada anteriormente, se o respondente acertar ao item do teste será associado o valor 1 e se errar o valor 0.

A análise de um teste pode ser fundamentada em dois principais modelos: a teoria clássica dos testes (TCT) e a teoria da resposta ao item (TRI).

A TCT vem sendo usada há um bom tempo, tendo sido bastante útil no desenvolvimento dos testes psicológicos. Porém esta teoria apresenta várias limitações como, por exemplo, ser dependente do conjunto de itens que compõem o instrumento de medida. Esta dependência limita sua aplicabilidade, tornando-se inviável a comparação entre os resultados obtidos por indivíduos que não responderam a um mesmo teste ou, pelo menos, ao que denomina de formas paralelas de testes, em que eles estão medindo a mesma coisa, porém com itens diferentes. Gulliksen (1950), Lord & Novick (1968) e Vianna (1987), entre outros autores, abordaram este tema.

Na clássica síntese de Gulliksen (1950), é preciso saber o escore bruto, o escore verdadeiro e o erro devido a vários fatores. Já a TRI é uma metodologia que parte da hipótese de que existe no respondente uma característica individual (traço latente ou habilidade do indivíduo) determinante de como responder aos itens de um teste, apresentando uma vantagem sobre a TCT, ou seja, a comparação de populações distintas, bastando para isso que as mesmas sejam submetidas a testes com itens comuns ou indivíduos de uma mesma população submetidos a testes distintos.

Uma vez que as variáveis inerentes à conduta humana não são observáveis, pelo menos diretamente, elas são denominadas latentes, logo sua medição se dará através das repostas dadas por um indivíduo a um conjunto de itens, hipotetizando relações entre suas repostas e seu traço latente.

Podemos traduzir estas variáveis através de um parâmetro, que denotará “a medida” de proficiência, que é uma característica do respondente. O teste apresentará também outros tipos de parâmetros característicos do item como, por exemplo, dificuldade e discriminação.

Considera-se que os parâmetros de cada item não dependem dos outros itens que compõem o teste, mas que o resultado se faz em função das respostas do respondente a cada item.

Uma vantagem que a TRI apresenta é a possibilidade de comparação de desempenho entre populações distintas, desde que submetida a testes com itens comuns, fato este que não ocorre na TCT, pois esta se encontra focada na prova como um todo.

O traço latente possui uma relação probabilística com cada um dos itens que fazem parte do teste, dessa forma podemos estabelecer uma relação entre o desempenho num teste (dado pelo escore) e a habilidade do respondente.

Assim, é possível apreender algumas características dos itens e dos respondentes: saber se os respondentes são mais ou menos hábeis, se os itens apresentam menor ou maior grau de dificuldade, assim como serem capazes de discriminar.

Os princípios da teoria clássica (TCI) continuam valendo. A TRI não surgiu para anular a TCT, mas sim para trazer uma nova proposta estatística, a de centralizar a análise não no teste, característica da TCT, e sim em cada item que o compõe.

Traduzindo, a TRI quer dizer o seguinte: apresente ao individuo um estímulo, que neste caso será o teste com seus respectivos itens e ele responderá aos mesmos. A partir dessas respostas, pode-se inferir sobre o seu traço latente, assim como estimar parâmetros que caracterizam o teste na sua dificuldade e na capacidade de discriminação.

Vários modelos matemáticos poderiam expressar essa relação, eles são conhecidos de uma forma geral por modelos de um, dois e três parâmetros.

A seguir será apresentado um modelo logito de dois parâmetros bastante utilizado e que aparece com uma grande incidência na literatura que aborda esta teoria e suas aplicações em vários tipos de testes.

A partir deste modelo serão explicados os pontos mais importantes para o entendimento da TRI e as relações que se dão entre a habilidade do respondente e os parâmetros do item.

1.2 Modelo

Os modelos que foram inicialmente desenvolvidos na década de 50 consideravam que uma única habilidade estava sendo aferida e se baseavam na distribuição normal (ogiva normal) como em Lord (1952) com o modelo de dois parâmetros e Rash (1960) com o modelo de um parâmetro. Posteriormente, Birnbaum (1968) substituiu essa função pela distribuição logística.

O modelo logístico que matematicamente se torna mais atraente, pois apresenta uma função explícita dos parâmetros do item e da habilidade sem que envolva integração, o que não ocorre com o modelo de ogiva normal. Esse modelo é predominante até hoje e inclusive encontram-se no mercado softwares que se baseiam no mesmo, assim como no modelo normal.

Outros modelos, como o proposto por Samejima (1969) tem como foco não somente se o respondente acertou ou errou o teste, mas tenta obter mais informações acerca de suas respostas, esse modelo é conhecido como modelo de resposta gradual (Graded Response Model) e assume que é possível colocar as respostas dadas a um item em categorias. Um caso particular desse modelo é o modelo de escala gradual (Rating Scale Model), sendo diferente do anterior quanto à suposição de que os escores das categorias são igualmente espaçados, ou seja, os escores das categorias de respostas devem ser eqüidistantes.

Um modelo com mesmo objetivo do anterior foi desenvolvido por Bock (1972) e ficou conhecido como modelo de resposta nominal (Nominal Categories Model).

O detalhamento dos modelos anteriores, assim como outros modelos podem ser encontrados no excelente resumo de Andrade, Tavares & Valle (2000).

Historicamente os dados inerentes a conjunto de respostas de diferentes respondentes eram e são modelados por modelos simétricos com ligações denominadas probito (usando a distribuição normal acumulada) e logito (usando a distribuição logística).

Na TRI, os modelos matemáticos propostos apresentam uma relação entre os parâmetros dos itens com o parâmetro que define a habilidade do respondente através de uma função que define a probabilidade de responder corretamente ao item do teste.

De uma forma geral, os modelos levam em consideração, basicamente, três fatores:

- 1) Da natureza do item (dicotômico ou não dicotômico);
- 2) Do número de populações envolvidas (apenas uma ou mais de uma);

- 3) É da quantidade de traços latentes que está sendo medida (um ou mais, caracterizando a unidimensionalidade ou multidimensionalidade).

Nessa dissertação será considerado a unidimensionalidade do teste, ou seja, supõe-se que o teste mede um único traço latente, aquele que é dominante para que possa ser respondido o item de forma correta.

No entanto têm-se também modelos multidimensionais que consideram mais de uma habilidade necessária para se responder a um item, como no caso de um teste de matemática que cobra os conceitos inerentes à disciplina, assim como a capacidade de intelecção de texto, como é característica de muitos exames vestibulares de hoje, como, por exemplo, o ENEM (exame nacional do ensino médio).

Outro fato importante é que será considerado um único grupo de respondentes. Entendamos grupo como um conjunto de indivíduos que fazem parte de uma amostra aleatória simples de uma população.

A população depende de certas características, que podem variar dependendo do objeto de estudo.

1.2.1 Axiomas fundamentais

- 1) O desempenho de um respondente a um item do teste se aplica em função de seu traço latente, tendo o desempenho como efeito e o traço latente com a causa;

- 2) O desempenho no item tem relação entre o traço latente e os parâmetros do item e é descrito por uma equação monotônica crescente chamada de CCI (Curva característica do item) exemplificada na figura que segue, onde se observa que quanto maior for a habilidade do respondente, maior será a probabilidade de responder corretamente ao item.

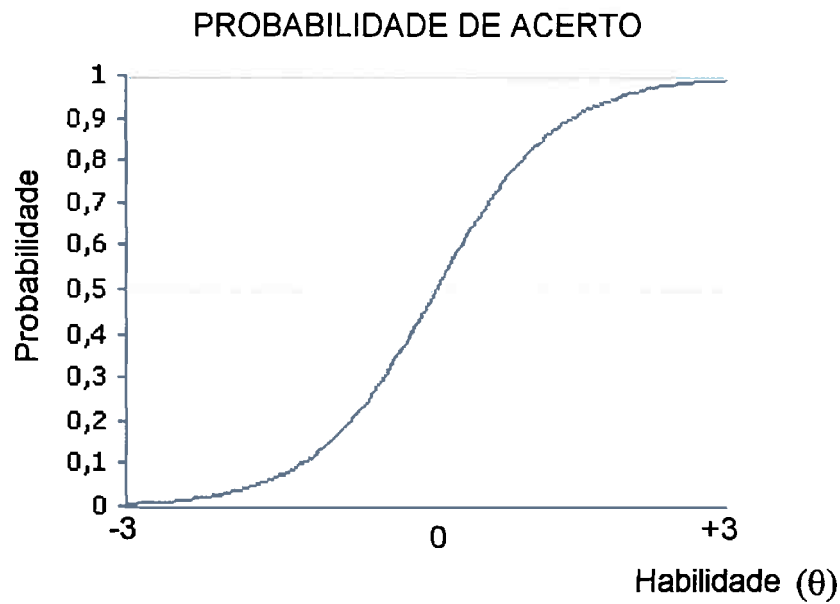


Figura 1.1 Gráfico da probabilidade de acerto em função da habilidade

1.3 Modelo logístico de dois parâmetros para itens dicotômicos

O modelo que segue abaixo é denominado modelo logístico de dois parâmetros e pode ser utilizado tanto em testes de múltipla escolha (quando corrigidos como certo ou errado) como para itens abertos (nos quais a resposta é aberta, mas corrigidos de forma certo ou errado). Nesse modelo serão considerados os parâmetros discriminação (a) e dificuldade (b).

$$P(u_{j,i} / \theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_i)}} \quad (1.1)$$

Com $i = 1, 2, 3, \dots, I$ e $j = 1, 2, 3, \dots, J$, sendo:

$U_{j,i}$ uma variável dicotômica que assume valores 1 ou 0, conforme o indivíduo j responde correta ou incorretamente o item i , respectivamente;

θ_j representa a habilidade ou traço latente do j -ésimo indivíduo;

$P(U_{j,i} = 1 / \theta_j)$ a probabilidade de um indivíduo j com habilidade θ_j responder corretamente o item i e é chamada função de resposta do item;

b_i o parâmetro de dificuldade ou de posição do item i , medido na mesma escala da habilidade;

a_i o parâmetro de discriminação ou inclinação do item i , medido na mesma escala da habilidade com valor proporcional à inclinação da CCI (curva característica do item) no ponto b_i .

I número de itens

J número de indivíduos

A equação acima pode ser entendida como a proporção de respostas corretas dadas ao item i dentre todos os respondentes que tem habilidade θ_j .

O modelo proposto leva em consideração que quanto maior for a habilidade do respondente, maior será sua probabilidade de acertar ao item e que a relação não é linear, mas uma curva em forma de S com inclinação e deslocamentos definidos pelos parâmetros a e b , respectivamente e na mesma escala das habilidades.

A função é conhecida como função de resposta ao item ou curva característica do item (CCI) e satisfaz a propriedade de ser monotônica, ou seja, é uma função estritamente não decrescente e a mesma para todo j e i e é obtida a partir de $P(U \mid \theta)$.

O modelo proposto tem como hipótese a unidimensionalidade do teste, ou seja, a suposição de que o conjunto dos itens que compõem o teste deve estar medindo uma única habilidade latente que será responsável pela resposta dada a todos os itens do teste.

Na verdade várias habilidades podem ser necessárias para se responder certo item, porém vamos considerar que exista uma dominante, pois assim o postulado da unidimensionalidade é satisfeito.

Outra suposição a ser feita nesse modelo é a independência local ou independência condicional, onde deve se assumir que, para uma habilidade dada, as respostas a diferentes itens do teste são independentes. Essas suposições são fundamentais para a implementação do modelo.

Embora pareça improvável que os comportamentos de um respondente não sejam correlacionados, a independência local diz que se houver correlação, essas se devem à influência de outros fatores. Se os mesmos forem controlados, o traço latente dominante será a única fonte de variação e as respostas se tornam independentes.

Um ponto importante é a escala de habilidade, pois é arbitrária, mas sua escolha define a escala em que estarão os parâmetros do item, o que nos permite sua interpretação. Assim como as relações dela com os parâmetros e não sua magnitude, ou seja, não faz sentido analisar itens a partir de seus valores sem conhecer a escala na qual foram determinados.

Caso seja de interesse, pode-se mudar a escala através de transformações lineares, como podem ser visto em Andrade, Tavares & Valle (2000). Uma nova escala assim será obtida e a probabilidade de um respondente dar uma resposta correta a um item é sempre a mesma, sendo, portanto invariante à escala de mediada adotada.

Um problema enfrentado na TRI, sobretudo para a estimação dos parâmetros usando máxima verossimilhança, é quando um respondente acerta todas as respostas ou erra todas, problema esse não encontrado quando a estimação é feita usando métodos bayesianos, como será abordado nessa dissertação.

Um questionamento a ser feito é se a velocidade com que a função converge para 1 é a mesma com o qual ela converge para 0, pois se elas forem diferentes, as funções simétricas, como a função distribuição acumulada da normal ou a logística, se tornam inadequadas, ou seja, não se verifica a suposição de normalidade para a distribuição dos escores, abrindo assim a necessidade de se trabalhar com funções do tipo assimétricas, ver Chen (1999) e Bazan (2005).

Muitos autores têm questionado esta suposição (Samejima, 1997, e Micceri, 1989), pois a consideram restritiva no que diz respeito à modelagem da conduta humana através de seu traço latente. Vários casos abordam essa situação, como Micceri (1989) em que são apresentados vários exemplos em que a normalidade para a variável latente não é observada.

Com o propósito de contornar o problema, e para se afastar dessa suposição de normalidade, Samejima (2000) propôs uma família de modelos assimétricos e inclui a CCI baseada na ligação probito como um caso particular.

Nessa dissertação, será considerada uma nova curva característica do item (CCI) Bazan (2005) considerando a função de distribuição acumulada da normal assimétrica (Azzalini, 1985, Dalla Valle, 2004), também sendo usada essa distribuição no caso de assimetria na distribuição da habilidade.

A distribuição normal assimétrica, assim como suas propriedades, serão apresentadas no apêndice A.

1.4 Interpretação do modelo

Para melhor interpretação do modelo destacado anteriormente, consideremos as figuras que seguem:

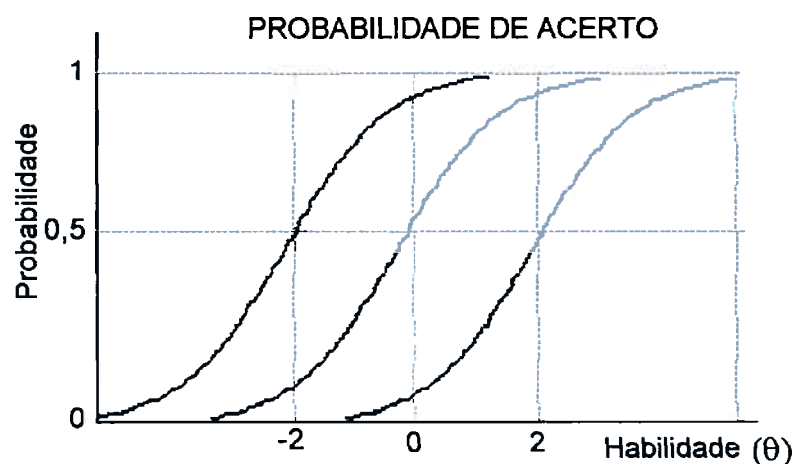


Figura 1.2 Relação do parâmetro b com a habilidade.

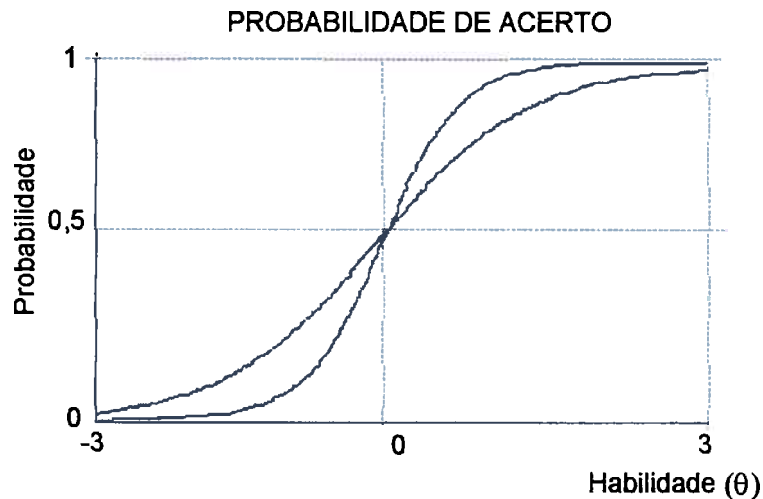


Figura 1.3 Influência do parâmetro a

Pode-se observar que quanto maior for a habilidade do respondente, maior será a probabilidade de acertar ao item.

No entanto se fixarmos o valor do parâmetro que mede a habilidade e variarmos o valor do parâmetro b teremos a seguinte interpretação: quanto maior for o valor de b , menor será a probabilidade de acerto, assim esse parâmetro caracteriza a dificuldade do item. A figura (1.2) mostra que para uma mesma probabilidade de acerto, se o item for mais difícil, deslocamento para a direita, maior deverá ser a habilidade do respondente para acertá-lo.

O parâmetro a é proporcional ao coeficiente angular da reta tangente no ponto de inflexão da CCI, assim quanto maior for o valor de a , mais acentuada será a inclinação da reta tangente, cuja consequência será uma maior diferença entre as probabilidades de dois indivíduos, com habilidades diferentes, acertarem o item, ou seja, o item tem a capacidade de discriminar os respondentes mais hábeis dos menos hábeis.

A seguir é apresentada uma tabela contendo valores de probabilidades, sendo feitas variações ou nos parâmetros dos itens ou no parâmetro que mede a habilidade, com base no modelo de dois parâmetros apresentado anteriormente.

Item	a	b	θ	$P(\theta)$
1	1	-2	-1	0,73
2	1	-3	-1	0,88
3	2	-1	1	0,98
4	3	-1	1	0,99
5	1	-1	1	0,88
6	1	-1	3	0,98
7	2	3	2	0,11

Tabela 1.1 Probabilidades para alguns valores de a , b e θ

Teoricamente os parâmetros de dificuldade e discriminação podem assumir qualquer valor no conjunto dos números reais, sendo que não se espera valores negativos para o parâmetro de discriminação, pois dessa forma teríamos a probabilidade de acertar um item diminuindo à medida que a habilidade do respondente aumentasse o que, de certa forma, não é aceitável.

Da mesma forma não se espera valores nulos para a , pois este se associa à inclinação da reta tangente a CCI no ponto de inflexão, teríamos que para qualquer habilidade a probabilidade de acerto do item seria a mesma.

1.5 Estimação dos parâmetros dos itens e das habilidades

Uma abordagem completa sobre métodos de estimação dos parâmetros dos itens e das habilidades dos respondentes pode ser encontrada em Baker (1992) e Andrade, Tavares & Valle (2000) usando estimação pelo método da máxima verossimilhança através da aplicação de métodos iterativos como os algoritmos de Newton-Raphson ou Scoring de Fisher.

Também são encontrados procedimentos bayesianos de estimação, diferentes do que será abordado nessa dissertação.

Na situação em que são conhecidos os parâmetros dos itens, pois os mesmos já foram calibrados e fazem parte, por exemplo, de um banco de itens, desejamos estimar as habilidades dos respondentes.

Podemos ter também uma situação onde sabemos a habilidades dos respondentes e queremos calibrar os itens estimando seus parâmetros.

Mas na maioria das vezes a estimação será feita tanto para os parâmetros dos itens como das habilidades dos respondentes.

1.6 Objetivos dessa dissertação

Essa dissertação tem como objetivo a aplicação dos modelos probito-normal (PN), probito-assimétrico normal (PAN) e probito normal-assimétrico (PNA) a um conjunto de dados observado de um grupo de estudantes que se prepara para o vestibular.

No capítulo 2 serão apresentados os modelos e suas características sob a perspectiva bayesiana, sobretudo na definição das prioris a serem usadas para as distribuições dos parâmetros e do parâmetro que mede a habilidade dos respondentes.

No capítulo 3 será feito um resumo do método bayesiano de estimação, assim como a implementação dos modelos no programa Bugs .

No capítulo 4 será feita uma aplicação ao conjunto de dados com a obtenção das estimativas e suas respectivas interpretações.

Ao final, no apêndice A, temos um resumo dos conceitos e propriedades da distribuição assimétrica. Já no apêndice B, é apresentado o conjunto de respostas dadas por quarenta e sete alunos que foram submetidos a um teste contendo quinze itens de múltipla escolha, apresentando somente uma resposta correta para cada item com o respectivo histograma e dot-plot.

Capítulo 2

Modelos probabilísticos na TRI

2.1 Introdução

Na teoria da resposta ao item (TRI) temos que a resposta dada por um indivíduo a um item de um teste depende, sobretudo, da sua habilidade (traço latente) e dos parâmetros que compõe o item, já chamados de discriminação e dificuldade.

Nos modelos dominantes e alvos de muitos trabalhos há uma constante no que diz respeito a assumir uma normalidade para a variável aleatória que mede a habilidade do indivíduo através de um traço latente dominante.

Não que essa suposição não seja verificada em muitos casos, até porque muitos modelos se baseiam nela, porém em alguns trabalhos podem-se encontrar resultados que diferem dessa suposição, assim modelos que se embasam nessa hipótese de normalidade podem não apresentar resultados verossímeis.

Os modelos caracterizados através da TRI não somente necessitam de equações matemáticas para descrevê-los com mais propriedade, levando em consideração a suposição dita anteriormente, mas também no que diz respeito à ligação utilizada.

Serão descritos na próxima seção modelos de respostas dicotômicas conhecidos como probito-normal (PN), com a suposição de simetria na ligação e na distribuição da habilidade, probito assimétrico-normal (PAN), com assimetria na ligação e simetria na distribuição da habilidade e por último o modelo probito normal-assimétrico (PNA), com simetria na ligação e assimetria na distribuição da habilidade.

Os dois últimos modelos destacados acima se justificam em muitos casos e permitem uma maior liberdade na adequação do modelo aos dados a serem estudados (Bazan, 2005).

Muitos modelos com ligação do tipo logito e probito aparecem em grande parte da literatura da área e já se tem em muitos casos sua implementação já contemplada em softwares comerciais como o Bilog.

Como a abordagem nessa dissertação será no contexto bayesiano, os modelos a serem utilizados serão os destacados acima. Esse fato se justifica pela maior facilidade do uso desses modelos, sobretudo pelo uso de conjugações quando a distribuição normal é usada, como por exemplo, no estabelecimento das distribuições a priori para os parâmetros do item.

Muitas referências na literatura dão destaque aos vários modelos da TRI, principalmente numa abordagem clássica como, por exemplo, Baker (1992) e na língua portuguesa Andrade, Tavares & Valle(2000).

Em todos os modelos destacados o uso de métodos de simulação se baseia em MCMC (Monte Carlo Markov Chain) para estimação do traço latente e dos parâmetros dos itens, sendo que em particular nessa dissertação será utilizado o método de dados aumentados (DAGS - data augmented Gibbs sampling), usado por Albert (1992), usando variáveis auxiliares.

Os métodos MCMC apresentam simulações de amostras aleatórias de uma distribuição multivariada normalmente desconhecida (Chen, 2000) e Patz e Junker (1999) e se baseiam em cadeias de Markov cuja distribuição estacionária multivariada de interesse queremos amostrar, sendo que no contexto bayesiano será uma distribuição a posteriori.

2.2 Modelo probito – normal (PN)

Considere $U_{j,i} \mid \theta_j, a_i, b_i \sim \text{Bernoulli}(p_{j,i})$

Com $U_{j,i} = \begin{cases} 1 \\ 0 \end{cases}$, sendo 1 para respostas corretas e 0 se o indivíduo erra a resposta.

Em que p_{ij} é uma função do parâmetro θ_j associado ao indivíduo j e a_i e b_i são os parâmetros associados ao item i .

Temos que a probabilidade do indivíduo j responder corretamente ao item i , com $j = 1, 2, 3, \dots, J$ e $i = 1, 2, 3, \dots, I$ é dada por:

$$p_{j,i} = P(u_{j,i} \mid \theta_j, a_i, b_i) = \Phi(a_i \theta_j - b_i) \quad (2.1)$$

Sendo $a_i \theta_j - b_i$ uma função linear de θ_j , com $j = 1, 2, 3, \dots, J$ e $i = 1, 2, 3, \dots, I$ (Johnson e Albert, 1999) com $a_i \in \mathbb{R}_+$ e $\theta_j, b_i \in \mathbb{R}$.

Para as variáveis $u_{j,i}$ que representam as respostas binárias, temos que $\mathbf{u} = (u_{1,1}, u_{1,2}, \dots, u_{1,I})$ é um vetor das respostas dadas pelo indivíduo 1 a todos os itens numerados de 1 a I e $\mathbf{u} = (u_{2,1}, u_{2,2}, \dots, u_{2,I})$ é um vetor das respostas dadas pelo indivíduo 2 a todos os itens numerados de 1 a I e assim sucessivamente para todos os indivíduos J , formando uma matriz de J linhas e I colunas, ou seja, nas linhas temos os indivíduos e nas colunas os itens.

A seguir temos uma tabela exemplificando as respostas dadas a um teste.

	1	2	3	4	...	I
1	0	1	1	0	...	1
2	1	1	1	1	...	1
3	0	0	0	0	...	0
4	1	0	1	0	...	1
5	0	0	1	1	...	0
...
J	1	1	0	0	...	1

Tabela 2.1 Respostas dadas por J indivíduos a I itens

Para o modelo em (2.1) é considerado que a resposta de certo individuo depende unicamente de sua habilidade ou traço latente θ_j , em que $\theta_1, \theta_2, \dots, \theta_j$ representa o traço latente dos J indivíduos.

Esse traço latente não é diretamente observado, mas sabemos qual foi a resposta dada pelo individuo J para cada item particular do teste.

Outro fato importante é que assumimos independência local nas respostas dadas pelos J indivíduos, ou seja, uma eventual resposta dada num item não depende da resposta dada a um outro item.

A função Φ dada em (2.1) é definida como curva característica do item (CCI) e deve satisfazer a propriedade de ser monotônica, sendo estritamente não decrescente em θ_j e tipicamente a mesma para todo j e i.

Usando o fato acima, temos que a probabilidade de se observar um particular conjunto de respostas para um individuo j é dada por:

$$P(u_{j,1}, u_{j,2}, \dots, u_{j,I} \mid \theta_j) = \prod_{i=1}^I [\Phi(a_i \theta_j - b_i)]^{u_{j,i}} [1 - \Phi(a_i \theta_j - b_i)]^{1-u_{j,i}} \quad (2.2)$$

Com os traços latentes dos J indivíduos dados pelo vetor $\theta_j = (\theta_1, \theta_2, \dots, \theta_j)$, os parâmetros de discriminação e dificuldade dados pelos vetores $\mathbf{a} = (a_1, a_2, \dots, a_I)$ e $\mathbf{b} = (b_1, b_2, \dots, b_I)$, respectivamente e com $j = 1, 2, 3, \dots, J$ e $i = 1, 2, 3, \dots, I$.

Considerando a que as respostas dadas por todos os indivíduos J são independentes, temos:

A distribuição conjunta de $(\theta, \mathbf{a}, \mathbf{b})$ é dada por:

$$P(u_{j,1}, u_{j,2}, \dots, u_{j,I} \mid \theta_j) = \prod_{j=1}^J \prod_{i=1}^I [\Phi(a_i \theta_j - b_i)]^{u_{j,i}} [1 - \Phi(a_i \theta_j - b_i)]^{1-u_{j,i}} \quad (2.3)$$

Com $j = 1, 2, 3, \dots, J$ e $i = 1, 2, 3, \dots, I$

Pode se observar que o modelo acima possui $J + 2 I$ parâmetros desconhecidos e é não identificável desde que seja preservado sob uma classe especial de transformações dos parâmetros (Albert, 1992).

Uma solução possível para contornar essa dificuldade é impor aos parâmetros dos itens certas restrições (Bock e Aitikin, 1981).

No modelo acima será considerada a distribuição normal para a distribuição associada a variável que mede o traço latente como segue:

Distribuição a priori para θ_j

$$\theta_j \sim N(\mu_\theta, \sigma_\theta^2)$$

Assim temos o caso em que se supõe que a variável latente apresenta um bom comportamento e que as habilidades são amostras aleatórias desta distribuição (Albert, 1992).

No capítulo 3 será considerado $\mu_\theta = 0$ e $\sigma_\theta^2 = 1$, pois dessa forma resolve-se o problema da indeterminação da escala (Suen, 1990) e conseqüentemente o problema da identificação dos parâmetros dos itens e suas interpretações (Lord, 1980).

Essa escolha de escala é arbitrária e podendo ser feitas transformações nessa escala, porém as probabilidades condicionais permanecem as mesmas, fato esse explicado por Andrade, Tavares & Valle (2000).

Para a inferência bayesinana temos que especificar a função de verossimilhança. Considere que \mathbf{u} representa o vetor com os dados observados, assim a função de verossimilhança para o modelo probito-normal é dada por:

$$L(\theta, \mathbf{a}, \mathbf{b} | D) = \prod_{j=1}^J \prod_{i=1}^I [\Phi(\mathbf{a}_i \theta_j - b_i)]^{u_{ji}} [1 - \Phi(\mathbf{a}_i \theta_j - b_i)]^{1-u_{ji}} \quad (2.3)$$

Com $j = 1, 2, 3, \dots, J$ e $i = 1, 2, 3, \dots, I$

Em que $\Phi(\cdot)$ dada em (2.3) é a função acumulada da distribuição normal padrão.

2.3 Modelo probito assimétrico-normal (PAN)

A seguir será discutido o uso de uma ligação assimétrica, uma vez que a probabilidade de sucesso pode-se se aproximar de 0 a uma taxa diferente da que se aproxima de 1 como já discutido anteriormente. Várias propostas de uso das distribuições assimétricas são encontradas nas literaturas da área e para maior entendimento podemos observar o gráfico que segue, apresentando curvas características do item (CCI) com assimetria e a curva característica do item do modelo simétrico como referência.

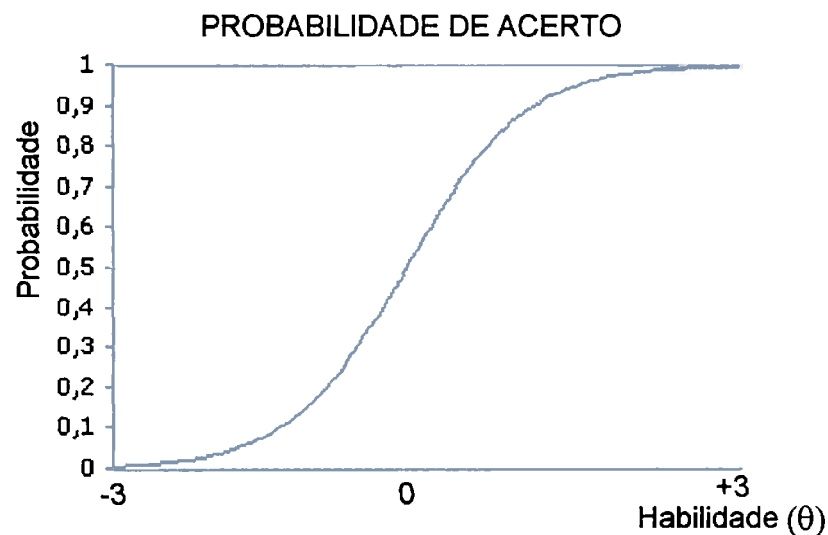


Figura 2.1 CCI no modelo assimétrico

Aqui será considerada a ligação probito assimétrica (Bazan, 2005), a qual se baseia na função de distribuição acumulada da distribuição normal assimétrica com vetor de parâmetros $\varphi = (\mu, \sigma^2, \lambda)$ em que μ é denominado parâmetro de locação, σ^2 parâmetro de escala e λ é um parâmetro de assimetria cuja notação utilizada é F_φ (Dalla Valle, 2004).

Assim, temos:

$$p_{i,j} = F_{\varphi}(a_i\theta_j - b_i) \quad (2.4)$$

Com $j=1, 2, 3, \dots, J$ e $i=1, 2, 3, \dots, I$

Para que não se tenha uma distribuição muito parametrizada, vamos assumir que $\mu = 0$ e $\sigma^2 = 1$, assim neste caso temos a distribuição normal assimétrica padrão (Azzalini, 1985), deixando livre λ , ou seja, assumindo valores para cada um dos itens, obtendo assim CCI simétricas ou assimétricas.

Dessa forma pode-se considerar que esse parâmetro ligado à função de ligação também terá sua interpretação como parâmetro do item, não mudando as interpretações já dadas anteriormente para os demais parâmetros, mas introduzindo um parâmetro novo com suas particulares interpretações que serão abordadas no capítulo 4 com a aplicação do modelo ao conjunto de dados.

A função $F_{\varphi}(\cdot)$ dada em (2.4) pode ser escrita da seguinte forma:

$$F_{\varphi}(a_i\theta_j - b_i) = 2\phi_2\left(\left(\begin{matrix} a_i\theta_j - b_i \\ 0 \end{matrix}\right), \left(\begin{matrix} 0 \\ 0 \end{matrix}\right), \left(\begin{matrix} 1 & -d_i \\ -d_i & 1 \end{matrix}\right)\right) \quad (2.5)$$

Onde $\phi_2(\cdot)$ é a função distribuição acumulada da normal bi variada e $|d_i| \leq 1$ é o coeficiente de correlação da normal bi variada com $d_i = \frac{\lambda_i}{\sqrt{1 + \lambda_i^2}}$, em que λ_i é o parâmetro de assimetria da normal assimétrica e d_i é uma parametrização cuja transformação é de 1-1, sendo dessa forma, também, um parâmetro de assimetria.

No caso acima considerando $\lambda_i = 0$ a expressão se reduz a $\phi(a_i\theta_j - b_i)$, o que leva ao modelo probito-normal.

A função de verossimilhança fica dada por:

$$L(a, b, \theta, \lambda / u) = \prod_{j=1}^J \prod_{i=1}^I [F_{\lambda_i}(a_i\theta_j - b_i)]^{u_{ij}} [1 - F_{\lambda_i}(a_i\theta_j - b_i)]^{1-u_{ij}} \quad (2.6)$$

A figura seguinte apresenta algumas CCI probito assimétrica para valores fixos de a_i e b_i e variando os valores para λ_i .

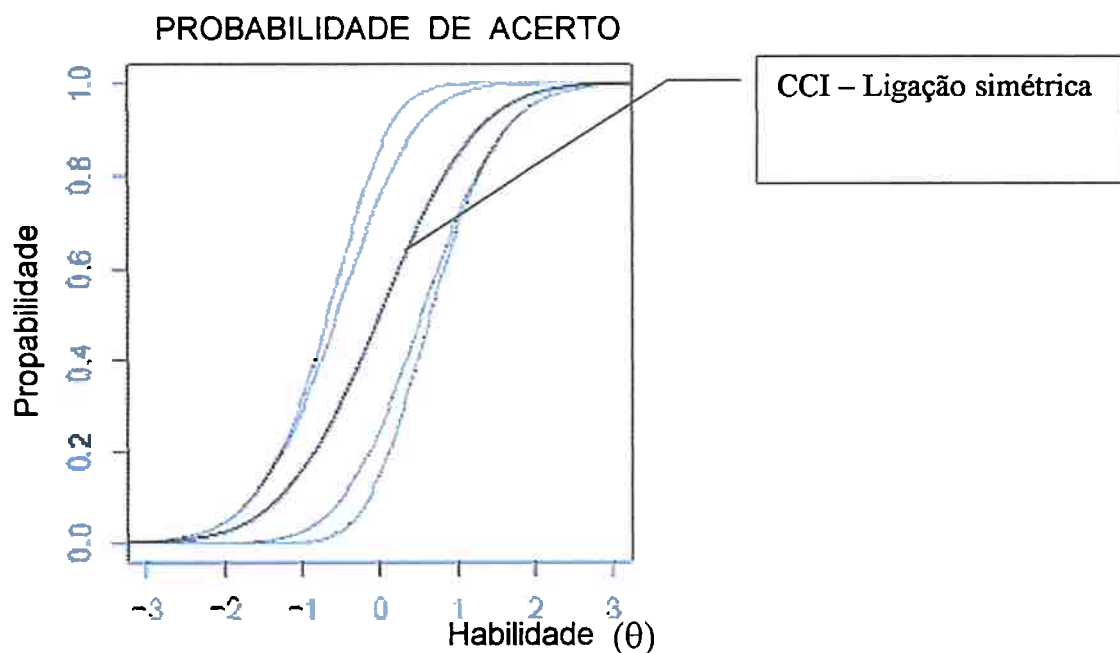


Figura 2.2 CCI assimétricas e simétrica

Neste modelo será considerada a mesma distribuição para os parâmetros que mede a habilidade, ou seja, neste modelo temos também que $\theta_j \sim N(\mu_\theta, \sigma_\theta^2)$, sendo $\mu_\theta = 0$ e $\sigma_\theta^2 = 1$.

2.4 Modelo probito-assimétrico normal (PAN)

Nesse caso será considerada uma ligação simétrica e uma assimetria na distribuição do parâmetro que mede a habilidade do j -ésimo indivíduo.

Será considerado que os θ_j serão independentes e identicamente distribuídos com distribuição dada por:

$$\theta_j \sim NA(\mu_\theta, \sigma_\theta^2, c)$$

Com $c = \frac{\alpha}{\sqrt{1+\alpha^2}}$ sendo uma parametrização de α , ou seja, uma transformação de 1-1

de α e $|c| \leq 1$ e com μ_θ sendo o parâmetro de locação e $\mu_\theta \in \mathbb{R}$, $\sigma_\theta^2 \in \mathbb{R}_+^*$ o parâmetro de escala e $\alpha \in \mathbb{R}$ sendo o parâmetro de assimetria e comum a todos os indivíduos.

No apêndice A será apresentado algumas propriedades dessa distribuição e uma forma de escrevê-la como combinação linear de uma variável aleatória com distribuição Normal (0.0,1.0) e uma outra variável aleatória com distribuição Half Normal (0.0,1.0) e independentes.

Assim ficam já estabelecidas as prioris a serem usadas na implementação do modelo probito-normal assimétrico (PNA). Essa necessidade se dá pelo fato de que no programa a ser usado Winbugs, ainda não se encontra implementada a distribuição normal assimétrica para ser usada diretamente.

2.5 Especificação das distribuições da priori para os parâmetros dos itens

Na intenção de se obter uma posteriori própria é necessário que se especifique prioris próprias. Além disso, distribuições a priori que sejam informativas podem ser usadas por não se

acreditar que os valores dos parâmetros dos itens não são extremos, ou seja, não se encontram no limite do espaço paramétrico.

Para o parâmetro de dificuldade tem-se que $a_i \sim N(\mu_a, \sigma_a^2) I(0, \infty)$ com $i = 1, 2, \dots, I$, pois não se espera valores menores ou iguais a zero, uma vez que se o valor fosse negativo a interpretação seria de que à medida que a habilidade do respondente fosse maior, a probabilidade seria menor e no caso de ser zero seria o mesmo que dizer que a probabilidade de acerto seria a mesma para qualquer valor do parâmetro que mede a habilidade.

Para a implementação do modelo no capítulo 3 será considerado que $\mu_a = 1$ e $\sigma_a^2 = 0,5$ como em Bazan (2005).

No caso do parâmetro que mede a dificuldade do item será considerado que $b_i \sim N(\mu_b, \sigma_b^2)$ com $i = 1, 2, \dots, I$.

Na priori acima, caso não se tenha muita informação a respeito do parâmetro de dificuldade do item, podemos adotar um valor para σ_b^2 grande, podendo ter um efeito pequeno na distribuição a posteriori para dados não extremos, quando um indivíduo responde correta ou incorretamente a todos os itens de um teste. Porém (Sahu, 2002) coloca que para valores grandes de μ_b encontramos estimativas instáveis.

Para estimação feita no capítulo 3 será considerado $\mu_b = 0$ e $\sigma_b^2 = 2$ como em Bazan (2005), pois em estudos anteriores elas se mostraram adequadas pois se obtém estimativas de grande precisão para o parâmetro de dificuldade do item.

Para o modelo probito assimétrico - normal, em que se introduz um novo parâmetro de item será considerado que $d_i \sim U(a, b)$ com $i = 1, 2, \dots, I$

Sendo d_i uma reparametrização para λ_i , parâmetro ligado a assimetria para o modelo probito assimétrico- normal (PAN), temos que d_i oferece a mesma interpretação que o parâmetro λ numa escala de valores em $[-1, 1]$,

Assim a priori que será usada para implementar o modelo e conseqüente estimação dos parâmetros terá $a = -1$ e $b = 1$ que se equivale a considerar que $\lambda_i \sim T(0.0, 0.5, 2.0)$, conforme será apresentado no apêndice A.

Capítulo 3

Inferência bayesiana nos modelos TRI

3.1 Inferência bayesiana via MCMC

O processo conhecido como MCMC (Monte Carlo Markov Chain) estabelece um algoritmo para simular uma distribuição a posteriori $\pi(\cdot)$ cuja função pode-se apresentar em três casos:

- 1) tem a sua forma conhecida;
- 2) a forma da distribuição é desconhecida, mas a forma da distribuição condicional dos parâmetros é conhecida;
- 3) a forma da distribuição a posteriori é não tratável analiticamente.

Vamos tratar aqui do terceiro caso, pois nos interessa essa situação, já que a mesma será a abordagem na implementação do programa para estimação dos parâmetros nesse trabalho usando o programa Winbugs.

No algoritmo MCMC temos três formas conhecidas: Algoritmo Metropolis Hastings, Algoritmo Metropolis e Gibbs sampler.

O amostrador de Gibbs é essencialmente um esquema interativo de amostragem de uma cadeia de Markov cujo núcleo é formado pelas distribuições condicionais completas, usando, assim, cadeias de Markov para gerar amostras de uma função de densidade de probabilidade.

Descrição do algoritmo do amostrador:

Suponha que a distribuição de interesse seja $\pi(\theta)$ onde $\theta = (\theta_1, \theta_2, \dots, \theta_j)$ são amostras dos parâmetros que se quer estimar e que as densidades condicionais completas a posteriori estão disponíveis.

Suponha agora que a geração direta de $\pi(\theta)$ é extremamente complicada, mas que a geração das marginais é possível de se realizarem.

O algoritmo de Gibbs nos fornece uma alternativa para essa questão, ou seja, fazendo sucessivas gerações das distribuições condicionais completas, cujo procedimento é o seguinte:

- 1) inicialize o contador de interações da cadeia $j=1$ e arbitre valores iniciais para θ^0 ;
- 2) obtenha um novo valor θ^j através de sucessivas gerações de valores;
- 3) mude o contador j para $(j+1)$ e retorne a (2) até atingir a convergência.

3.2 Modelos de dados aumentados

3.2.1 Modelo probito-normal

Em Albert (1992) é apresentado um modelo que considera não só os dados observados, mas também um conjunto de dados de uma variável latente auxiliar, pois não é observada, denotada por Z .

Para a distribuição definida em () será introduzida um conjunto de variáveis Latentes $Z_{j,i}$, tal que:

$$Z_{j,i} | \theta, a, b \sim N(a_i \theta_j - b_i, 1) \quad (3.1)$$

Sendo $u_{j,i} = 1$ se $Z_{j,i} > 0$ e $u_{j,i} = 0$ se $Z_{j,i} \leq 0$

Em Albert (1992) o modelo TRI probito normal, envolvendo I itens e J indivíduos, com $u_{j,i} \sim \text{Bernoulli}(p_{j,i})$ e $p_{j,i} = \phi(a_i \theta_j - b_i)$ é equivalentemente definido considerando que

$$Z_{j,i} = \alpha_i \theta_j - b_i + e_{j,i} \quad (3.2)$$

onde $e_{j,i} \sim N(0,1)$,

No modelo dado em (3.1) a equivalência segue ao verificar

$$\text{que } p_{j,i} = P(U_{j,i} = 1 | \theta_j, \alpha_i, b_i, z_{j,i}) = P(Z_{j,i} > 0 | \theta_j, \alpha_i, b_i) = \Phi(\alpha_i \theta_j - b_i), \quad (3.3)$$

com $i=1,2,\dots,I$ e $j=1,2,\dots,J$, pois $Z_{j,i} | \theta_j, \alpha_i, b_i \sim N(\alpha_i \theta_j - b_i, 1)$.

O resultado anterior mostra que essa estrutura linear para a variável $Z_{j,i}$ leva a um modelo equivalente ao modelo probito-normal.

Assim a função de verossimilhança considerando agora também os dados s aumentados com vetor $Z = (z_{1,1}, \dots, z_{J,I})$ é dada por:

$$L(\theta, \alpha, b | u, z) = \prod_{i=1}^I \prod_{j=1}^J [\Phi(z_{j,i}, \alpha_i \theta_j - b_i, \sigma^2) I^*(z_{j,i}, u_{j,i})] \quad (3.4)$$

Onde $\Phi(z_{j,i}, \alpha_i \theta_j - b_i, \sigma^2)$ denota a densidade da distribuição normal da variável z com média $\alpha_i \theta_j - b_i$ e variância σ^2 e $I^*(z_{j,i}, u_{j,i}) = I(Z_{j,i} > 0)I(u_{j,i} = 1) + I(Z_{j,i} \leq 0)I(u_{j,i} = 0)$, $i=1,2,\dots,I$ e $j=1,2,\dots,J$ e assume valores 1 ou 0, dependendo se o argumento é verdadeiro ou não.

Considerando a função de dados aumentados dada em (3.1) e as destituições a priori definidas no capítulo 2 e considerando uma distribuição normal para o parâmetro de dificuldade, temos que as distribuições condicionais completas para o amostrador de Gibbs são dadas por (Johnson e Albert, 1999):

$$1) Z_{j,i} | \theta_j, \alpha_i, b_i, u_{j,i} \propto N(\alpha_i \theta_j - b_i) I(Z_{j,i}^*, u_{j,i}) \quad i=1,2,\dots,I \text{ e } j=1,2,\dots,J \quad (3.5)$$

$$2) \theta_j | Z_j, a_i, b_i, y_{j,i} \propto N(m_{\theta_j}, v_{\theta_j}) \quad j=1, 2, \dots, J \quad (3.6)$$

$$\text{Com } m_{\theta_j} = \frac{\sum_{i=1}^I a_i (Z_{j,i} + b_i)}{\sum_{i=1}^I a_i^2 + 1} \text{ e } v_{\theta_j} = \frac{1}{\sum_{i=1}^I a_i^2 + 1}, \quad j=1, 2, \dots, J$$

$$3) a_i, b_i | \theta, Z_i, y_{j,i} \propto N_2(m_{a_i, b_i}, v_{a_i, b_i}) \quad i=1, 2, \dots, I \quad (3.7)$$

$$\text{Com } m_{a_i, b_i} = [W^T W + \Sigma_{a_i, b_i}^{-1}] [W^T Z_i + \Sigma_{a_i, b_i}^{-1} \mu_{a_i, b_i}] \quad v_{a_i, b_i} = [W^T W + \Sigma_{a_i, b_i}^{-1}]^{-1}$$

$$\text{e } \mu_{a_i, b_i} = \begin{pmatrix} \mu_{a_i} \\ 0 \end{pmatrix}, \Sigma_{a_i, b_i} = \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{bmatrix} \text{ e } W = (\theta - 1) \text{ e } W_j^T = (\theta_j, -1), \quad j=1, 2, \dots, J$$

Assim a distribuição posteriori conjunta, considerando as prioris dadas no capítulo 2, é dada por:

$$f(Z, \theta, a, b | u_{j,i}) \propto \exp \left[-\frac{1}{2} \left(\sum_{j=1}^J \sum_{i=1}^I [Z_{j,i} - (a_i \theta_j - b_i)]^2 + \sum_{j=1}^J \theta_j^2 + \frac{1}{\sigma_b^2} \sum_{i=1}^I b_i^2 + \frac{1}{\sigma_a^2} \sum_{i=1}^I (a_i - \mu_a)^2 \right) \right] I(Z_{j,i}, u_{j,i}) \quad (3.8)$$

Para distribuição condicional de $Z_{j,i}$ é suficiente os termos correspondentes, assim, temos:

$$f(Z | \theta, a, b, u_{j,i}) \propto \exp \left[-\frac{1}{2} \left(\sum_{j=1}^J \sum_{i=1}^I (Z_{j,i} - (a_i \theta_j - b_i))^2 \right) \right] I(Z_{j,i}, u_{j,i}) \quad (3.9)$$

que corresponde ao núcleo de $i \times k$ distribuições normal truncada univariada para cada $Z_{j,i}$ dependendo do sinal de $Z_{j,i}$.

Finalmente para se obter as distribuições completas para θ e para os parâmetros a e b do item, devemos, da mesma forma anterior, considerar os termos correspondentes. Logo:

Para θ , temos:

$$f(\theta | Z, \alpha, b, u_{j,i}) \propto \exp \left[-\frac{1}{2} \sum_{j=1}^J \left(\left(1 + \sum_{i=1}^I \alpha_i^2 \right) \theta_j^2 - 2\theta_j \sum_{i=1}^I \alpha_i (b_i + z_{j,i}) \right) \right] \quad (3.10)$$

que corresponde ao núcleo de j distribuições normal univariadas para θ_j com média e desvio padrão dados, os quais são obtidos completando quadrados convenientemente.

Para a e b temos:

$$f(a, b | Z, \theta, u_{j,i}) \propto \exp \left[-\frac{1}{2} \left(\sum_{j=1}^J \sum_{i=1}^I \left[-2Z_{j,i} (a_i \theta_j - b_i) + (a_i \theta_j - b_i)^2 + \frac{1}{\sigma_a^2} \sum_{i=1}^I (a_i - \mu_a)^2 \right] \right) \right] \quad (3.11)$$

Que corresponde ao núcleo de i distribuições normal bivariadas para a e b .

Com as distribuições condicionais dadas anteriormente, é possível a obtenção de uma amostragem usando o amostrador de Gibbs e considerando que foram introduzidas variáveis latentes auxiliares, tem se uma jeito de ser feita no programa Winbugs, porém isso não é possível de certo feito de forma direta, requerendo uma correta especificação para as variáveis indicadoras como visto em Bazan (2005).

3.2.2 Modelo probito-assimétrico normal

O algoritmo abaixo (Bazan, 2005) foi desenvolvido para o caso de distribuição assimétrica e é similar ao dado em Albert (1992) para um conjunto de dados aumentados.

O modelo TRI probito assimétrico, envolvendo um total de I itens e J indivíduos, com $u_{j,i} \sim \text{Bernoulli}(p_{j,i})$ e $p_{j,i} = \phi(a_i \theta_j - b_i)$ é equivalente definido considerando que $u_{j,i} = 1$ se $Z_{j,i} > 0$ e $u_{j,i} = 0$ se $Z_{j,i} \leq 0$ em que $Z_{j,i} \sim NA(a_i \theta_j - b_i, 1, -\lambda_i)$, $j=1,2,\dots, J$ e $i=1,2,\dots, I$.

Para verificar usa-se o fato que $1 - F_{-\lambda}(z) = F_{\lambda}(-z)$ (Azzalini, 1985) e é similar a dada em Albert (1992).

O motivo de se introduzir a variável auxiliar $Z_{j,i}$ é para se evitar trabalhar com a verossimilhança tipo Bernoulli. Para a função de verossimilhança aumentada para o modelo TRI probito assimétrico dada abaixo, considere o vetor $Z = (z_{1,1}, \dots, z_{J,I})$, os parâmetros dos itens a e b e o parâmetro de habilidade θ .

$$L(\theta, a, b, \lambda | u, z) = \prod_{i=1}^I \prod_{j=1}^J [f_{\varphi}(z_{j,i}) I(z_{j,i}, u_{j,i})] \quad (3.12)$$

Em que $\varphi = (a_i \theta_j - b_i, 1, -\lambda_i)$ e $I^*(z_{j,i}, u_{j,i}) = I(Z_{j,i} > 0) I(u_{j,i} = 1) + I(Z_{j,i} \leq 0) I(u_{j,i} = 0)$

$I = 1, 2, \dots, I$ e $j = 1, 2, \dots, J$ e assume valores 1 ou 0, dependendo se o argumento é verdadeiro ou não.

O modelo é correspondente a densidade condicional $p(u_{j,i} | Z_{j,i})$ usado para derivação de (3.3).

O modelo TRI probito normal assimétrico para um teste contento I itens e aplicados a J indivíduos com $u_{j,i} \sim \text{Bernoulli}(p_{j,i})$ e $p_{j,i} = \phi_{NA}(a_i \theta_j - b_i, \lambda_i)$ é equivalente a considerar que:

$u_{j,i} = 1$ se $Z_{j,i}^* > 0$ e $u_{j,i} = 0$ se $Z_{j,i}^* \leq 0$, em que $Z_{j,i}^* \sim N(-\delta_i V_{j,i} + a_i \theta_j - b_i, 1 - \delta_i^2)$ e $V_{j,i} \sim HN(0, 1)$, $j = 1, 2, \dots, J$ e $i = 1, 2, \dots, I$

Pode-se verificar definindo $Z_{j,i}^* = \frac{Z_{j,i}}{V_{j,i}}$ como sendo a representação estocástica para a distribuição normal assimétrica (Henze, 1986) para $Z_{j,i}$ em que $Z_{j,i} \sim NA(a_i \theta_j - b_i, 1, -\lambda_i)$, $j = 1, 2, \dots, J$ e $i = 1, 2, \dots, I$, e então a prova segue como dada no para apêndice A para a distribuição normal assimétrica.

Assim, a função de verossimilhança aumentada envolvendo as variáveis latentes auxiliares dadas pelo vetor $Z^* = (z_{1,1}^*, \dots, z_{J,I}^*)$ e pelos parâmetros dos itens, assim como das habilidades é definida por:

$$L(\theta, a, b, d | Z_{j,i}^*, u_{j,i}) = \prod_{j=1}^J \prod_{i=1}^I \phi(Z_{j,i}^*, -d_i V_{j,i} + a_i \theta_j - b_i, 1 - d_i^2) I(Z_{j,i}^*, u_{j,i}) \phi(V_{j,i}, 0, 1) I(V_{j,i} > 0) \quad (3.13)$$

Considerando a função de verossimilhança dada em (3.13) e as distribuições a priori especificadas para os parâmetros dos itens e das habilidades dadas no capítulo 2, temos no que segue as distribuições condicionais completas para o amostrador de Gibbs (Bazan, 2005).

$$1) Z_{j,i} | \theta_j, a_i, b_i, u_{j,i}, \lambda_i \propto \phi_{NA}(Z_{j,i}, a_i \theta_j - b_i, 1, -\lambda_i) I(Z_{j,i}, u_{j,i}) \quad i=1, 2, \dots, I \text{ e } j=1, 2, \dots, J \quad (3.14)$$

$$2) \theta_j | Z_j, a_i, b_i, y_{j,i}, \lambda \propto \phi(\theta_j, m_{\theta_j}, v_{\theta_j}) \psi(\theta_j) \quad j=1, 2, \dots, J \quad (3.15)$$

$$\text{Com } m_{\theta_j} = \frac{\sum_{i=1}^I a_i (Z_{j,i} + b_i)}{\sum_{i=1}^I a_i^2 + 1} \text{ e } v_{\theta_j} = \frac{1}{\sum_{i=1}^I a_i^2 + 1}, \quad j=1, 2, \dots, J \quad \text{e}$$

$$\psi(\theta_j) = \prod_{i=1}^I \Phi(-\lambda_i Z_{j,i} - (a_i \theta_j - b_i)) \Phi(\kappa \theta_j)$$

$$3) a_i, b_i | \theta, Z_i, y_{j,i}, \lambda_i \propto N_2(a_i, b_i; m_{a_i, b_i}; v_{a_i, b_i}) \psi(a_i, b_i) \quad i=1, 2, \dots, I \text{ e} \quad (3.16)$$

$$\psi(a_i, b_i) = \prod_{j=1}^J \Phi(-\lambda_j Z_{j,i} - W_j^T (a_i, b_i))$$

$$\text{Com } m_{a_i, b_i} = [W^T W + \Sigma_{a_i, b_i}^{-1}]^{-1} [W^T Z_i + \Sigma_{a_i, b_i}^{-1} \mu_{a_i, b_i}] \quad v_{a_i, b_i} = [W^T W + \Sigma_{a_i, b_i}^{-1}]^{-1}$$

$$\text{e } \mu_{a_i, b_i} = \begin{pmatrix} \mu_{a_i} \\ 0 \end{pmatrix}, \Sigma_{a_i, b_i} = \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{bmatrix} \text{ e } W = (\theta, -1) \text{ e } W_j^T = (\theta_j, -1), \quad j=1, 2, \dots, J$$

$$4) \lambda_i | \theta, a, b, Z_i, u_{j,i} \propto \phi(\lambda_i) \psi(\lambda_i) \text{ com } \psi(\lambda_i) = \prod_{j=1}^J \Phi(-\lambda_j Z_{j,i} - (a_i \theta_j - b_i)) \Phi(\omega \lambda_i) \quad (3.17)$$

3.3 Implementação no Winbugs

O modelo probito normal (PN) pode ser escrito de forma hierárquica como dada em Albert(1992) como:

$$Z_{j,i} \mid u_{j,i}, \theta_j, \alpha_i, b_i \sim N(\alpha_i \theta_j - b_i, 1) I(z_{j,i}, u_{j,i})$$

$$\theta_j \sim N(0.0, 1.0), j=1, 2, \dots, J$$

$$\alpha_i \sim N(1.0, 2.0) I(0,)$$

$$b_i \sim N(0.0, 0.5)$$

O modelo probito assimétrico- normal pode ser escrito de forma hierárquica como dada em Bazan (2005) e usa o fato de poder ser escrito como uma combinação linear de uma normal padrão e uma half normal, assim temos:

$$Z_{j,i}^* \mid u_{j,i}, \theta_j, \alpha_i, b_i, d_i \sim N(\alpha_i \theta_j - b_i - d_i V_{j,i}, 1 - d_i^2) I(z_{j,i}^*, u_{j,i})$$

$$V_{j,i} \sim HN(0.0, 1.0)$$

$$d_i \sim U(-1.0, 1.0)$$

$$\theta_j \sim N(0.0, 1.0), j=1, 2, \dots, J$$

$$\alpha_i \sim N(1.0, 2.0) I(0,)$$

$$b_i \sim N(0.0, 0.5)$$

No modelo probito-normal assimétrico (PNA), cuja a assimetria é considerada na distribuição das habilidades a hierarquia para a sua correta implementação também leva em consideração que o mesmo pode ser escrito como uma combinação linear de uma normal padrão e de uma half normal e é dado como segue:

$$Z_{j,i} \mid u_{j,i}, \theta_j, \alpha_i, b_i \sim N(\alpha_i \theta_j - b_i, 1) I(z_{j,i}, u_{j,i})$$

$$\theta_j \sim NA(0.0, 1.0, \alpha), j=1, 2, \dots, J$$

$$\alpha_i \sim N(1.0, 2.0) I(0,)$$

$$b_i \sim N(0.0, 2.0)$$

$$\alpha \sim U(-1.0, 1.0)$$

3.4 Programas para implementação no Winbugs

O programa Winbugs (Bayesian Analysis Using Gibbs Sampler) é um software livre, desenvolvido para se fazer inferência bayesina e que vem sendo usado pela comunidade estatística devido a sua facilidade de manuseio e a grande variedade de problemas que ele consegue resolver utilizando o algoritmo de Gibbs.

A seguir serão apresentados os programas que foram implementados para cada um dos modelos abordados anteriormente e que já tiveram sua implementação em Bazan (2005).

```

Model;
{ #modelo PN
for(j in 1 : J) {
for(l in 1 : l) {
z[j ,i] ~ dnorm(m[j , i], 1.0)|(lo[u[j , i]+1],up[u[ j , i]+1])
m[j, i] <- a[i] * teta[j] - b[i]
}
}
for( i in 1 : l ) {
b[i] ~ dnorm( 0.0, 0.5)
teta[i] ~ dnorm( 1.0, 2.0)|(0,)
}
for( j in 1 : J ) {
teta[j] ~ dnorm( 0.0, 1.0)
}
lo[1]<-50
up[1]<-0

```

```
lo[2]<-0
up[2]<-50 }
```

Valores iniciais

```
list(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0), a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1))`
```

```
Model;
{#modelo PAN
for( j in 1 : J ) {
for( i in 1 : I ) {
z[j ,i] ~ dnorm(muz[j , i], preczs[i])l(lo[u[ j , i]+1],up[u[ j , i]+1])
v[j ,i] ~ dnorm(0.0, 1.0)l(0,)
muz[j , i] <- a[i] * teta[j] - b[i]-d[i]*v[j,i]
}
}
for(i in 1 : I ) {
b[i] ~ dnorm( 0.0, 0.5)
a[i] ~ dnorm( 1.0, 2.0)l(0,)
d[i] ~ dunif(-1.0, 1.0)
preczs[i]<-1/(1-pow(d[i],2))
lambda[i]<-d[i]*sqrt(preczs[i])
}
for( j in 1 : J ) {
teta[j] ~ dnorm( 0.0, 1.0)
}
lo[1]<-50
up[1]<-0
lo[2]<-0
up[2]<-50 }
```

Valores iniciais

```
list(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0), a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1),
d=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0))
```

```
Model;
{#Modelo PNA
for(j in 1 : J ) {
for( i in 1 : I ) {
z[j,i] ~ dnorm(m[ j, i], 1.0)|(lo[u[ j, i]+1],up[u[ j, i]+1])
m[j,i] <- a[i] * teta[j] - b[i]
}
}
for( i in 1 : I ) {
b[i] ~ dnorm( 0.0, 0.5)
a[i] ~ dnorm( 1.0, 2.0)|(0,)
}
for( j in 1 : J ) {
teta[j]<-d*v[j]+sqrt(1-pow(d,2))*w[j]
v[j] ~ dnorm( 0.0, 1.0)|(0,)
w[j] ~dnorm( 0.0, 1.0)
}
d ~ dunif(-1.0 , 1.0)
lo[1]<--50
up[1]<-0
lo[2]<-0
up[2]<-50 }
```


Valores iniciais

`list(b=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0), a=c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1))`

Capítulo 4

Aplicações dos modelos

4.1 Introdução

Neste capítulo será feita a aplicação dos modelos abordados no capítulo anterior a um conjunto de dados obtidos das respostas que 47 alunos deram a conjunto de 15 itens de múltipla escolha, tendo somente uma resposta correta, portanto sendo corrigidos de forma dicotômica.

Estes alunos, oriundos da classe média, passaram 5 meses estudando num tradicional curso pré-vestibular da cidade de São Paulo e se preparavam para prestar vestibular em universidades públicas e privadas para o curso de administração, como USP, GV, IBMEC, etc. Estas instituições apresentam grande concorrência em seus vestibulares, o que requer do candidato um preparo muito maior.

Para cada um dos três modelos abordados nesta dissertação serão apresentadas as estatísticas obtidas após o processo estimação utilizando, como já destacado, o programa Winbugs, assim como os gráficos importantes para análise e interpretação dos parâmetros dos itens e das habilidades dos respondentes.

No apêndice B será apresentada uma tabela com proporções de acertos em cada item e também gráficos do tipo histograma e dot-plot com a frequência de estudantes em cada pontuação.

4.2 Valores iniciais

Na implementação dos modelos no Winbugs pode-se ou não considerar valores iniciais para os parâmetros a serem estimados. Caso não seja atribuído nenhum valor inicial, o próprio sistema gera os valores. Entretanto, isso torna o processo de estimação muito lento e dependo do computador que se utiliza, até inviável. Dessa forma serão considerados valores iniciais para os parâmetros dos itens, com 1 e 0 para a_i e b_i , $i = 1, 2, \dots, I$, respectivamente.

Tanto nos modelos simétricos como nos assimétricos, estes valores foram fixados com base nas médias das distribuições a priori estabelecidas no capítulo 3.

Em particular, no modelo probito assimétrico-normal em que se estabelece um parâmetro d_i que mede a assimetria, o valor inicial adotado foi 0, uma vez que este valor corresponde a média da distribuição a priori para d_i .

Da mesma forma que em Bazan (2005), valores iniciais para θ_j (habilidades dos respondentes) e $Z_{j,i}^*$ e $V_{j,i}$ foram aleatoriamente gerados pelo sistema, porém poderiam ser atribuídos valores iniciais com base nas suas distribuições a priori.

Serão considerados em todos os casos um total de 25000 interações com saltos de 5 em 5 e sendo descartadas as 1000 interações iniciais para assim diminuir possíveis correlações.

4.3 Aplicação dos dados no modelo probito normal

4.3.1 Estimativas obtidas para os parâmetros dos itens

Parâmetro de discriminação a

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
a[1]	0.8389	0.3795	0.01205	0.2162	0.7975	1.711	5000
a[2]	1.488	0.4558	0.01258	0.6546	1.453	2.464	5000
a[3]	1.01	0.4049	0.01534	0.3101	0.9776	1.861	5000
a[4]	0.9217	0.3651	0.01242	0.298	0.8989	1.712	5000
a[5]	0.9836	0.3757	0.01216	0.3126	0.9671	1.757	5000
a[6]	0.7606	0.3233	0.01353	0.2336	0.7296	1.482	5000
a[7]	0.8118	0.3624	0.01356	0.1878	0.7773	1.583	5000
a[8]	0.8874	0.3732	0.01098	0.2208	0.8629	1.683	5000
a[9]	1.084	0.3916	0.01435	0.3992	1.049	1.892	5000
a[10]	0.4413	0.3064	0.01051	0.02381	0.3883	1.167	5000
a[11]	1.239	0.4823	0.02008	0.3878	1.21	2.288	5000
a[12]	0.8208	0.3838	0.01177	0.1807	0.784	1.635	5000
a[13]	1.014	0.4359	0.01441	0.2746	0.9756	1.975	5000
a[14]	1.258	0.4607	0.01724	0.4648	1.228	2.287	5000
a[15]	1.183	0.4754	0.01662	0.3355	1.151	2.213	5000

Tabela 4.1 Estimativas para o parâmetro de discriminação

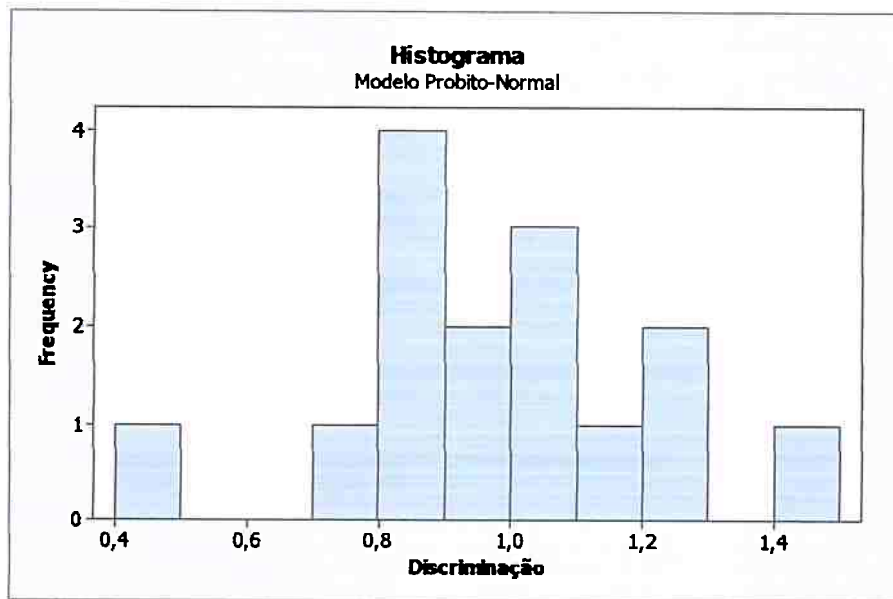


Figura 4.1 Histograma dos valores parâmetro de discriminação

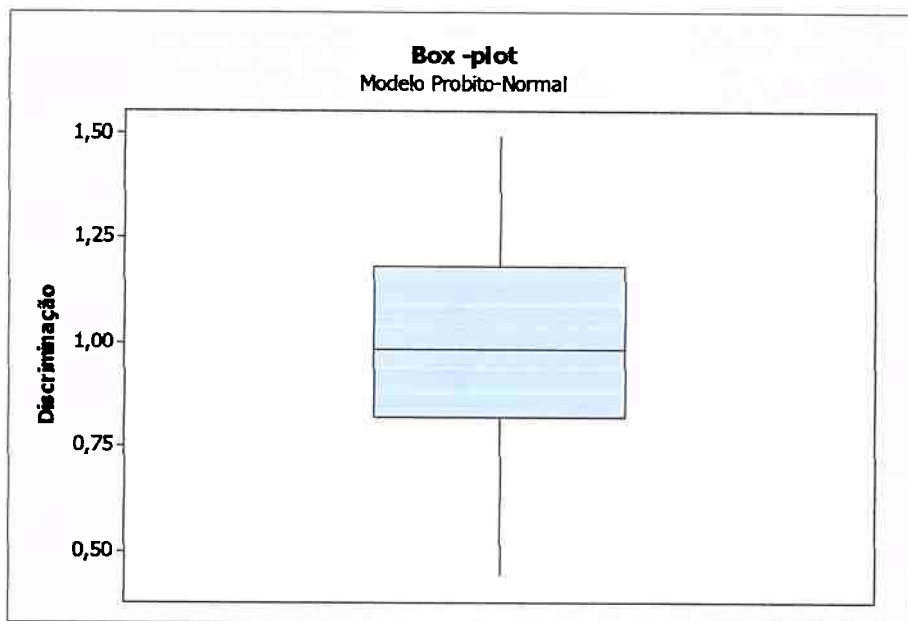


Figura 4.2 Box-Plot para o parâmetro de discriminação

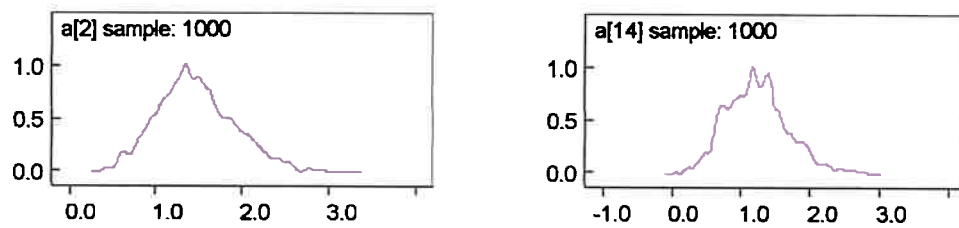


Figura 4.3 Densidades a posteriori do parâmetro de discriminação dos itens 2 e 14

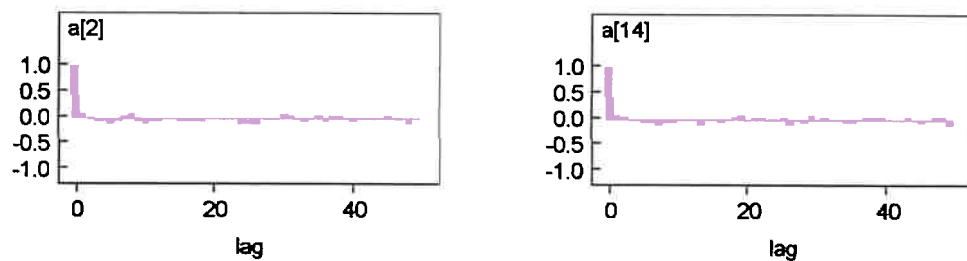


Figura 4.4 Gráficos das correlações para o parâmetro de discriminação dos itens 2 e 14

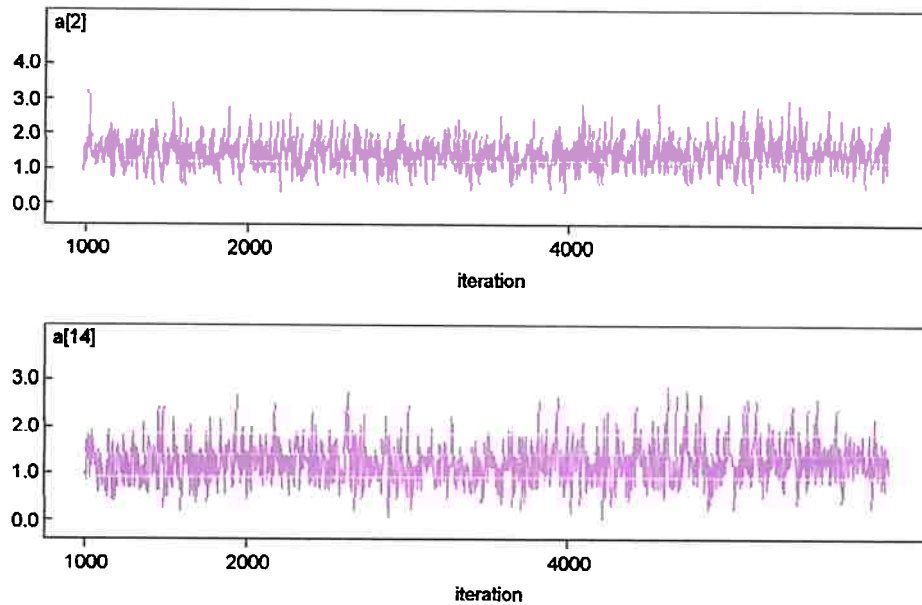


Figura 4.5 Histórico de convergência para os parâmetro de discriminação dos itens 2 e 14

Como se pôde observar na tabela 4.1, a mediana a posteriori ficou em torno de 1, sendo que os itens com maior capacidade de discriminação foram os de número 2, e 14. Já os que apresentaram menor discriminação foram os de número 6 e 10.

Analisando os gráficos das densidades a posteriori pode-se afirmar que apresentaram características de normalidade com média próxima de 1.

Da mesma forma os gráficos de correlações mostram que durante a estimação quase não se observa correlação nos resultados finais enquanto os gráficos do histórico mostram que houve boa convergência devido a sua oscilação.

Parâmetro de dificuldade b

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
b[2]	-0.399	0.3301	0.01162	-1.046	-0.3819	0.226	5000
b[3]	-0.9012	0.2971	0.009718	-1.538	-0.8927	-0.3935	5000
b[4]	-0.1882	0.2353	0.008059	-0.6355	-0.1936	0.2606	5000
b[5]	-0.3323	0.2484	0.008144	-0.8403	-0.3277	0.1662	5000
b[6]	0.06346	0.2271	0.008348	-0.3725	0.05553	0.5294	5000
b[7]	-0.7817	0.2533	0.008818	-1.294	-0.7654	-0.314	5000
b[8]	-0.7865	0.2581	0.006779	-1.313	-0.7788	-0.3243	5000
b[9]	-0.3508	0.2666	0.009196	-0.9077	-0.3359	0.1458	5000
b[10]	-1.242	0.273	0.008607	-1.799	-1.223	-0.7465	5000
b[11]	-2.186	0.5117	0.01672	-3.308	-2.125	-1.291	5000
b[12]	-1.367	0.3219	0.00922	-2.042	-1.353	-0.7755	5000
b[13]	-1.346	0.3434	0.01063	-2.105	-1.326	-0.7462	5000
b[14]	-1.31	0.3485	0.01101	-2.083	-1.266	-0.6939	5000
b[15]	-1.411	0.3811	0.01486	-2.275	-1.378	-0.7396	5000

Tabela 4.2 Estimativas para o parâmetro de dificuldade

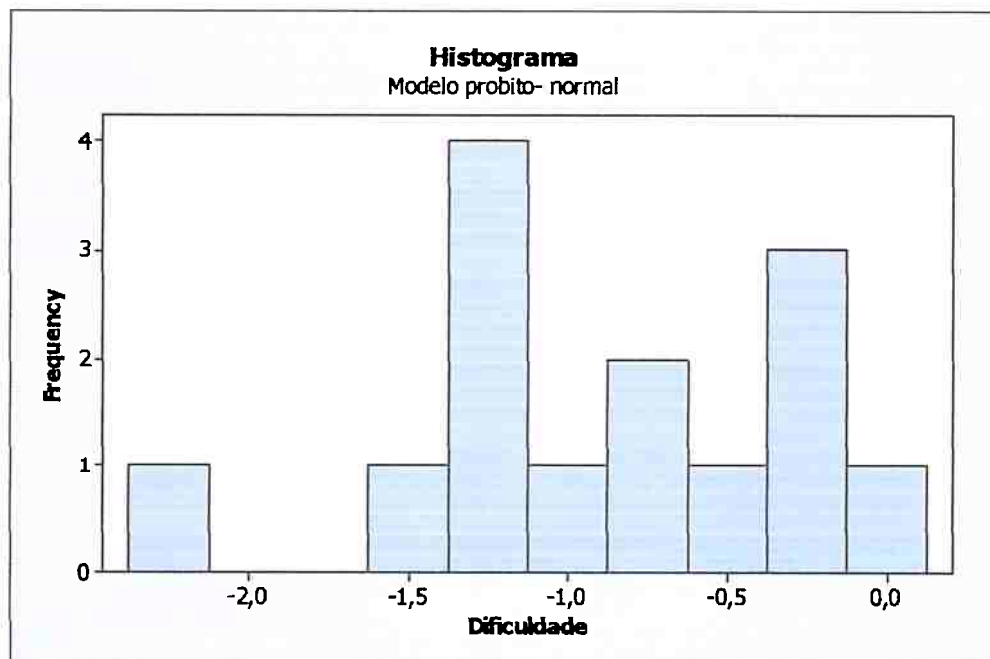


Figura 4.6 Histograma dos valores do parâmetro de dificuldade

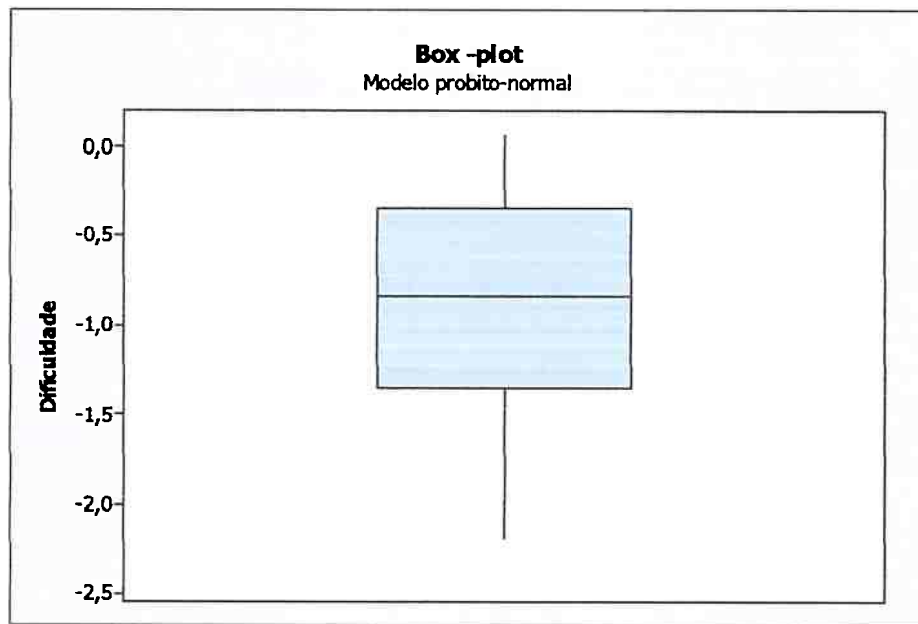


Figura 4.7 Box-Plot dos valores parâmetro de dificuldade

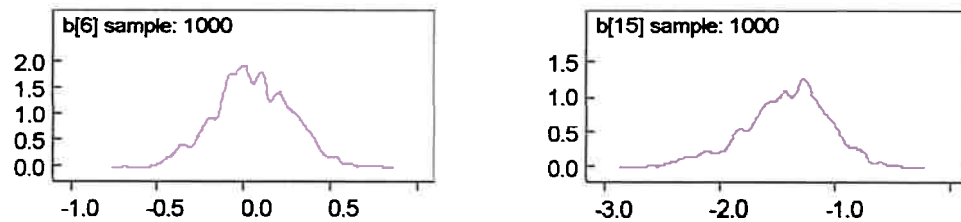


Figura 4.8 Densidades a posteriori do parâmetro de dificuldade dos itens 6 e 15

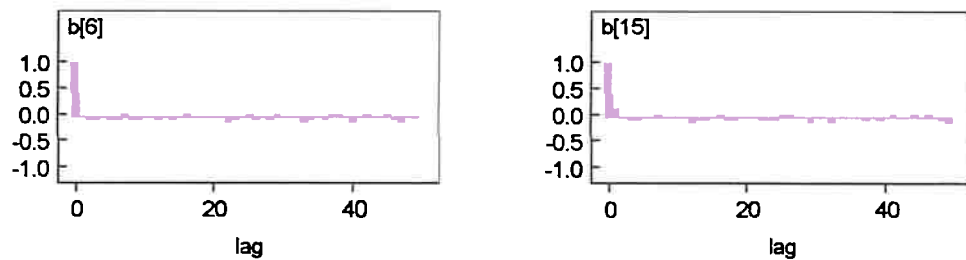


Figura 4.9 Gráfico das correlações para o parâmetro de dificuldade dos itens 6 e 15

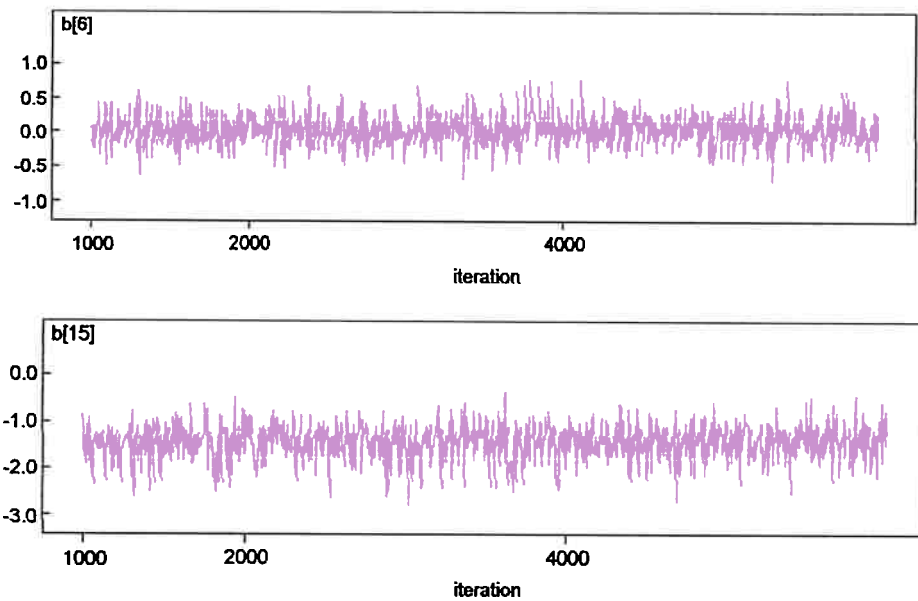


Figura 4.9 Histórico de convergência para o parâmetro de dificuldade dos itens 6 e 15

Analisando a tabela (4.2), os itens, de uma forma geral, não apresentam dificuldades para os respondentes, uma vez que os valores obtidos foram baixos. Isso pode ser confirmado pela alta proporção de acertos como é apresentado no apêndice B.

Os itens que apresentaram maior dificuldade foram os de números 4 e 6 e os de menor dificuldade foram os de número 11 e 15.

Analisando os gráficos das densidades a posteriori, pode-se concluir que nos dois itens a suposição de normalidade se verifica, porém verifica-se que no item 15 a média ficou distante do valor estabelecido como média da priori.

Já os gráficos das correlações apresentam boa variação, o caracteriza boa convergência.

Parâmetro de habilidade $\theta = u$

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
u[1]	-0.5669	0.3789	0.01069	-1.363	-0.5535	0.144	5000
u[2]	1.335	0.6606	0.01926	0.1966	1.274	2.817	5000
u[3]	-0.9366	0.387	0.01355	-1.71	-0.9253	-0.2272	5000
u[4]	-0.7155	0.372	0.01184	-1.465	-0.7076	-0,009767	5000
u[5]	-0.2505	0.3813	0.01303	-0.9344	-0.2696	0.5758	5000
u[6]	1.302	0.6528	0.02203	0.1675	1.274	2.732	5000
u[7]	-0.4749	0.3735	0.01312	-1.17	-0.4774	0.2583	5000
u[8]	0.2339	0.4501	0.01383	-0.6061	0.2297	1.192	5000
u[9]	0.8588	0.5713	0.01209	-0.1376	0.8155	2.098	5000
u[10]	0.8498	0.5771	0.02075	-0.1828	0.819	2.072	5000
u[11]	1.32	0.6604	0.02237	0.2096	1.279	2.731	5000
u[12]	0.2752	0.4696	0.01758	-0.5542	0.2531	1.328	5000
u[13]	-0.3641	0.3848	0.01287	-1.116	-0.371	0.3803	5000
u[14]	0.8578	0.5581	0.01896	-0.1336	0.8253	2.113	5000
u[15]	1.298	0.6464	0.02466	0.1661	1.258	2.673	5000
u[16]	0.2167	0.4667	0.01357	-0.6261	0.169	1.181	5000
u[17]	-0.1846	0.3939	0.01127	-0.9323	-0.1746	0.5878	5000
u[18]	0.1677	0.4222	0.01507	-0.6494	0.162	1.032	5000
u[19]	0.3686	0.4625	0.01546	-0.4711	0.337	1.358	5000
u[20]	1.308	0.6452	0.02137	0.1785	1.268	2.761	5000
u[21]	-0.03768	0.4015	0.01141	-0.7788	-0.04983	0.8276	5000
u[22]	0.8549	0.5656	0.01652	-0.09119	0.8239	2.083	5000
u[23]	0.9514	0.5795	0.01817	-0.08231	0.914	2.187	5000
u[24]	0.7328	0.5334	0.01576	-0.1658	0.6766	1.892	5000
u[25]	-0.4655	0.3734	0.01215	-1.169	-0.4707	0.2896	5000
u[26]	0.4543	0.4747	0.01525	-0.4175	0.4345	1.49	5000
u[27]	-0.1843	0.3745	0.01299	-0.8595	-0.1916	0.6197	5000
u[28]	-0.5811	0.3632	0.01075	-1.301	-0.5818	0.1331	5000
u[29]	0.7124	0.5326	0.0159	-0.1836	0.6748	1.9	5000
u[30]	-0.4797	0.3792	0.01181	-1.225	-0.4815	0.276	5000
u[31]	0.1474	0.4336	0.01397	-0.6102	0.1464	1.072	5000
u[32]	0.7573	0.5338	0.01772	-0.1473	0.7251	1.968	5000
u[33]	-0.679	0.3588	0.01278	-1.388	-0.6765	-0.001777	5000
u[34]	0.348	0.4585	0.01425	-0.4881	0.3348	1.3	5000
u[35]	-1.34	0.4174	0.01737	-2.181	-1.308	-0.5925	5000
u[36]	-0.03137	0.4046	0.01525	-0.8139	-0.03681	0.7919	5000
u[37]	1.318	0.6307	0.02076	0.2282	1.288	2.664	5000
u[38]	-1.552	0.4378	0.01318	-2.499	-1.545	-0.7446	5000
u[39]	-0.00323	0.3974	0.01291	-0.7555	-0.02502	0.8036	5000
u[40]	-0.426	0.379	0.01176	-1.165	-0.4209	0.3312	5000
u[41]	-0.1655	0.3899	0.01292	-0.9022	-0.1689	0.6213	5000
u[42]	0.3493	0.4723	0.01309	-0.5056	0.3127	1.326	5000
u[43]	-0.5898	0.3816	0.01282	-1.382	-0.5812	0.1268	5000
u[44]	0.1329	0.4318	0.01342	-0.6273	0.1137	1.033	5000
u[45]	-0.3795	0.3802	0.01372	-1.102	-0.3781	0.3868	5000
u[46]	1.337	0.6452	0.0192	0.2057	1.291	2.812	5000
u[47]	-0.6625	0.3853	0.01229	-1.462	-0.662	0.1212	5000

Tabela 4.3 Estimativas para o parâmetro de habilidades

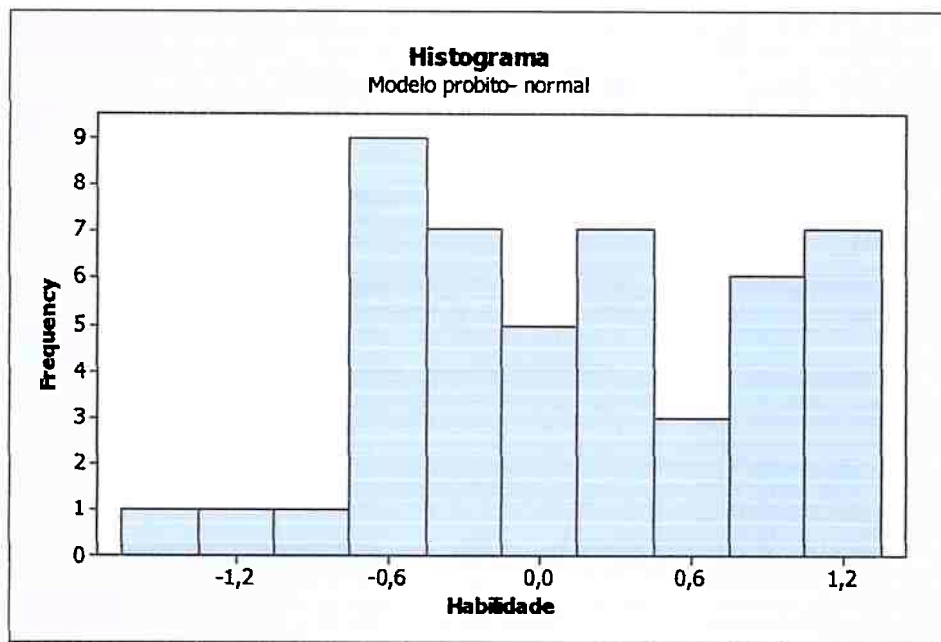


Figura 4.10 Histograma dos valores do parâmetro das habilidades

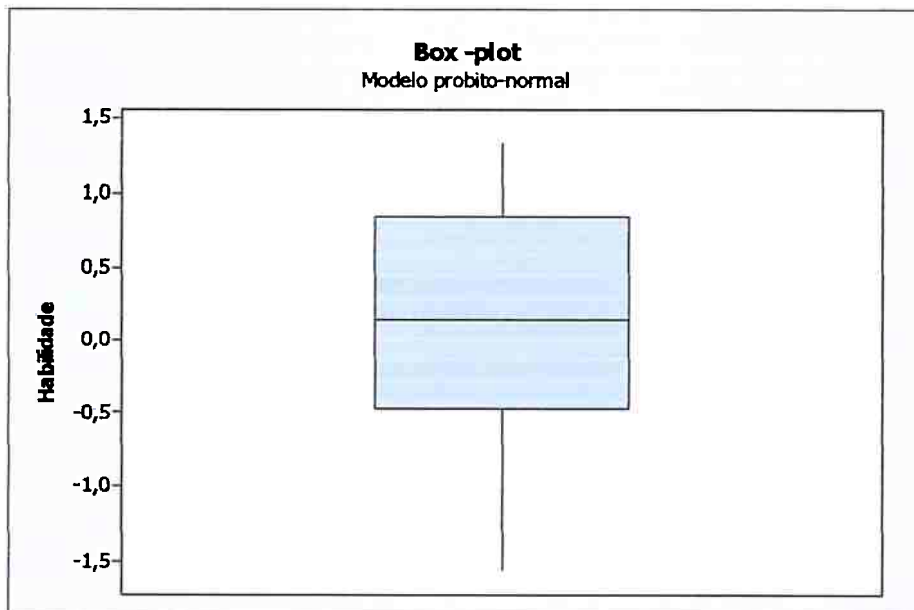


Figura 4.11 Box - Plot dos valores do parâmetro das habilidades

De uma forma geral, a tabela (4.3) apresenta valores positivos e negativos para as habilidades dos respondentes, mas em alguns casos são identificados indivíduos com grande habilidade (indivíduos 2 e 46). Estes indivíduos tiveram escores com valores iguais a 15, ou seja, acertaram todos os itens.

Alguns indivíduos apresentaram baixa habilidade (indivíduos 35 e 38) com escores iguais a 4.

Aqui cabe um comentário que talvez se torna útil para a instituição na qual esses alunos estudam, pois, por uma questão pedagógica, os indivíduos são separados em salas usando como critério os acertos que os estudantes obtêm em seus testes.

Pode-se, portanto, identificar alunos com habilidades semelhantes para se formar grupos de estudo com um enfoque pedagógico mais adequado, ou seja, respeitando-se as habilidades de cada um e sua capacidade para assimilar as informações passadas em sala de aula.

4.4 Modelo probito-assimétrico normal

Parâmetro de discriminação a

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
a[1]	0.7971	0.3817	0.007224	0.1529	0.7655	1.662	5000
a[2]	1.402	0.4471	0.009123	0.629	1.369	2.373	5000
a[3]	0.9577	0.406	0.008175	0.2835	0.9092	1.871	5000
a[4]	0.8432	0.3431	0.006075	0.2715	0.8088	1.605	5000
a[5]	0.8857	0.3515	0.005827	0.2824	0.8517	1.675	5000
a[6]	0.6944	0.3052	0.005596	0.178	0.6696	1.366	5000
a[7]	0.7644	0.3445	0.005865	0.1752	0.7344	1.519	5000
a[8]	0.8414	0.3569	0.006301	0.2371	0.8038	1.648	5000
a[9]	1.018	0.3818	0.006601	0.3737	0.9789	1.838	5000
a[10]	0.4127	0.2838	0.005505	0.02079	0.3666	1.082	5000
a[11]	1.268	0.4931	0.01198	0.3827	1.249	2.29	5000
a[12]	0.7869	0.3981	0.006504	0.1432	0.7416	1.673	5000
a[13]	0.9802	0.4267	0.007775	0.248	0.9414	1.9	5000
a[14]	1.181	0.4558	0.00906	0.4057	1.141	2.2	5000
a[15]	1.177	0.4698	0.01044	0.3699	1.146	2.195	5000

Tabela (4.4) Estimativa para o parâmetro de discriminação

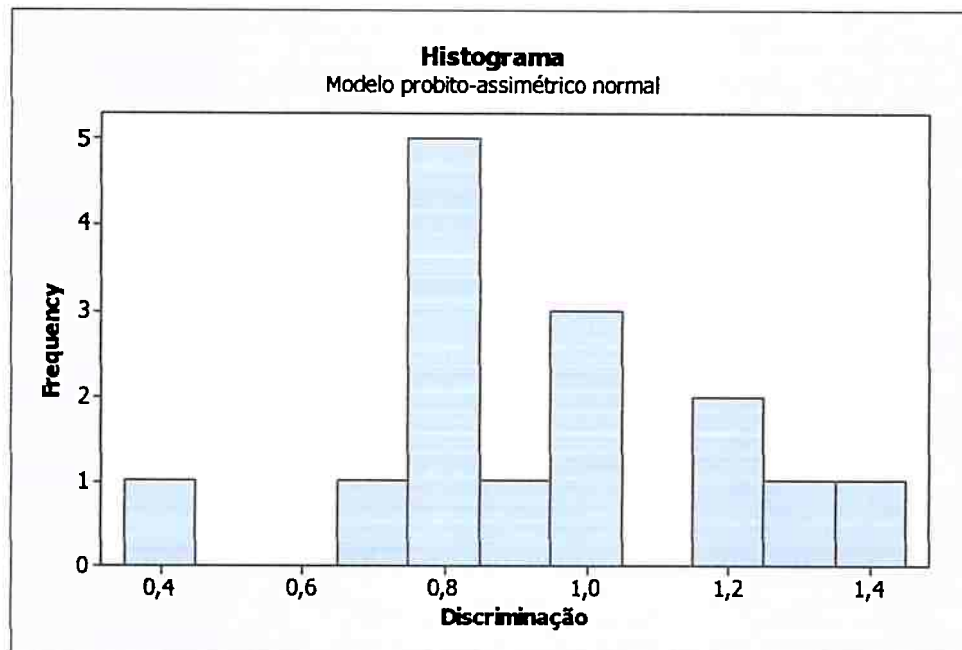


Figura 4.12 Histograma dos valores do parâmetro de discriminação

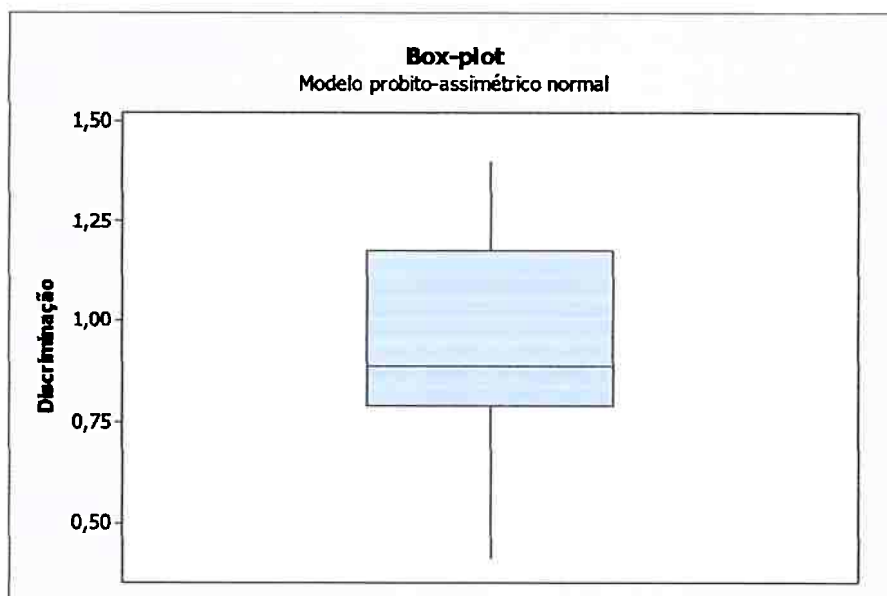


Figura 4.13 Box-Plot dos valores do parâmetro de discriminação

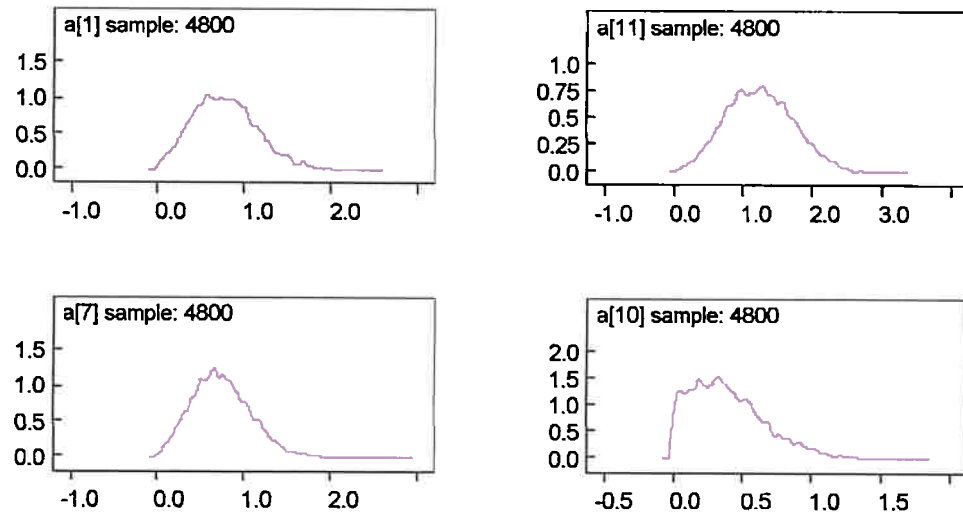


Figura 4.14 Densidades a posteriori para o parâmetro de dificuldade dos itens 1, 7, 10 e 11

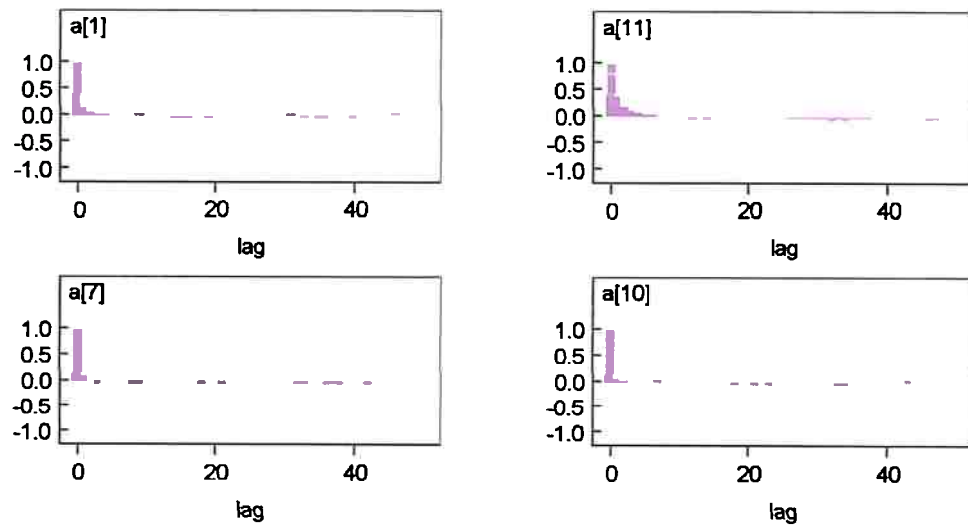


Figura 4.15 Gráficos das correlações para o parâmetro de dificuldade dos itens 1, 7, 10 e 11

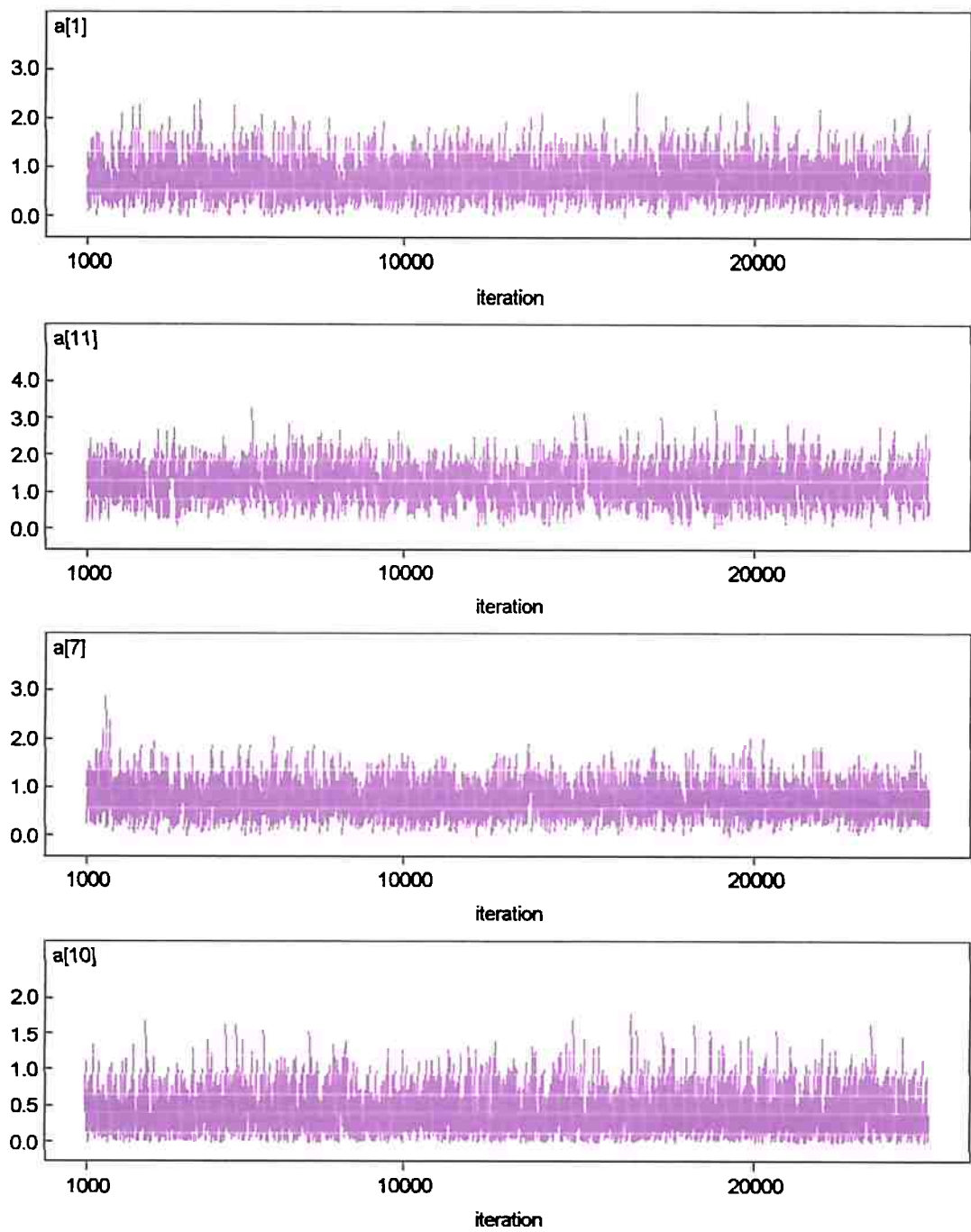


Figura 4.16 Histórico de convergência para o parâmetro de dificuldade dos itens 1,7,10 e 11

Da tabela (4.4) pode-se observar que, de uma forma geral, os valores obtidos para as estimativas estão próximos da média da priori estabelecida para a distribuição dos α_i . Neste modelo os itens que mais discriminaram foram os de número 2 e 11 e os que menos discriminaram foram os de número 6 e 12

Analisando os gráficos das densidades e das correlações, pode-se afirmar que tiveram comportamento apresentando normalidade com simetria com exceção do item 10, cuja distribuição a posteriori apresentou certa simetria.

Já os gráficos das correlações, assim como do histórico tiveram comportamento iguais aos comentados no modelo anterior.

Parâmetro de dificuldade b

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
b[1]	-0.9224	0.5078	0.01362	-1.815	-0.9635	0.1057	5000
b[2]	-0.3256	0.5207	0.01297	-1.298	-0.3205	0.6062	5000
b[3]	-0.7973	0.4954	0.01165	-1.714	-0.8222	0.1622	5000
b[4]	-0.1764	0.4472	0.01022	-0.9722	-0.1728	0.6637	5000
b[5]	-0.3148	0.4469	0.009393	-1.092	-0.3244	0.5313	5000
b[6]	0.04233	0.4292	0.01078	-0.7402	0.04619	0.8247	5000
b[7]	-0.6808	0.4649	0.01076	-1.478	-0.7123	0.2314	5000
b[8]	-0.7037	0.4791	0.01074	-1.551	-0.7326	0.2393	5000
b[9]	-0.3196	0.4515	0.01055	-1.139	-0.3232	0.5253	5000
b[10]	-1.128	0.4823	0.01073	-1.968	-1.174	-0.1297	5000
b[11]	-1.865	0.7118	0.02078	-3.305	-1.855	-0.5169	5000
b[12]	-1.198	0.543	0.01207	-2.189	-1.229	-0.1115	5000
b[13]	-1.156	0.5567	0.015	-2.149	-1.202	-0.05632	5000
b[14]	-1.134	0.5649	0.01448	-2.188	-1.156	-0.04582	5000
b[15]	-1.26	0.5863	0.01347	-2.371	-1.283	-0.1003	5000

Tabela 4.5 Estimativas para o parâmetro de dificuldade

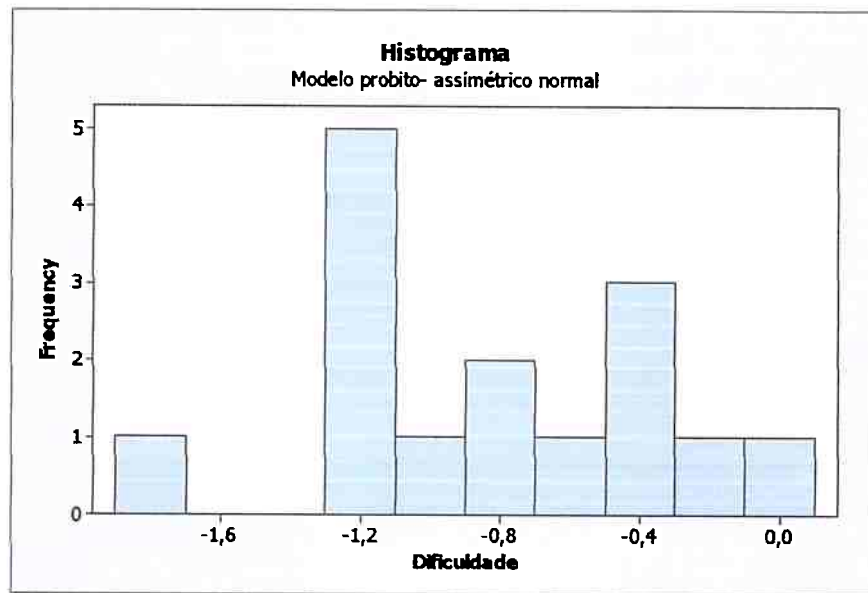


Figura 4.17 Histograma dos valores do parâmetro de dificuldade

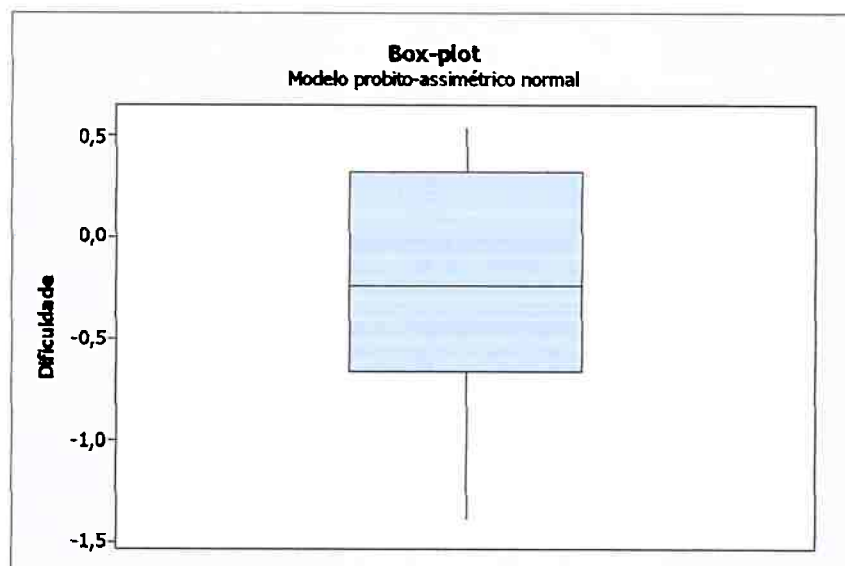


Figura 4.18 Box-Plot para os valores do parâmetro de dificuldade

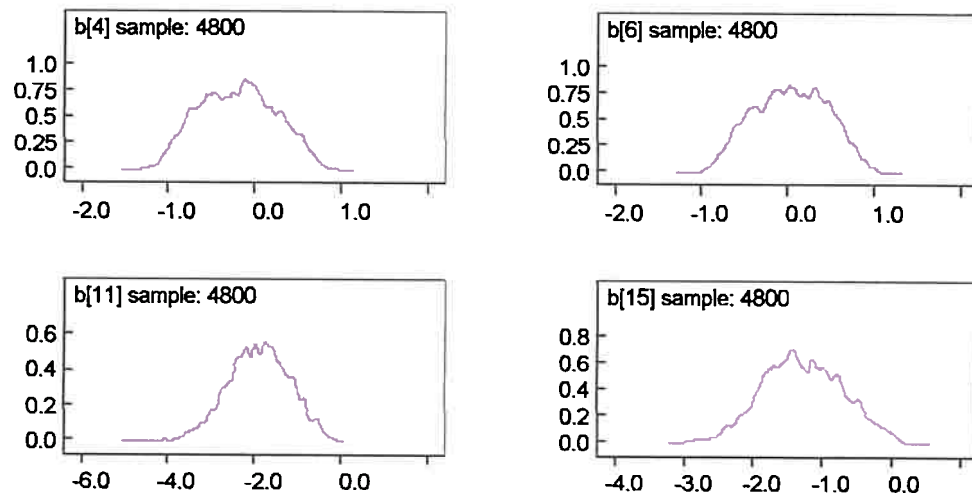
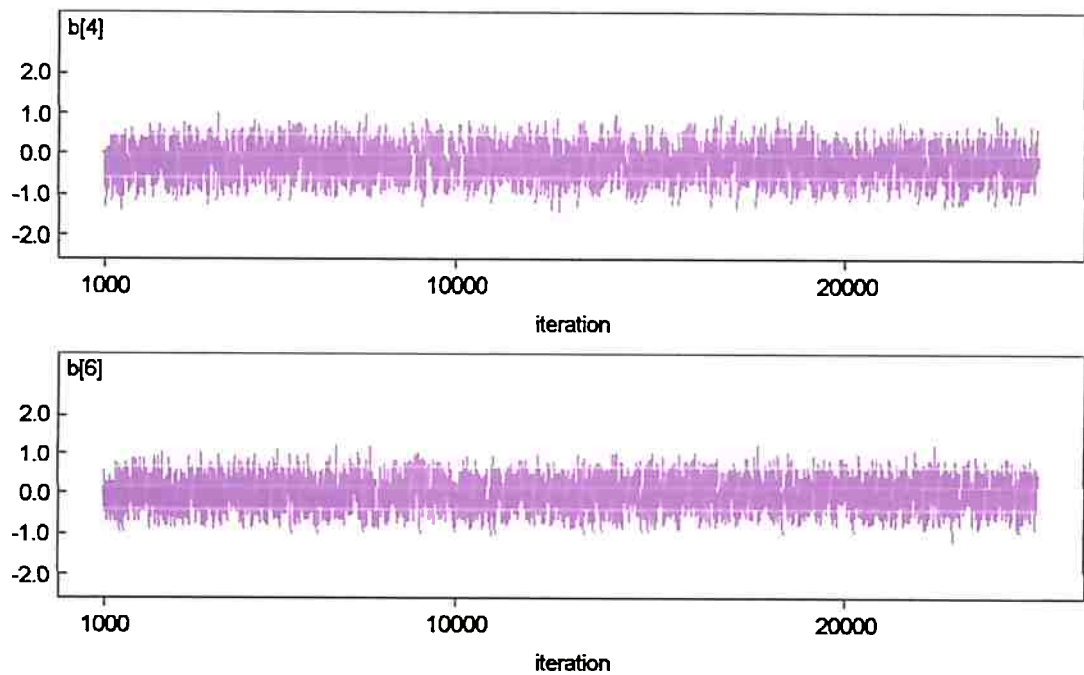


Figura 4.19 Densidades a posteriori para o parâmetro de dificuldade dos itens 4, 6, 11 e 15



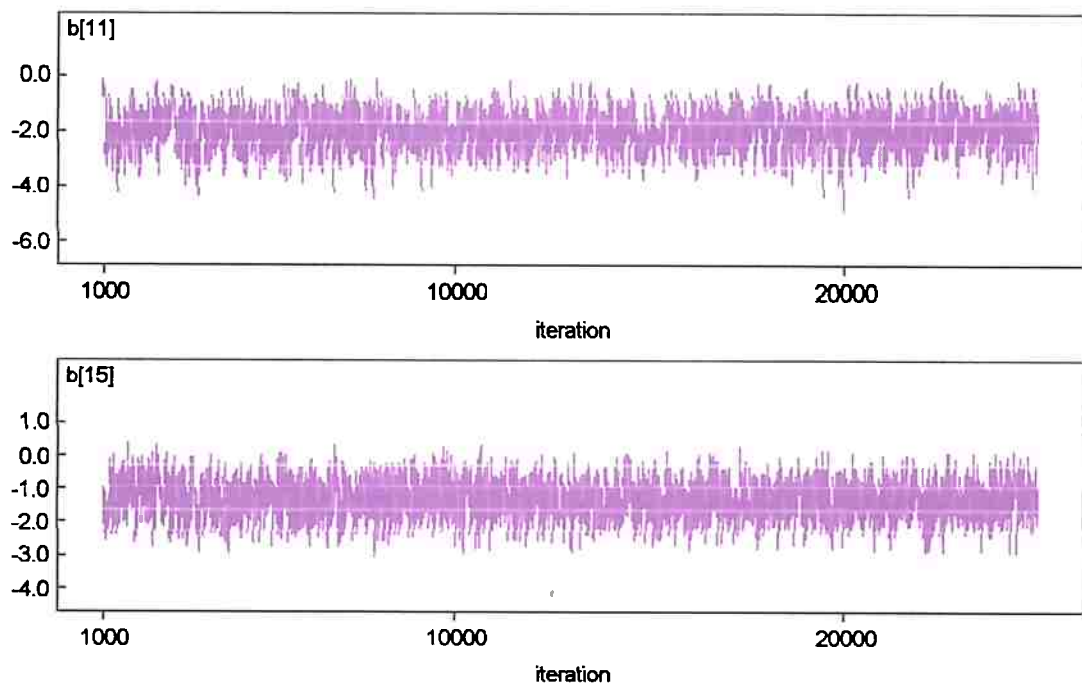


Figura 4.20 Histórico de convergência para o parâmetro de dificuldade dos itens 4, 6, 11 e 15

Observando os valores das estimativas obtidas para o parâmetro que mede a dificuldade do item, temos, de uma forma geral, que os itens foram facilmente respondidos e assim como no modelo anterior foram encontrados baixos valores para os b_i e em todos os casos abaixo da média da priori para a distribuição associada a este parâmetro.

Parâmetro da assimetria d_i

Neste modelo foi introduzido mais um parâmetro, estabelecendo uma assimetria da distribuição usada para a ligação e abrindo a possibilidade de que a probabilidade se aproxima de 1 a uma taxa diferente da que se aproxima de 0.

Este fato pode ser entendido da seguinte forma: quanto maior o valor de d significa que o item penaliza os bons alunos e bonifica os maus alunos. Lembremos que $|d| \leq 1$, assim quanto mais próximo de 1 estiver d , mais o item está penalizando os indivíduos mais hábeis e quanto mais próximo de -1 , mais penaliza os indivíduos menos hábeis.

Para entender melhor o que se aborda, acima vejamos o exemplo a seguir:

Suponha que um professor submeta seus alunos a um item que cobre a resolução de uma equação do segundo grau, em que o respondente deverá saber a famosa fórmula de Bháskara e que no enunciado seja fornecida a fórmula resolutive. Ora, para os alunos mais hábeis que sabem o conceito e a fórmula, esta informação é praticamente desnecessária, porém para os indivíduos menos hábeis pode ser que seja extremamente relevante a informação para a resolução da questão. Dessa forma a informação estaria penalizando os bons alunos, pois não precisam da informação, e bonificando os menos hábeis.

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
d[1]	-0.0647	0.5134	0.01342	-0.9252	-0.09293	0.9066	5000
d[2]	-0.1025	0.5543	0.01516	-0.9575	-0.1393	0.8937	5000
d[3]	-0.08218	0.5105	0.01144	-0.9183	-0.1286	0.9048	5000
d[4]	-0.02113	0.5059	0.01143	-0.9219	-0.03327	0.9067	5000
d[5]	-0.01415	0.496	0.01092	-0.8803	-0.01594	0.8937	5000
d[6]	0.01238	0.49	0.01241	-0.8724	0.005541	0.8963	5000
d[7]	-0.05689	0.4971	0.01271	-0.8892	-0.07846	0.8734	5000
d[8]	-0.05097	0.501	0.01148	-0.9058	-0.07445	0.8931	5000
d[9]	-0.02648	0.5016	0.01167	-0.8839	-0.04749	0.8969	5000
d[10]	-0.03225	0.5042	0.01168	-0.8643	-0.06587	0.9196	5000
d[11]	-0.2989	0.5572	0.01558	-0.983	-0.4506	0.9061	5000
d[12]	-0.1155	0.5086	0.01209	-0.924	-0.1643	0.8735	5000
d[13]	-0.1154	0.5238	0.01253	-0.9363	-0.1699	0.8839	5000
d[14]	-0.1171	0.5256	0.01242	-0.9239	-0.1786	0.8976	5000
d[15]	-0.1001	0.5482	0.01388	-0.929	-0.1674	0.9268	5000

Tabela 4.6 Estimativas para o parâmetro de assimetria

Analisando a tabela (4.6), em que são fornecidas as estatísticas para o parâmetro de assimetria, pode-se observar que os valores obtidos para d_i não apresentam valores muito distantes de 0, média adotada para a distribuição a priori deste parâmetro. No entanto, alguns itens merecem uma melhor análise.

No item 10, o valor de d nos diz que este item está penalizando os indivíduos mais hábeis e bonificando os menos hábeis. Já no item 11, encontramos o inverso, ou seja, penaliza os menos hábeis e bonifica os mais hábeis.

Quanto às densidades a posteriori, foram encontrados gráficos que permitem observar a simetria da distribuição a posteriori para o item 11 e certa simetria para o item 6. Isso nos permite dizer que a hipótese de assimetria pode ser verificada ou não.

Quanto ao intervalo de confiança, foram encontrados como limite inferior e superior valores cuja média da priori para d se nos encontra mesmos em todos os casos, porém aqui se pode destacar que estes valores estão bem próximos dos mínimos e máximos estabelecidos para a priori de d , como no capítulo 3.

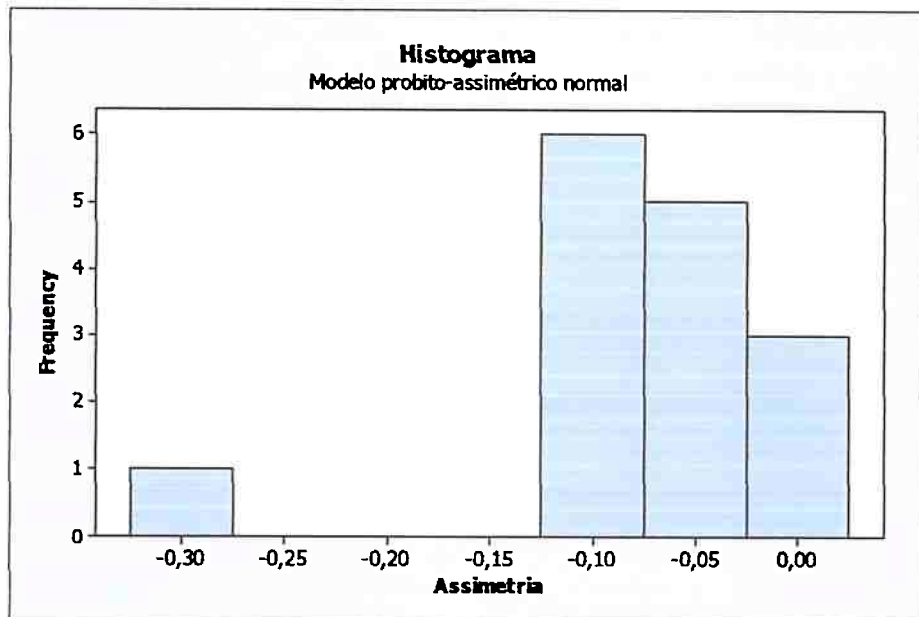


Figura 4.21 Histograma dos valores do parâmetro de assimetria

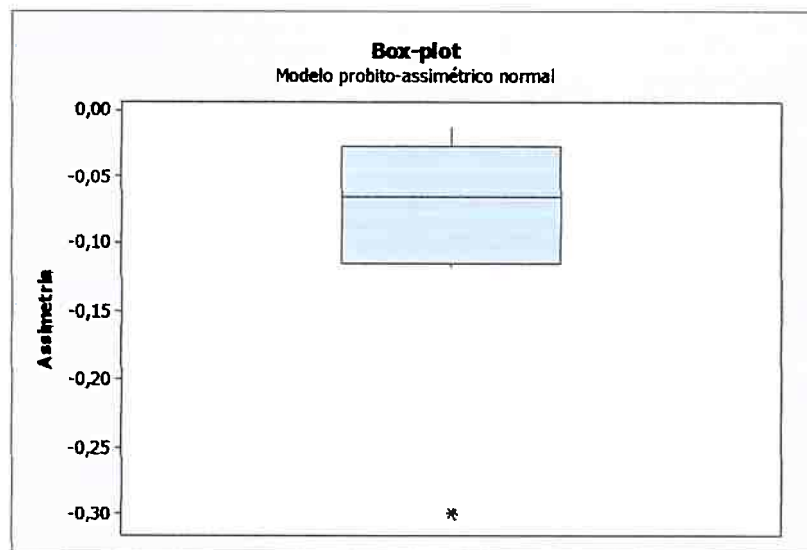


Figura 4.22 Box – plot dos valores do parâmetro de assimetria

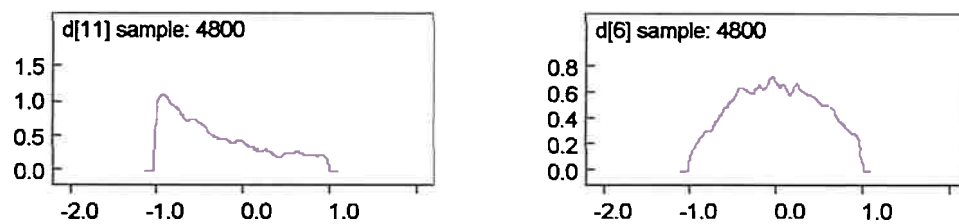


Figura 4.23 Densidades a posteriori do parâmetro de assimetria para os itens 6 e 11

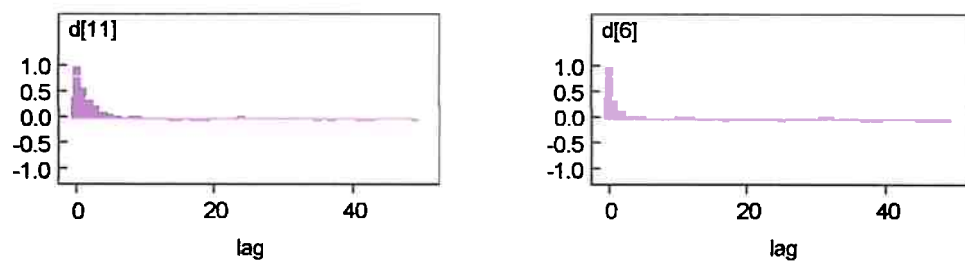


Figura 4.24 Gráfico das correlações para o parâmetro de assimetria para os itens 6 e 11

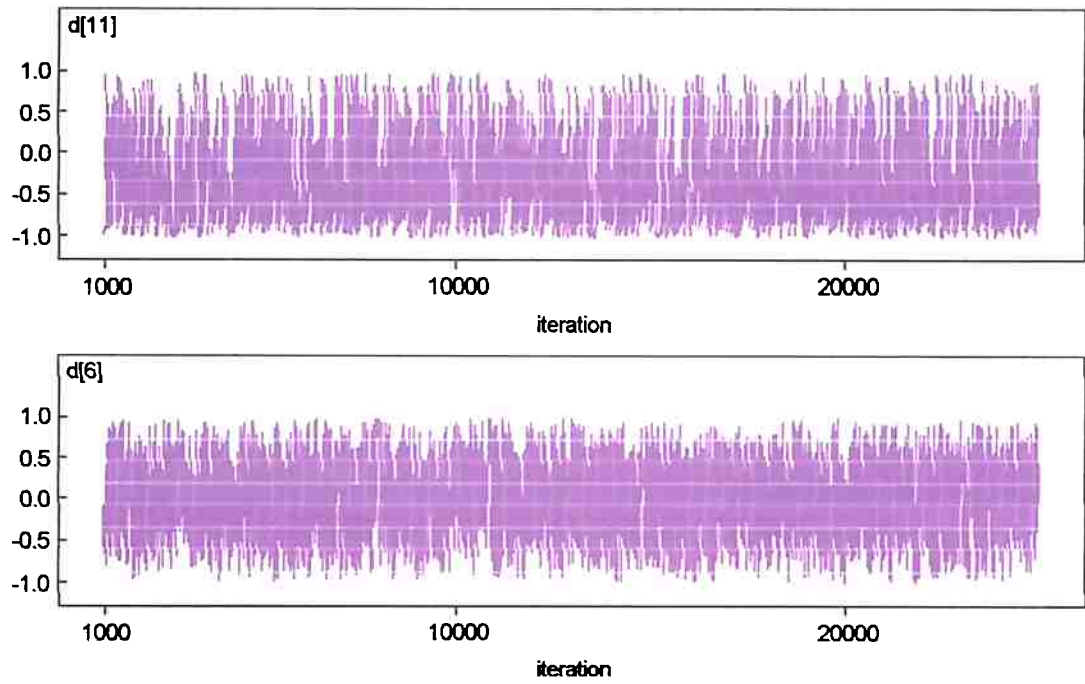


Figura 4.25 Gráfico do histórico de convergência para o parâmetro de assimetria dos itens 6 e 11

Parâmetro de habilidade $\theta = u$

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
u[1]	-0.5679	0.3698	0.006797	-1.315	-0.5729	0.1415	5000
u[2]	1.27	0.6478	0.0105	0.1372	1.224	2.642	5000
u[3]	-0.9857	0.3725	0.00741	-1.741	-0.9732	-0.2749	5000
u[4]	-0.723	0.3616	0.005755	-1.445	-0.7067	-0.0283	5000
u[5]	-0.2825	0.3653	0.006926	-0.9982	-0.2854	0.4519	5000
u[6]	1.258	0.6444	0.01081	0.1623	1.2	2.686	5000
u[7]	-0.5022	0.3648	0.006802	-1.211	-0.4998	0.2045	5000
u[8]	0.208	0.4326	0.008247	-0.5725	0.1864	1.112	5000
u[9]	0.8328	0.5579	0.009428	-0.1238	0.7792	2.036	5000
u[10]	0.8346	0.5543	0.01047	-0.1364	0.7977	2.032	5000
u[11]	1.271	0.6528	0.01377	0.1411	1.23	2.647	5000
u[12]	0.2508	0.4413	0.007301	-0.5458	0.2281	1.196	5000
u[13]	-0.3534	0.3644	0.005758	-1.059	-0.3533	0.3713	5000
u[14]	0.8395	0.5598	0.01113	-0.1352	0.8	2.073	5000
u[15]	1.257	0.647	0.0126	0.1273	1.215	2.67	5000
u[16]	0.1296	0.4599	0.007962	-0.713	0.0981	1.129	5000
u[17]	-0.2395	0.3705	0.005831	-0.9463	-0.2472	0.5	5000
u[18]	0.1517	0.4181	0.006928	-0.641	0.1371	1.018	5000
u[19]	0.3181	0.4563	0.007181	-0.4989	0.2899	1.323	5000
u[20]	1.28	0.651	0.01135	0.1264	1.233	2.65	5000
u[21]	-0.08966	0.3914	0.006046	-0.8152	-0.1074	0.7211	5000
u[22]	0.8394	0.5629	0.01076	-0.1204	0.788	2.075	5000
u[23]	0.9431	0.614	0.0109	-0.1001	0.8933	2.319	5000
u[24]	0.6627	0.5305	0.00867	-0.2483	0.6155	1.835	5000
u[25]	-0.4813	0.3672	0.006411	-1.212	-0.4783	0.2338	5000
u[26]	0.4355	0.468	0.00811	-0.4043	0.3954	1.463	5000
u[27]	-0.2217	0.3711	0.006263	-0.947	-0.234	0.5175	5000
u[28]	-0.5892	0.3519	0.006336	-1.303	-0.5798	0.07299	5000
u[29]	0.7086	0.5356	0.008484	-0.2174	0.6682	1.868	5000
u[30]	-0.5195	0.3544	0.006541	-1.235	-0.5206	0.18	5000
u[31]	0.09855	0.4108	0.007569	-0.6699	0.08174	0.968	5000
u[32]	0.7337	0.535	0.00847	-0.2088	0.6962	1.88	5000
u[33]	-0.7065	0.3427	0.005839	-1.407	-0.7062	-0.04992	5000
u[34]	0.3358	0.445	0.007807	-0.4622	0.308	1.29	5000
u[35]	-1.295	0.4108	0.007192	-2.164	-1.276	-0.5395	5000
u[36]	-0.05998	0.3854	0.006796	-0.8077	-0.06371	0.714	5000
u[37]	1.271	0.642	0.01252	0.1776	1.222	2.658	5000
u[38]	-1.556	0.4295	0.006759	-2.468	-1.536	-0.7946	5000
u[39]	-0.06181	0.3932	0.007397	-0.7896	-0.0686	0.7323	5000
u[40]	-0.4507	0.3667	0.006032	-1.17	-0.4544	0.2812	5000
u[41]	-0.2012	0.3884	0.006495	-0.9367	-0.214	0.5916	5000
u[42]	0.2987	0.4561	0.007519	-0.5074	0.268	1.29	5000
u[43]	-0.5801	0.3629	0.007615	-1.324	-0.5801	0.1205	5000
u[44]	0.1362	0.415	0.007422	-0.6225	0.1204	1.001	5000
u[45]	-0.4089	0.3626	0.006257	-1.102	-0.4062	0.3016	5000
u[46]	1.255	0.6457	0.01165	0.1451	1.201	2.644	5000
u[47]	-0.6518	0.3587	0.006079	-1.376	-0.6455	0.0505	5000

Tabela 4.7 Estimativas para as habilidades dos indivíduos

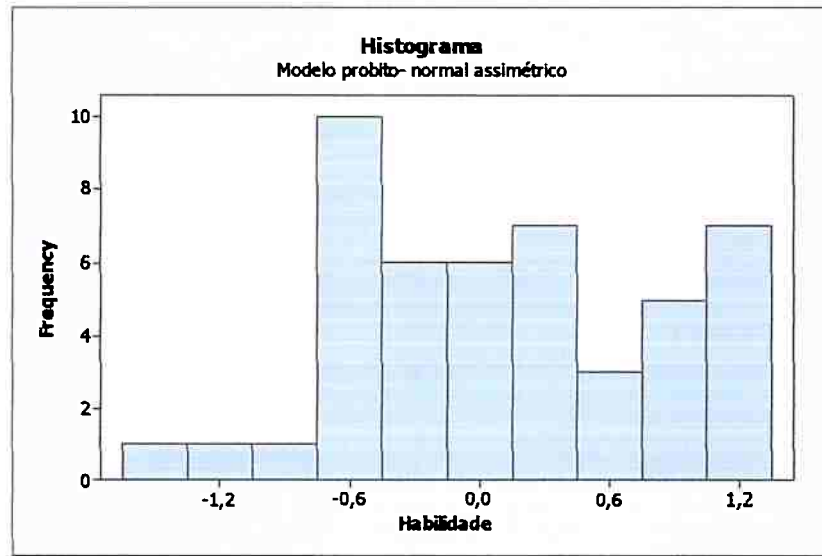


Figura 4.26 Histograma com os valores das habilidades

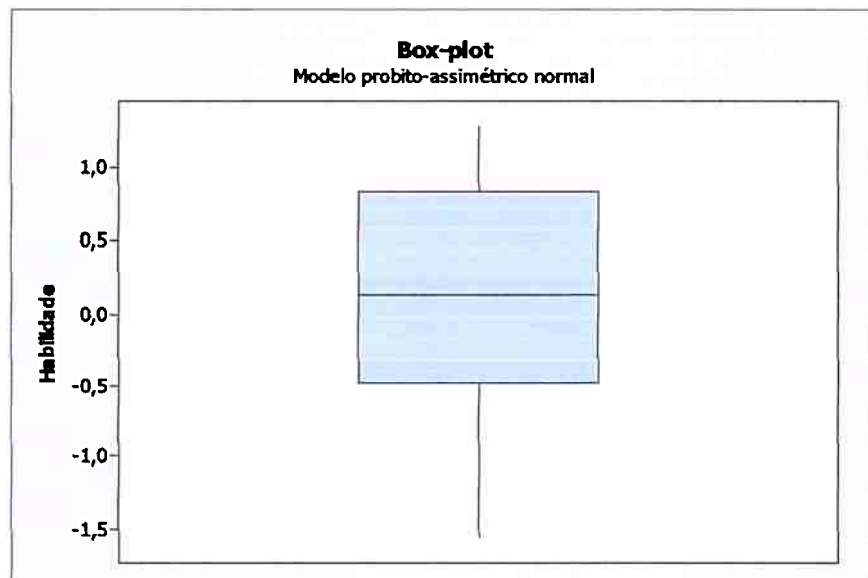


Figura 4.27 Box-Plot dos valores das habilidades

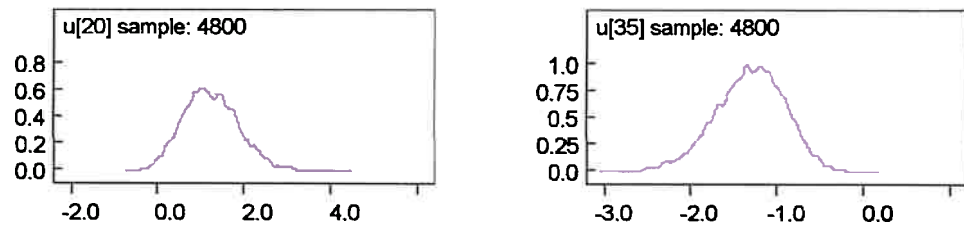


Figura 4.28 Densidades das posteriori para as habilidades dos indivíduos 20 e 35

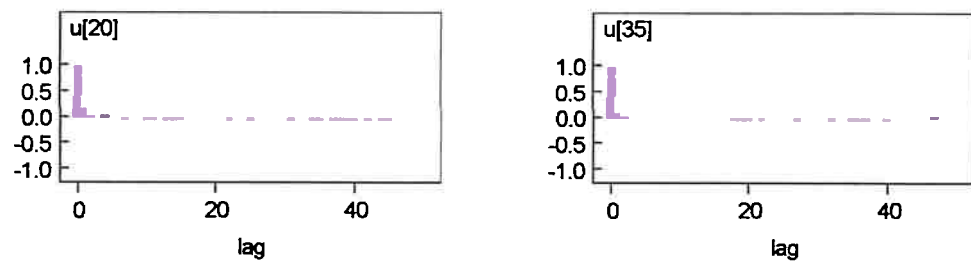


Figura 4.29 Gráfico das correlações das habilidades dos indivíduos 20 e 35

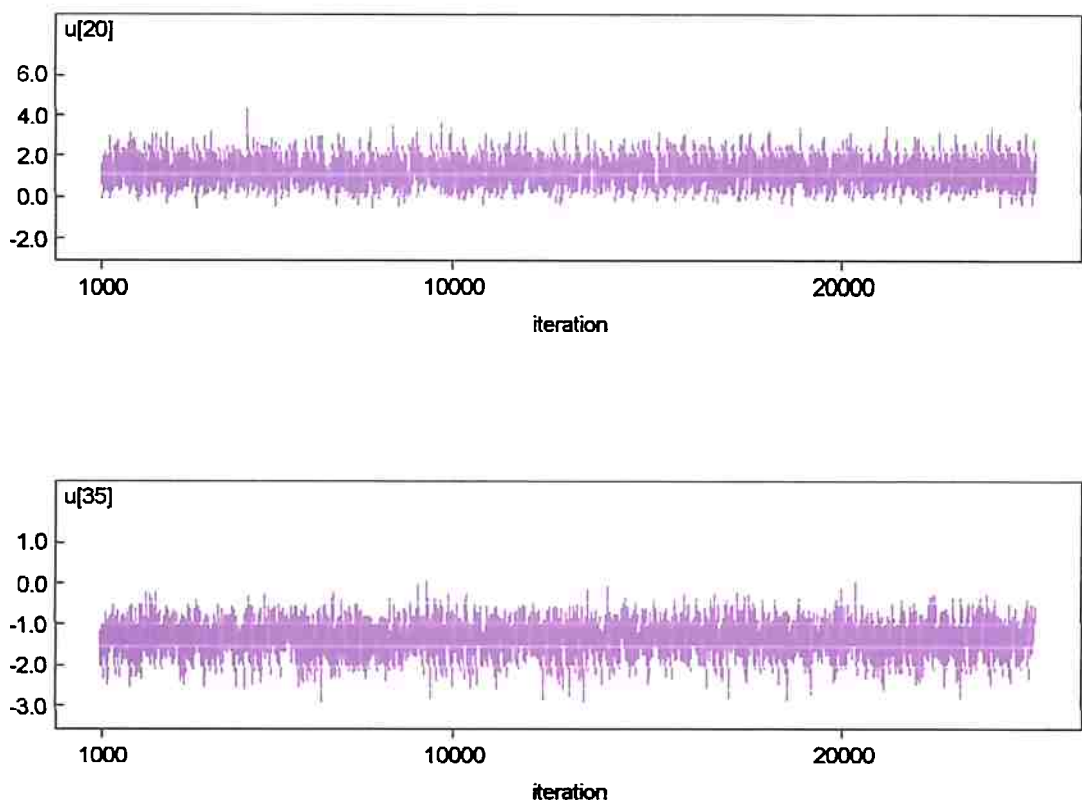


Figura 4.30 Histórico de convergência para as habilidades dos indivíduos 20 e 35

Na tabela (4.7) pode-se observar que neste modelo encontramos estimativas para as habilidades cujos valores identificam indivíduos com habilidades acima e abaixo da média da distribuição a priori para o parâmetro de habilidade. Assim como no modelo probito-normal (PN) os indivíduos com mais habilidades são os indivíduos 2 e 14 e os de menos habilidades são 6 e 10, sendo que na maioria dos casos o intervalo de credibilidade de 95% contém o valor da média a priori para θ .

4.5 Modelo probito normal-assimétrico

Parâmetro de discriminação α

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
a[1]	1.019	0.4466	0.007839	0.2261	0.987	1.972	5000
a[2]	1.623	0.4673	0.007115	0.8	1.588	2.637	5000
a[3]	1.193	0.46	0.006854	0.3727	1.171	2.171	5000
a[4]	1.082	0.4013	0.00661	0.3778	1.046	1.944	5000
a[5]	1.111	0.4141	0.005973	0.3732	1.091	1.996	5000
a[6]	0.8578	0.3626	0.005667	0.2251	0.8262	1.631	5000
a[7]	1.034	0.4338	0.007311	0.2728	1.003	1.967	5000
a[8]	1.103	0.4384	0.008127	0.3212	1.066	2.055	5000
a[9]	1.224	0.4211	0.005758	0.4813	1.196	2.12	5000
a[10]	0.5801	0.372	0.005375	0.03713	0.5308	1.445	5000
a[11]	1.52	0.5715	0.009033	0.482	1.486	2.721	5000
a[12]	1.062	0.4828	0.007916	0.2318	1.028	2.097	5000
a[13]	1.25	0.5007	0.007747	0.3426	1.223	2.295	5000
a[14]	1.46	0.5126	0.007675	0.5445	1.435	2.525	5000
a[15]	1.406	0.5243	0.009092	0.4561	1.371	2.512	5000

Tabela 4.8 Estimativas para o parâmetro de discriminação

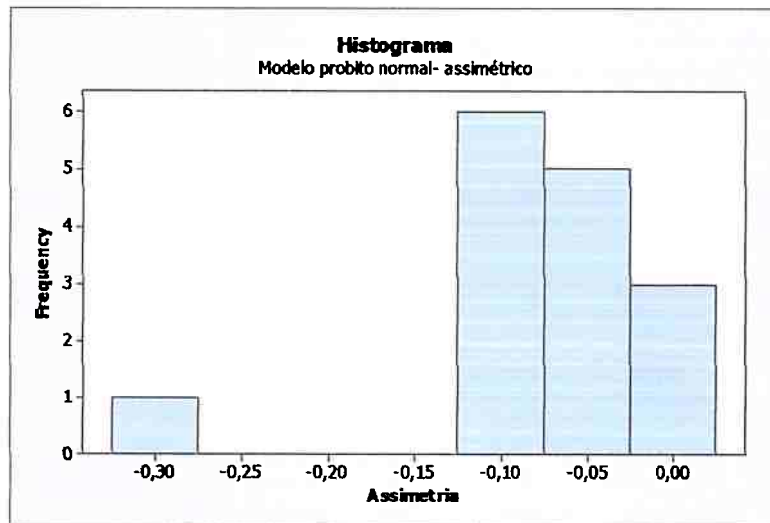


Figura 4.31 Histograma dos valores do parâmetro de discriminação

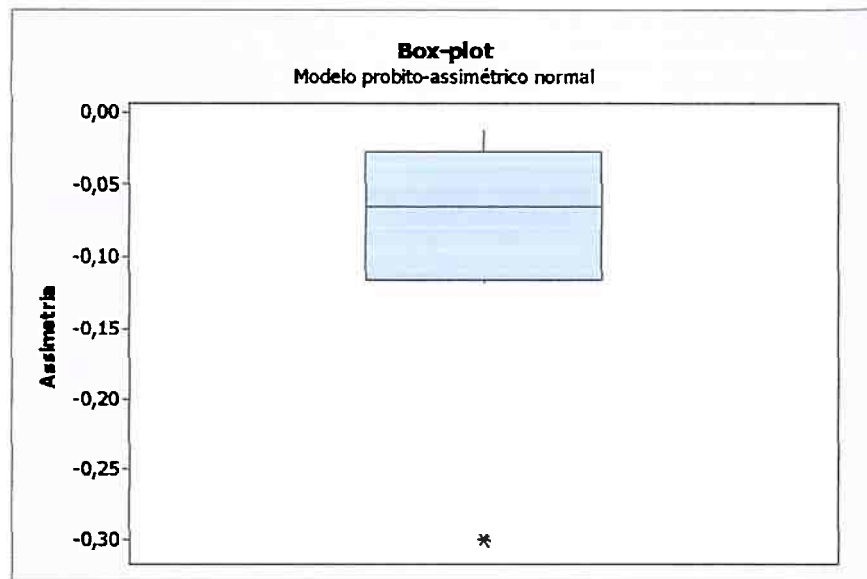


Figura 4.32 Box -plot dos valores do parâmetro de discriminação

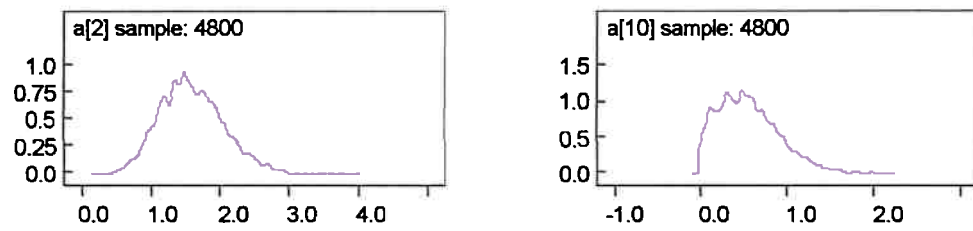


Figura 4.33 Densidade a posteriori do parâmetro de discriminação dos itens 2 e 10

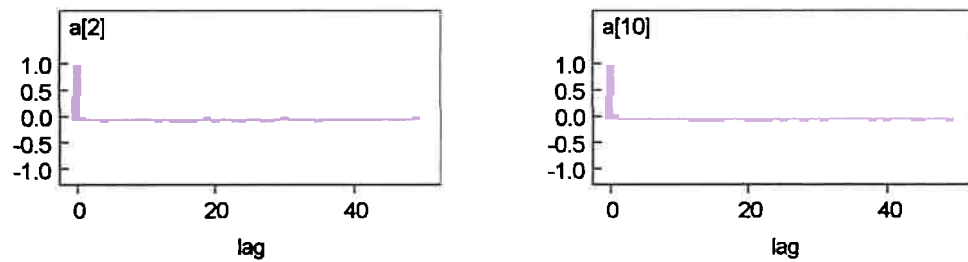


Figura 4.34 Gráfico das correlações para o parâmetro de discriminação dos itens 2 e 10

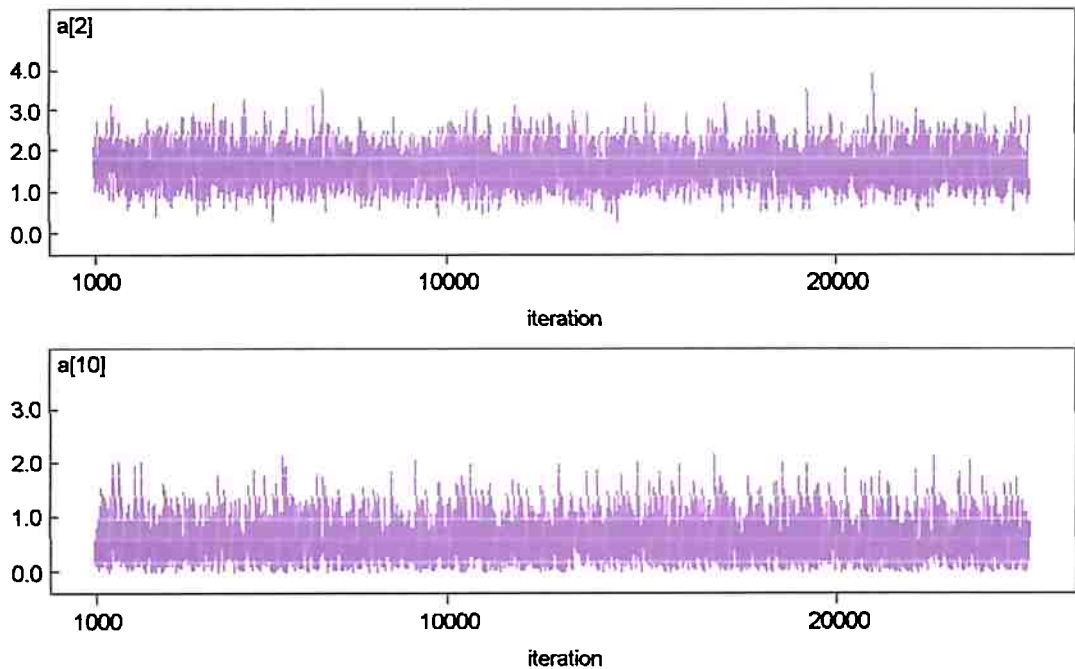


Figura 4.35 Histórico de convergência para o parâmetro de discriminação dos itens 2 e 10

Neste modelo pode –se observar através da tabela 4.8 que alguns itens apresentaram boa discriminação como por exemplo os de número 2 e 10, sendo que no item 10 o gráfico da densidade a posteriori apresentou uma ligeira assimetria.

Parâmetro de dificuldade b

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
b[1]	-0.4739	0.332	0.008322	-1.081	-0.4811	0.2239	5000
b[2]	0.488	0.3978	0.01063	-0.2635	0.4872	1.273	5000
b[3]	-0.2463	0.3478	0.008319	-0.9129	-0.2472	0.4637	5000
b[4]	0.3957	0.3454	0.007819	-0.2294	0.3837	1.098	5000
b[5]	0.2774	0.3426	0.007305	-0.3436	0.2582	0.9722	5000
b[6]	0.5355	0.3304	0.006728	-0.04631	0.5118	1.246	5000
b[7]	-0.2069	0.3347	0.008168	-0.8333	-0.22	0.4883	5000
b[8]	-0.1909	0.3425	0.008728	-0.8117	-0.201	0.5136	5000
b[9]	0.3254	0.3501	0.007375	-0.334	0.3156	1.027	5000
b[10]	-0.9363	0.314	0.005342	-1.524	-0.9456	-0.2944	5000
b[11]	-1.39	0.489	0.01105	-2.43	-1.365	-0.4927	5000
b[12]	-0.8159	0.3506	0.007138	-1.508	-0.8155	-0.141	5000
b[13]	-0.6573	0.3735	0.009411	-1.398	-0.6558	0.05731	5000
b[14]	-0.5002	0.3897	0.009668	-1.268	-0.4979	0.253	5000
b[15]	-0.6312	0.3886	0.01015	-1.408	-0.625	0.1037	5000

Tabela 4.9 Estimativas para o parâmetro de dificuldade

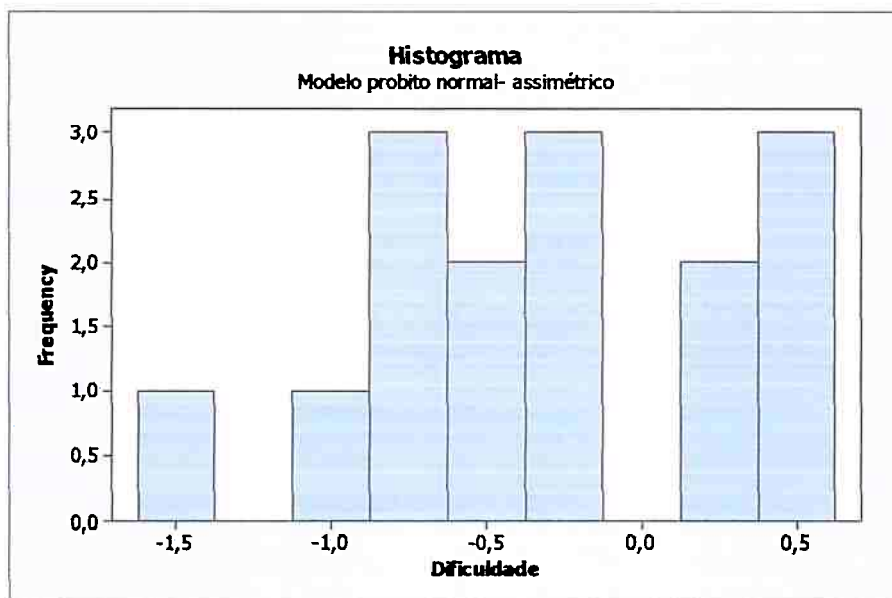


Figura 4.36 Histograma dos valores do parâmetro de dificuldade

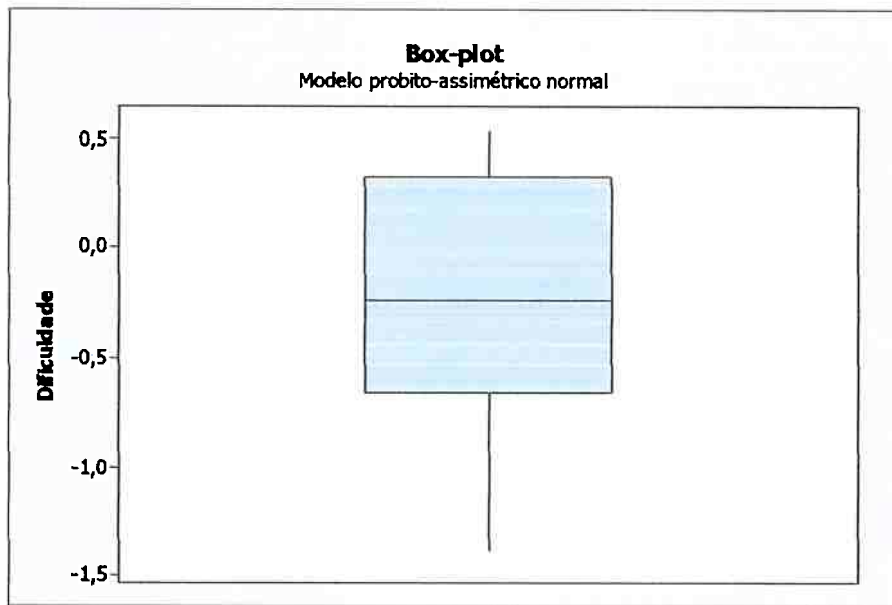


Figura 4.37 Box – Plot dos valores do parâmetro de dificuldade



Figura 4.38 Densidades a posteriori para o parâmetro de dificuldade dos itens 6 e 10

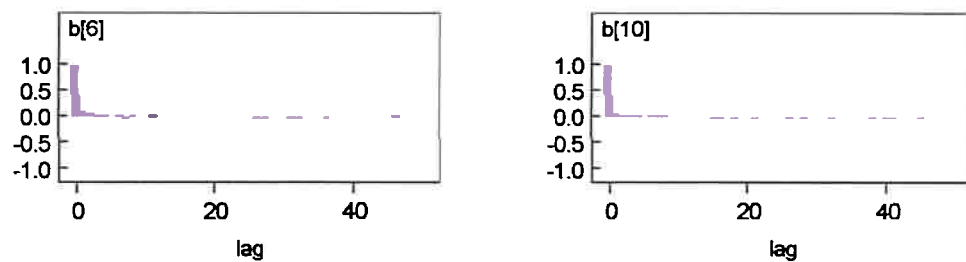


Figura 4.39 Gráfico das correlações para o parâmetro de dificuldade dos itens 6 e 10

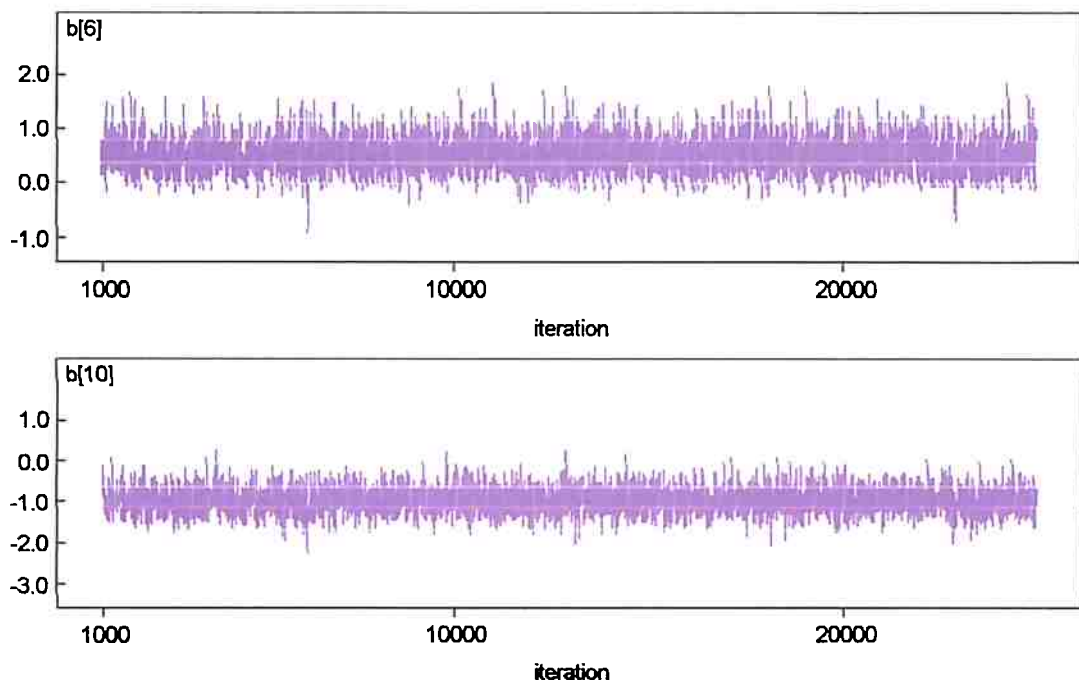


Figura 4.40 Histórico de convergência para o parâmetro de dificuldade dos itens 6 e 10

Com base nos resultados obtidos anteriormente pode-se afirmar que os itens não apresentaram grandes dificuldades para os respondentes, pois os mesmos apresentaram baixos valores para o parâmetro de dificuldade.

É importante observar que nem todos os intervalos de credibilidade contêm o valor da média a priori definido para este parâmetro.

Os gráficos apresentados mostram posterioris com características de normalidade e os gráficos das correlações mostram que as mesmas foram eliminadas ao final.

O histórico durante as iterações mostra que devido a grande oscilação a convergência foi atingida de forma satisfatória.

Parâmetro de habilidade $\theta = u$

Item []	Média	Desvio Padrão	Erro	$P_{2,5}$	Mediana	$P_{97,5}$	Amostra
u[1]	0.08558	0.3181	0.00796	-0.608	0.1014	0.6681	5000
u[2]	1.63	0.5682	0.009292	0.6497	1.579	2.853	5000
u[3]	-0.1908	0.3475	0.00887	-0.9754	-0.1515	0.4065	5000
u[4]	-3,10E-01	0.3303	0.008952	-0.7209	0.02618	0.5752	5000
u[5]	0.3332	0.3234	0.006204	-0.318	0.3409	0.9575	5000
u[6]	1.649	0.5729	0.008168	0.6209	1.618	2.883	5000
u[7]	0.1582	0.3259	0.007723	-0.5433	0.1718	0.7654	5000
u[8]	0.716	0.3727	0.005845	0.03525	0.7003	1.498	5000
u[9]	1.272	0.4836	0.007749	0.4291	1.234	2.335	5000
u[10]	1.27	0.4918	0.007616	0.4003	1.236	2.38	5000
u[11]	1.652	0.5737	0.007517	0.6628	1.603	2.902	5000
u[12]	0.7401	0.3876	0.007061	0.01485	0.7223	1.561	5000
u[13]	0.2882	0.3368	0.007741	-0.4053	0.2916	0.9212	5000
u[14]	1.275	0.5079	0.007382	0.3774	1.245	2.408	5000
u[15]	1.649	0.5722	0.008919	0.6518	1.608	2.885	5000
u[16]	0.6956	0.3904	0.006781	-0.03241	0.6796	1.524	5000
u[17]	0.4073	0.3352	0.006959	-0.256	0.4071	1.065	5000
u[18]	0.6675	0.3673	0.006543	-0.02768	0.6565	1.453	5000
u[19]	0.8361	0.3898	0.006525	0.1075	0.8266	1.633	5000
u[20]	1.656	0.5814	0.009941	0.6404	1.613	2.931	5000
u[21]	0.5143	0.3433	0.006416	-0.1509	0.5143	1.187	5000
u[22]	1.273	0.4869	0.007462	0.4183	1.233	2.344	5000
u[23]	1.316	0.5206	0.008141	0.4425	1.269	2.542	5000
u[24]	1.113	0.4506	0.00722	0.3212	1.074	2.072	5000
u[25]	0.1719	0.3189	0.007863	-0.509	0.1807	0.763	5000
u[26]	0.9294	0.4127	0.006652	0.1913	0.9041	1.807	5000
u[27]	0.3909	0.3392	0.006779	-0.2956	0.3973	1.049	5000
u[28]	0.09964	0.3328	0.008614	-0.6198	0.113	0.7048	5000
u[29]	1.154	0.4644	0.008502	0.3299	1.123	2.162	5000
u[30]	0.1523	0.3226	0.007509	-0.5281	0.1634	0.7638	5000
u[31]	0.6219	0.3668	0.00675	-0.07978	0.6132	1.363	5000
u[32]	1.178	0.4671	0.006313	0.3595	1.148	2.193	5000
u[33]	-0.01901	0.315	0.006698	-0.6666	-0.001992	0.5498	5000
u[34]	0.8519	0.396	0.007278	0.1239	0.8385	1.685	5000
u[35]	-0.4459	0.3871	0.01234	-1.305	-0.4023	0.1819	5000
u[36]	0.5024	0.3427	0.006516	-0.1721	0.5011	1.191	5000
u[37]	1.644	0.5731	0.00877	0.6439	1.609	2.88	5000
u[38]	-0.5809	0.4233	0.01287	-1.534	-0.5321	0.1012	5000
u[39]	0.5409	0.3514	0.006763	-0.155	0.5352	1.25	5000
u[40]	0.2198	0.3336	0.007879	-0.4795	0.2242	0.8492	5000
u[41]	0.42	0.3315	0.006765	-0.2456	0.4173	1.073	5000
u[42]	0.8124	0.3881	0.005816	0.1009	0.7945	1.629	5000
u[43]	0.06692	0.3238	0.007929	-0.5999	0.0846	0.6604	5000
u[44]	0.6756	0.3639	0.005871	-0.01628	0.6641	1.439	5000
u[45]	0.2345	0.3272	0.007652	-0.4458	0.2426	0.8422	5000
u[46]	1.64	0.582	0.01035	0.6194	1.6	2.925	5000
u[47]	0.03984	0.3349	0.007604	-0.7004	0.06431	0.6304	5000

Tabela 4.10 Estimativas para o parâmetro de habilidade

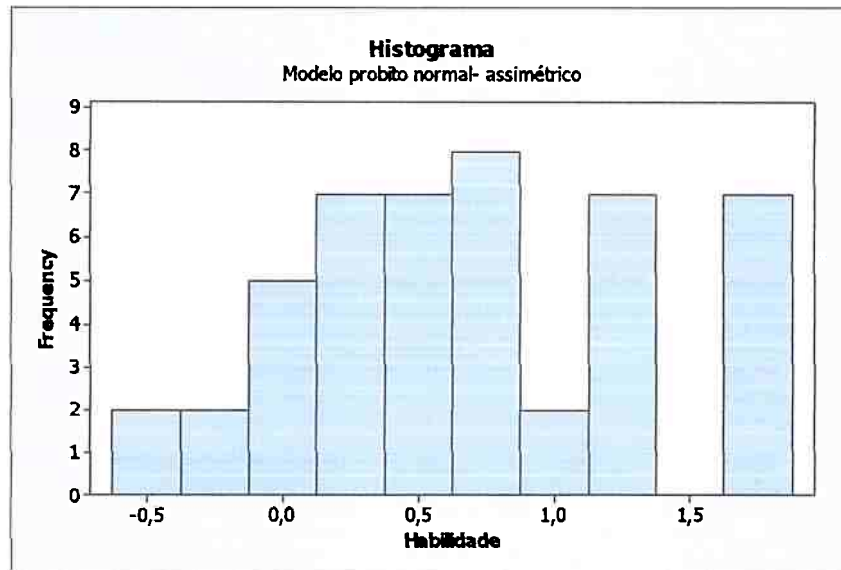


Figura 4.41 Histograma com os valores das habilidades.

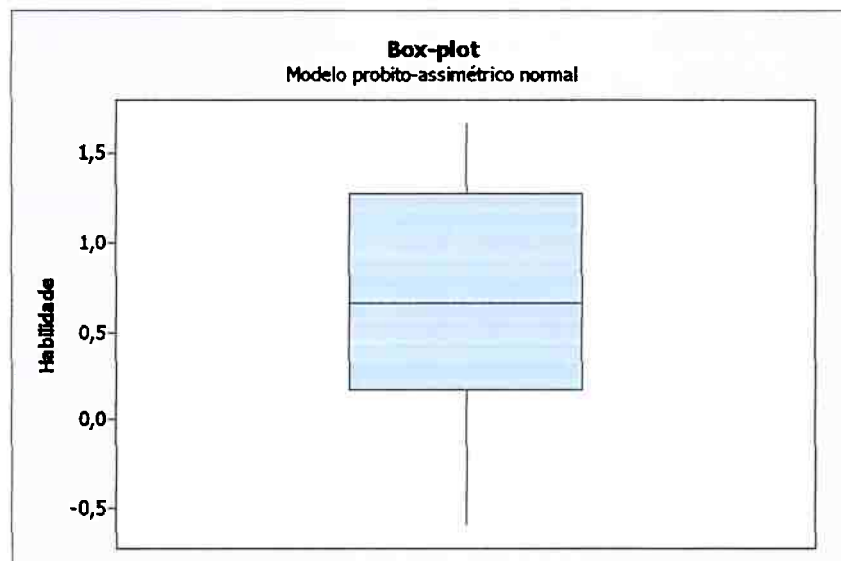


Figura 4.42 Box-Plot dos valores das habilidades

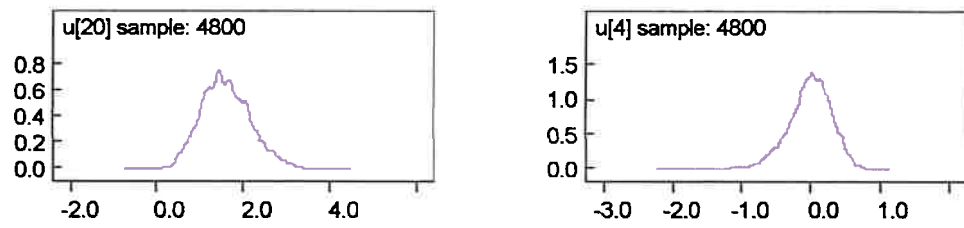


Figura 4.43 Densidades a posteriori para o parâmetro de habilidades dos indivíduos 4 e 20

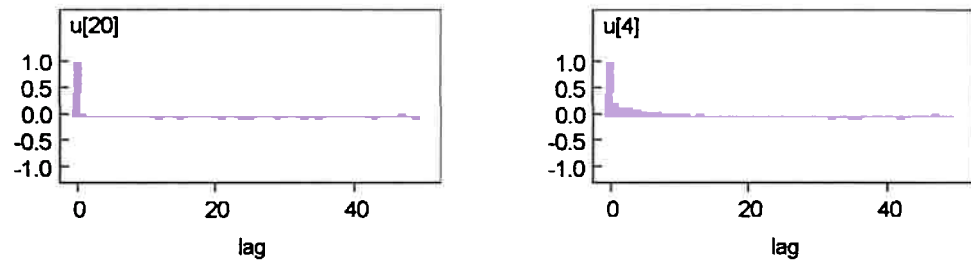


Figura 4.44 Gráfico das correlações para o parâmetro de habilidades dos indivíduos 4 e 20

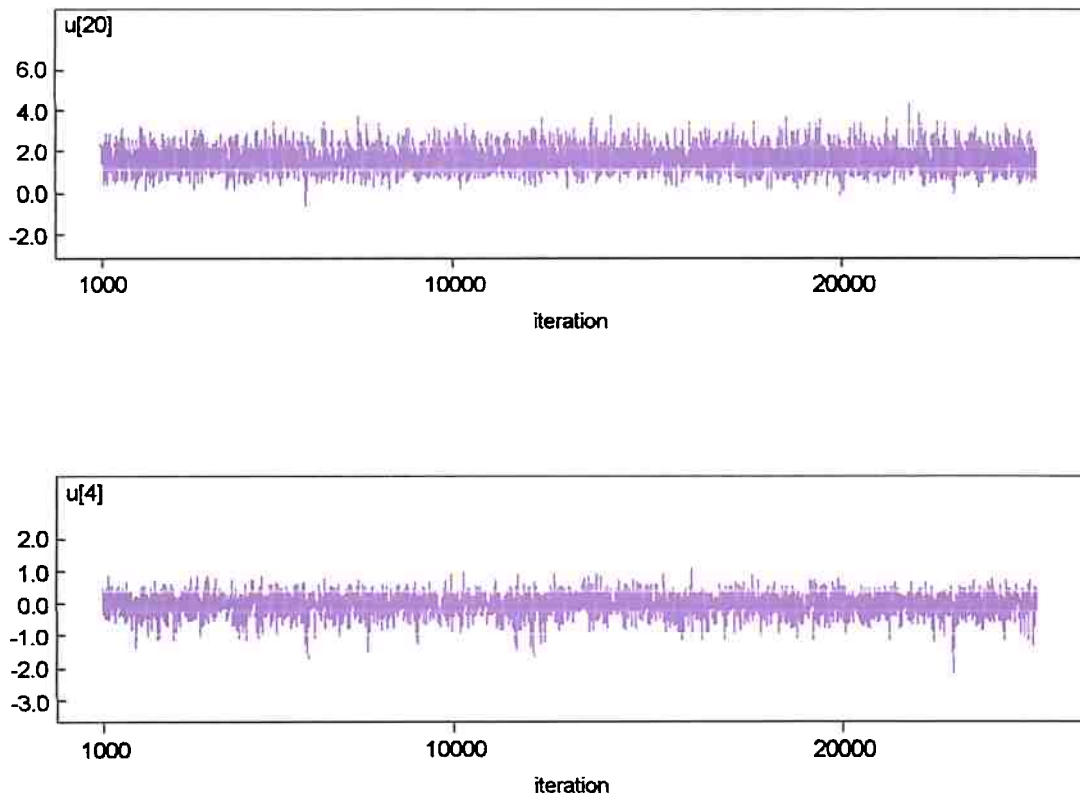


Figura 4.45 Histórico de convergência para as habilidades dos indivíduos 4 e 20

Da análise dos valores obtidos na tabela 4.10, pode se verificar que os indivíduos apresentaram de uma forma geral valores altos para o parâmetro das habilidades de cada um, sendo que os que obtiveram os maiores valores são aqueles que na tabela dada no apêndice esses indivíduos são os que apresentam os maiores valores de escore.

Apêndice A

A distribuição normal assimétrica

Uma variável R segue uma distribuição normal assimétrica (Dalla Valle, 2004), com vetor de parâmetros $\theta = (\mu, \sigma^2, \lambda)$ em que $\mu \in \mathbb{R}$ é um parâmetro de locação e $\sigma^2 > 0$ é um parâmetro de escala e $\lambda \in \mathbb{R}$ é um parâmetro de assimetria, se a função de densidade de R é dada por:

$$f_{\theta}(r) = \frac{2}{\sigma} \phi\left(\frac{r-\mu}{\sigma}\right) \Phi\left(\lambda \frac{r-\mu}{\sigma}\right), \quad (A.1)$$

onde $\phi(\cdot)$ e $\Phi(\cdot)$ denotam, respectivamente, a função de densidade e de distribuição da distribuição normal padrão. A notação considerada é $R \sim \text{NA}(\theta)$ onde $\theta = (\mu, \sigma^2, \lambda)$, mas também se pode usar a notação $\phi\text{NA}(r; \mu, \sigma^2, \lambda)$ para representar a densidade. $F_{\theta}(r)$, com $\theta = (\mu, \sigma^2, \lambda)$, denota a fda da distribuição normal assimétrica. Quando considerada a transformação $d = \frac{\lambda}{\sqrt{1+\lambda^2}}$ a distribuição normal assimétrica tem outra parametrização em

que $-1 \leq d \leq 1$. Para voltar à parametrização original considere $\lambda = \frac{d}{(1-d^2)^{1/2}}$.

Casos particulares da normal assimétrica

- Se $\lambda = 0$, i.e se $\theta = (\mu, \sigma^2, 0)$ a densidade de R em (A.1) se reduz a densidade da $N(\mu, \sigma^2)$. Em particular se $\theta = (0, 1, 0)$ obtemos a normal padrão.
- No caso especial de $\theta = (0, 1, \lambda)$, a densidade de R em (A.1) se reduz a densidade da normal assimétrica padrão dada por Azzalini (1985) e neste caso escrevemos que $S \sim \text{NA}(\lambda)$ cuja função de densidade é dada por:

$$f_{\lambda}(s) = 2\phi(s)\Phi(\lambda s) \quad (A.2)$$

- No caso especial de $\theta = (0, 1+\lambda^2, \lambda)$, a densidade de R em (A.1) se reduz a densidade da normal assimétrica padrão dada por Chen et al (1999) e Branco e Dey (2002) a qual por sua vez é um caso particular da normal assimétrica dada em Sahu et al. (2003).

Representações da normal assimétrica padrão de Azzalini

Embora simular da normal assimétrica padrão seja simples devido as diversas bibliotecas que podem ser obtidas da WWW page: <http://azzalini.stat.unipd.it/SN> w, as seguintes representações são úteis para entender a relação da normal assimétrica com a normal.

- Seja $W \sim N(0, 1)$ e $V \sim HN(0, 1)$ v.a independentes e $|d| < 1$. Temos que a distribuição de

$$S = dV + (1 - d^2)^{1/2}W$$

é $NA(\lambda)$. Esta é a representação estocástica marginal e foi provada por Henze (1986), Azzalini e Dalla Valle (1996).

- Se $V, W \sim N(0, 1)$ são v.a independentes, a distribuição de $S = W|V < \lambda W'$ e $NA(\lambda)$.

Veja Henze (1986), Azzalini e Dalla Valle (1996).

- Se $S|V \sim N(dv, 1 - d^2)$ e $V \sim HN(0, 1)$ (densidade normal positiva), a distribuição de S é $NA(\lambda)$.

- Seja

$$\begin{pmatrix} V \\ W \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & d \\ d & 1 \end{pmatrix}\right)$$

e $|d| < 1$. Temos que $W|V > 0 \sim NA(\lambda)$. Esta é a representação estocástica condicional e foi provada por Arnold et al (1993) e Azzalini e Dalla Valle (1996).

- Seja

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & d \\ d & 1 \end{pmatrix}\right)$$

e $|d| < 1$

Temos que

$$S = \begin{cases} W & \text{Se } V > 0, \\ -W & \text{c.c.} \end{cases} \text{ é } NA(\lambda).$$

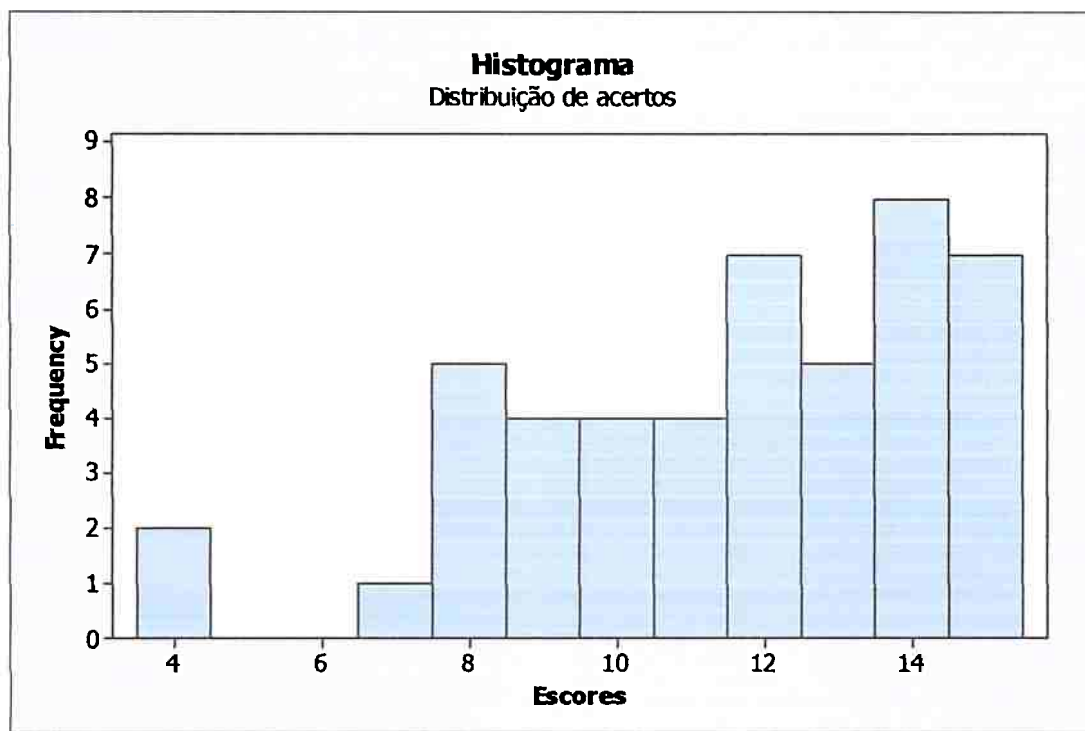
Esta outra representação estocástica condicional foi provada por Azzalini e Capitanio (1999).

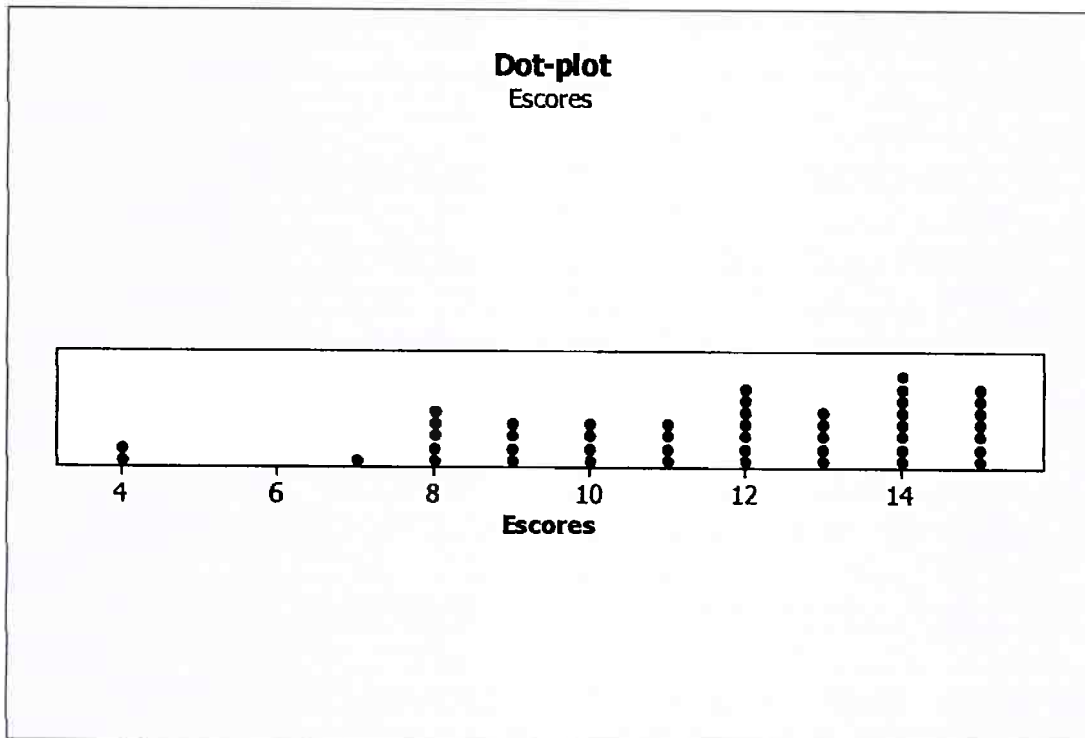
Apêndice B

	Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q15	TOT
A1	1	1	0	0	0	0	0	1	0	1	1	0	1	1	1	8
A2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
A3	0	0	0	1	1	1	0	0	0	1	0	1	1	1	0	7
A4	1	0	0	0	0	0	1	1	0	1	1	1	0	1	1	8
A5	1	1	0	1	0	0	1	0	0	1	1	1	1	1	1	10
A6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
A7	1	0	1	1	0	0	0	0	0	1	1	1	1	1	1	9
A8	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	12
A9	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	14
A10	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	14
A11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
A12	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	12
A13	1	0	1	0	0	0	1	1	0	1	1	1	1	1	1	10
A14	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	14
A15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
A16	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	13
A17	1	0	0	1	1	0	1	1	0	1	1	1	1	1	1	11
A18	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	12
A19	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	13
A20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
A21	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	12
A22	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	14
A23	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	14
A24	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	14
A25	1	0	1	1	0	-	0	1	1	0	1	1	1	0	1	9
A26	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	13
A27	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	11
A28	1	0	1	0	0	0	1	1	0	1	1	1	1	0	1	9
A29	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	14
A30	0	0	1	0	1	0	1	0	1	1	1	1	0	1	1	9
A31	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	12
A32	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	14
A33	0	1	0	0	1	1	1	0	1	1	1	1	0	0	0	8
A34	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	13
A35	1	0	1	0	0	0	0	0	0	0	1	0	1	0	0	4
A36	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	11
A37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
A38	0	0	0	0	0	0	1	1	0	1	0	1	0	0	0	4
A39	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	12
A40	1	0	1	1	1	0	1	1	0	1	1	0	1	1	0	10
A41	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	11
A42	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	13
A43	1	1	1	0	0	0	0	0	1	1	1	0	1	0	1	8

A44	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	12
A45	1	0	1	0	0	1	1	0	1	1	1	1	0	1	1	10
A46	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	15
A47	1	0	1	0	0	0	0	0	0	1	1	1	1	1	1	8
total de acertos	39	30	37	28	30	24	36	36	30	42	45	42	41	40	41	

Respostas dadas pelos 47 estudantes aos 15 itens. 1=acertou , 0= errou





Porcentagem de acertos em cada item.

Q01	Q02	Q03	Q04	Q05	Q06	Q07	Q08	Q09	Q10	Q11	Q12	Q13	Q14	Q15
82	63	78	59	63	51	76	76	63	89	95	89	87	85	87

REFERÊNCIAS BIBLIOGRÁFICAS

- Albert, J. H. (1992). Bayesian Estimation of Normal Ogive Item Response Curves Using Gibbs Sampling. *Journal of Educational Statistics*, 17, 251 - 269.
- Albert, J. H. e Chib, S. (1993). Bayesian Analysis of binary and polytomous response data. *Journal of the American Statistical Association*, 88, 669-679.
- Albert, J.H. e Ghosh, M. (2000). Item response modeling. *Generalized Linear Models: A Bayesian Perspective*(D. Dey, S. Ghosh & Mallick, eds.), Marcel-Dekker, New York. 173-193.
- Andrade, D. F. , Tavares, H. R. e Valle, R. C. (2000). Introdução à Teoria da resposta ao Item : Conceitos e Aplicações. 14º SINAPE : Caxambu, MG.
- Akaike, H. (1973). Information Theory and an extension of the maximum likelihood principle. In B. Petrov & F. Csáki (Eds.) 2nd International Symposium on Information Theory (pp. 267-281). Budapest: Akadémiai Kiadó.
- Arellano-Valle, R., Bolfarine,H., e Iglesias, P.(1994).Preditivistic Interpretation to the multivariate t distribution. *Test/Test*, 3, 1-16.
- Assunção, C. (1999). Estimacão dos parâmetros de modelos da teoria de resposta ao item e aplicações. Mestrado em Estatística - UFMG.
- Azzalini A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statistical*, 12, 171-178.
- Azzalini A. e Capitanio A. (1999). Statistical applications of the multivariate skew normal distributions. *J. R. Statist. Soc. B*, 61, 579-602.
- Azzalini A. e Dalla Valle A. (1996). The multivariate skew-normal distribution. *Biometrika*, 83, 715-726.
- Baker, F.B. (1992). *Item Response Theory - Parameter Estimation Techniques*. New York: Marcel Dekker, Inc.
- Bartholomew,D.J. e Knoot, M. (1999). *Latent variable models and factor analysis*. (2nd ed.). London: Arnold. (Kendall´s Library of Statistics 7).
- Basu, S. e Mukhopadhyay, S. (2000). Binary Response Regression With Normal Scale Mixtures Links. in *Generalized Linear Models: A Bayesian Perspective*,eds. D.K. Dey, S.K. Ghosh, and B.K: Mallick, New York: Marcel Dekker.

- Bazán, J., Espínosa G. e Farro Ch. (2002). Rendimiento y actitudes hacia la matemática en el sistema escolar peruano. In Rodríguez, J., Vargas, S., (eds.). *Análisis de los Resultados y Metodología de las pruebas Crecer 1998. Documento de trabajo 13*. Lima:MECEP-Ministerio de Educación. 55-70.
- Bazán, J., Branco, M. e Bolfarine, H. (2004a). A skew item response model. ISBA 2004 World Meeting, Viña del Mar, Chile, May 23-27, 2004. ISBA (International Society for Bayesian Analysis).
- Bazán, J., Bolfarine, H., e Branco, D. M. (2004b). A new family of asymmetric models for item response theory: A SKEW-NORMAL IRT FAMILY. 49p. *Technical report* (RT-MAE-2004-17). Department of Statistics. University of São Paulo.
- Bazán, J., Bolfarine, H. e Aparecida, R. (2004a). Bayesian estimation via MCMC for probit-normal model in item response theory. 26p. *Technical report* (RT-MAE-2004-15). Department of Statistics. University of São Paulo.
- Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In F. M. Lord & M. R. Novick. *Statistical Theories of Mental Test Scores*, eading, MA : Addison-Wesley.
- Bock, R. D. e Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika* 64,153-168.
- Branco, M.D. and Dey, D.K.(2002). Regression model under skew elliptical error distribution. *The Journal of Mathematical Sciences, Delhi, New Series, 1*, 151-169.
- Branco,M.D. and Dey, D.K. (2001). A class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis, 79*, 99-113.
- Chen, M-H, Shao, Q. M, e Ibrahim, J. G (2000). *Monte Carlo Methods in Bayesian Computation*. New York: Springer Verlag.
- Chen, M-H, Dey, D. K., e Shao, Q-M. (1999). A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association, 94*, 448, 1172-1186.
- Chen, M-H.,Dey, D.K., e Shao, Q-M. (2001). Bayesian analysis of binary data using skewed logit models. *Calcutta Statistical Association Bulletin, 51*, 201-202.
- Chib, S. e Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician, 49*, 327-335.
- Dalla Valle, A. (2004). The skew-normal distribution, in *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Genton, M. G., Ed. Chapman & Hall / CRC, Boca Raton, FL, pp. 3-24.

- Ghosh, M., Ghosh, A., & Chen, Ming-Hui & Agresti, A. (2000). Noninformative priors for one parameter item response models. *Journal of Statistical Planning and Inference*, 88, 99-115.
- Hambleton, R.K., Swaminathan, H. (1985). *Item Response Theory: Principles and Applications*. Boston: Kluwer Academic Publishers.
- Hashimoto, Y. (2002). Motivation and willingness to communicate as predictors of reported L2 use: the Japanese context. *Second Language Studies*, 20 (2), 20-70.
- Henze, N. (1986). A probabilistic representation of the "skew-normal" distribution. *Scandinavian Journal Statistical* 13, 271-275.
- Johnson, V. e Albert, J. (1999). *Ordinal Data Modeling*. New York, MA: Springer-Verlag.
- Johnson (2003). On the use of Heterogenous thresholds ordinal regression models to account for individual differences in response style. *Psychometrika*, 68, 563-583.
- Junker, B. W. e Ellis, J. L. (1997). A characterization of monotone unidimensional latent variable models. *Annals of Statistics*, 25, 1327-1343.
- Kim, S., Cohen, A. S., Baker, F. B., Subkoviak, M. J. e Leonard, T. (1984). An investigation of hierarchical Bayes procedures in item response theory. *Psychometrika*, 59, 405-421.
- Lord, F. M. (1952). *A theory of test scores*. New York: Psychometric Society.
- Lord, F., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Micceri, T. (1989). The Unicorn, The Normal Curve, and Other Improbable Creatures. *Psychological Bulletin*, 105(1), 156-166
- Mislevy, R. J. e Bock, R.D. (1990). *BILOC: Maximum likelihood item analysis and test scoring with logistic models for binary items* [Computer program]. Chicago: Scientific Software International.
- Mislevy, J. M. (1984). Estimating Latent Distributions. *Psychometrika*, Vol. 49, 359 - 381.
- Mislevy, R.J. (1986). Bayes modal estimation in item response models. *Psychometrika*, 51, 177-195.
- Mislevy, R.J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177-196.
- Mislevy, R. J. e Bock, R. D. (1990). *BILOG 3 : Item Analysis and Test Scoring with Binary Logistic Models*. Chicago : Scientific Software, Inc.

Patz, R. J. e Junker, B. W. (1999). A straightforward approach to Markov Chain Monte Carlo methods for item response models. *Journal of Educational and Behavioral Statistics*, 24, 146-178.

Prentice, R. L. (1976). A Generalization of the Probit and Logit Methods for Dose-Response Curves. *Biometrika*, 32,761-768.

Rupp, A., Dey, D. K., e Zumbo, B. (2004) (in press). To Bayes or Not to Bayes, from Whether to When: Applications of Bayesian Methodology To Item Response Modeling. *Structural Equations Modeling*.

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

Sahu, S. K., Dey, D.E. e Branco, M. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics*, Vol. 29,217-232.

Samejima, F. (1997). Departure from normal assumptions: a promise for future psychometrics with substantive mathematical modeling. *Psychometrika*, 62, 4,471-493.

Samejima, F. (2000). Logistic positive exponent family of models: virtue of asymmetric item characteristics curves. *Psychometrika*, 65(3), 319-335.

Spiegelhalter, D. J., Thomas, A., Best, N. G. e Gilks, W.R.(1996). *BUGS 0.5 examples* (Vol. 1 Version 1). Cambridge, UK: University of Cambridge.

Spiegelhalter D., Thomas, A., Best N. e Lunn, D. (2003). *WinBUGS version 1.4* [Computer program]. Imperial College & Medical Research Council Biostatistics Unit, Institute of Public Health, Cambridge University.

Swaminathan, H. e Gifford, J. A. (1982). Bayesian Estimation in the Rasch model. *Journal of Educational Statistics*, 7, 175 - 192.

Swaminathan, H. e Gifford, J. A. (1985). Bayesian Estimation in the two-parameter logistic model. *Psychometrika*, 50, 349 - 364.

Swaminathan, H. & Gifford, J. A. (1986). Bayesian Estimation in the three-parameter logistic model. *Psychometrika*, 51, 589 - 601.

Tanner, M. A. (1996). *Tools for Statistical Inference*. Springer-Verlag: Heidelberg.

Tavares, H. R. *Teoria da Resposta ao Item para Dados Longitudinais*. Tese de Doutorado. São Paulo: IME/USP.

Taylor, J. e Siqueira, A. (1996). The cost of Adding Parameters to a Model. *Journal of the Royal Statistical Society*, 58,593-607.

Tsutakawa, R. K. e Lin, H. Y. (1986). Bayesian estimation of item response curves. *Psychometrika*, 51, 251-267.

Tsutakawa, R. M. (1984). Estimation of two-parameter logistic item response curves. *Journal of Educational Statistics*, 9, 263-276.

Valle, R. C. (1999). *Teoria da Resposta ao Item*. Dissertação de Mestrado. São Paulo : IIE/USP.

Van der Linden, W.J. e Hambleton, R. K. (1997). *Handbook of Modern Item response Theory*. New York: Springer-Verlag.

Vianna, H. (2003). Avaliações Nacionais em Larga Escala: análise e propostas. *Estudos em avaliação Educacional*, 27, 41-76.