Análise bayesiana em modelos TRI
de três parâmetros

Katia Antunes Marques

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

Área de Concentração: **Estatística**

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

*Durante a elaboração deste trabalho o autor
recebeu apoio financeiro do CNPq*

São Paulo, maio de 2008

Agradecimentos

Não teríamos nascido sem que houvesse Amor entre duas pessoas; não teríamos crescido se não tivéssemos alguém que trabalhasse para nos alimentar, vestir e educar; não teríamos estudado se alguém não tivesse dito um dia: "você precisa fazer isso.". Quando atingimos uma certa idade, até tentamos caminhar sozinhos, mas logo percebemos que isso é praticamente impossível. Logo surgem os primeiros desafios, barreiras e dificuldades. Só ficamos empé após uma queda graças àqueles que aparecem para nos levantar, só ultrapassamos os obstáculos, se alguém mais forte do que nós, que caminha ao nosso lado no momento, nos oferece os braços para nos apoiar. Sim, muitas vezes ficamos parados e perdidos. Podemos ter tudo nas mãos, mas precisamos de alguém para nos fazer enxergar isso. Outras vezes falta-nos coragem para seguir determinada direção e, tomamos a decisão quando ouvimos: "siga por aqui, você é capaz. ". Se algum dia ouvimos de alguém que não somos capazes, então outra voz se sobrepõe: "vá em frente!". Já dizia Albert Eistein que algo só é impossível até que alguém duvide e prove o contrário.. Sempre contamos com o apoio de alguém que esteja ao nosso lado. E assim, um belo dia nos formamos, concluímos o primeiro, o segundo, o terceiro graus e, chegamos à defesa de nossa primeira tese. Será que se ninguém estivesse conosco nesses momentos de conquista, tantos anos de luta fariam sentido? É por isso que dedico e agradeço este trabalho a todos os que me ajudaram a construí-lo, que fizeram parte dele de alguma forma, que estiveram

presentes em alguma ou algumas de suas etapas. Acima de tudo, agradeço a Deus por mais este sonho realizado, pela força, coragem, determinação e fé, por todas as ferramentas a mim fornecidas para chegar até aqui. A meus pais, Alfredo e Maria José, a quem devo tudo, a começar por minha existência. A meus irmãos, Renato e Ricardo, muito presentes desde minha infância, agradeço por todo o carinho, proteção e essa amizade tão bonita entre nós. A meus avós, Merciano e Maria, obrigada por tantos ensinamentos! Ao Sam, os olhos que ganhei, por seu carinho, seu amor, e seu maravilhoso trabalho, quem me guiou até a linha de chegada de mais esta etapa cumprida. Sem palavras para agradecer quem se tornou uma parte inseparável de meu corpo! Aos demais familiares que, de uma forma ou de outra, me deram apoio e incentivo. Aos amigos que se tornaram pessoas muito especiais, quem adotei eternamente como membros de minha família. Meus padrinhos de crisma e de formatura da faculdade, bem como seus familiares. A meu namorado, pelo excessivo carinho e compreensão nas horas de estudo. A todos os amigos da infância na praia, da escola, da graduação, do mestrado, da Estação Ciências, do INSS e do Itaú, além de outros lugares, pelos momentos de descontração que passamos juntos, pela paciência, pela leitura dos livros, pelos estudos que muitas vezes nos tomaram domingos, feriados e madrugadas inteiras. Em especial aos que se tornaram verdadeiros amigos: Priscila, Karina, Camila, Lucinara, Fernanda, Cristiana, Lilian, Luanna, Fabiana, Marcelo, Márcia, Filipe, Luciana e Ronald. Aos funcionários e professores do colégio João XXIII e dos departamentos de matemática, computação e estatística do IME/USP, pela prontidão em diversos momentos, esclarecimentos prestados e ensinamentos recebidos. Às professoras Elisabeth Bigena e Márcia Branco, pela força e incentivo incondicional à continuidade de meus estudos, grandes responsáveis por esta conquista, tornaram visível a todos o que nem eu mesma era capaz de enxergar. A todos os colegas da escola e da USP, pela companhia, pelas experiências e pelas conversas de corredor. Aos membros do instituto IRIS e a todos da *Leader Dogs school*, por

tornarem real o sonho de ter um cão-guia. Ao Regis e à Layla que, juntamente com o Sam, trouxeram de volta a sensação de enxergar um pouco mais longe, mais vontade de seguir em frente e principalmente mais alegria a minha vida. Aos colegas que conheci na Estação Ciência e na Gerência Executiva Oeste do INS. Ao CNPq pelo apoio financeiro em grande parte do mestrado. Aos supervisores, gestores e pares da diretoria de Crédito-Cartões do banco Itaú, pela receptividade, pela paciência, por todos os ensinamentos, compreensão, apoio e incentivo nos últimos meses. Aos membros da banca examinadora por disponibilizarem seus tempos avaliando este trabalho e pelas valiosas sugestões e comentários, em especial à professora Márcia Branco, pela excelente orientação desde a graduação, na iniciação científica e no mestrado. Obrigada pelo desafio proposto, pela sua dedicação, paciência, compreensão, apoio,, por fim, pela crença em minha capacidade. Finalmente, não poderia deixar de registrar minha homenagem mais do que especial àquela que, com probabilidade 1, esteve presente em todas as etapas de minha vida. Muito além de meu nascimento, participou de meus primeiros passos, das primeiras trombadas, das quedas que levantou comigo, das perdas que choramos juntas, das barreiras que enfrentamos, das formaturas e outras vitórias que conquistamos... Seus olhos, como se fossem os meus, iam transformando papeis em conhecimento, ao passo que as letras iam sumindo de meu alcance. Foi quem, por muitas vezes me fez voltar a acreditar que nada seria impossível, mesmo com tantas dificuldades. Por sua coragem, dedicação e paciência, por seu colo, por seu carinho, por seus ensinamentos, por seu amor em tempo integral, resumindo, por ser a maior responsável por eu estar aqui, ter chegado onde cheguei, conquistado o que conquistei, tendo lutado por tudo o que precisei, sem palavras para agradecer: obrigada mãe!

Resumo

Neste trabalho discutimos a análise bayesiana em modelos TRI (Teoria da Resposta ao Item) de três parâmetros com respostas binárias e ordinais, considerando a ligação probito. Em ambos os casos usamos técnicas baseadas em MCCM (método de Monte Carlo baseado em Cadeias de Markov) para estimação dos parâmetros dos itens. No modelo com respostas binárias, consideramos dois conjuntos de dados resultantes de provas com itens de múltipla-escolha. Para esses dados, foi feito um estudo da sensibilidade à escolha de distribuições *a priori*, além de uma análise das estimativas *a posteriori* para os parâmetros dos itens: discriminação, dificuldade e probabilidade de acerto ao acaso. Um terceiro conjunto de dados foi utilizado no estudo do modelo com respostas ordinais. Estes dados são provenientes de uma disciplina básica de estatística, onde a prova contém itens dissertativos. As respostas foram classificadas nas categorias: certa, errada ou parcialmente certa. Utilizamos o programa WinBugs para a estimação dos parâmetros do modelo binário e a função *MCMCordfactanal* do programa *R* para estimar os parâmetros do modelo ordinal. Ambos os *softwares* são não proprietários e gratuitos (livres).

Abstract

In this dissertation the bayesian analysis for three parameters IRT (Item Response Theory) models with binaries and ordinals responses, considering the probit model, was discussed. For both cases, binary and ordinal, techniques based on MCCM (Monte Carlo Markov Chain) were used to estimate the items parameters. For binary response model, was considered two data sets from tests with multipla choices items. For these two data sets, a sensibility study of the *priori* distributions choice was considered, and also, an analyses of a *posteriori* estimates of the items parameters: discrimination, difficulties and guessing. A third data set is used to illustrate the ordinal response model. This come from an elementar statistical course, where a test with open items is considered. The responses are classified in the following categories: correct, wrong or partial correct. The WinBugs software was used to estimate the parameters for the binary model and, for the ordinal model was considered the function *MCMCordfactanal* from *R* program.

Sumário

1	Introdução	1
2	O modelo de três parâmetros com respostas binárias	6
2.1	Estimação dos Parâmetros	9
2.2	A escolha das distribuições <i>a priori</i>	12
2.3	Inferência Bayesiana usando MCCM para o modelo 3PN	15
2.4	Modelo de Dados Aumentados para o 3PN	16
2.5	Construindo a amostra MCCM	19
3	Aplicações	21
3.1	Aplicação I: prova de matemática para o ensino básico	22
3.1.1	Um estudo sobre a influência do salto na construção da amostra MCCM	22
3.1.2	Análise de sensibilidade às distribuições <i>a priori</i>	24

3.1.3	Análise das estimativas	28
3.2	Aplicação II: simulado em um cursinho pré-vestibular	32
3.2.1	Um estudo sobre a influência do salto na construção da amostra MCCM	32
3.2.2	Análise de sensibilidade às distribuições <i>a priori</i>	34
3.2.3	Análise das estimativas	37
3.3	Similaridades e diferenças entre as aplicações I e II	39
4	Modelos com respostas ordinais	40
4.1	Estimação dos parâmetros	44
4.2	Aplicação III: Provas do curso de Noções de Estatística	45
5	Considerações finais	48
A	Tabelas: Aplicação I	53
B	Tabelas: Aplicação II	68
C	Tabela: Aplicação III	77

Lista de Tabelas

3.1	Valores de CCK.	28
3.2	Estimativas <i>a posteriori</i> para os parâmetros α e β , ordenadas pela mediana	30
3.3	Estimativas <i>a posteriori</i> para os parâmetros de acerto ao acaso (δ), ordenadas pela mediana.	31
3.4	Tabela de CCK.	36
3.5	Estimativas <i>a posteriori</i> para os parâmetros α e β , ordenadas pela mediana.	37
3.6	Resumo das estimativas <i>a posteriori</i> para os parâmetros de acerto ao acaso (δ), ordenadas pela mediana.	38
A.1	Médias <i>a posteriori</i> para α e β * com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 5000. Aplicação I.	54
A.2	Médias <i>a posteriori</i> para δ com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 5000. Aplicação I.	55

A.3	Médias <i>a posteriori</i> para α e β^* com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 10000. Aplicação I.	56
A.4	Médias <i>a posteriori</i> para δ com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 10000. Aplicação I.	57
A.5	Relações entre SD e MC Error para δ com os dados da aplicação I. . .	58
A.6	Relações entre SD e MC Error para α e β^* dos dados da aplicação I. . .	59
A.7	Médias <i>a posteriori</i> para δ considerando diferentes distribuições <i>a priori</i> . Aplicação I.	60
A.8	Médias <i>a posteriori</i> para α e β^* considerando diferentes distribuições <i>a priori</i> . Aplicação I.	61
A.9	Medianas <i>a posteriori</i> para δ considerando diferentes distribuições <i>a priori</i> . Aplicação I.	62
A.10	Medianas <i>a posteriori</i> para α e β^* considerando diferentes distribuições <i>a priori</i> . Aplicação I.	63
A.11	Desvio padrão <i>a posteriori</i> para δ com diferentes distribuições <i>a priori</i> . Aplicação I.	64
A.12	Desvio padrão <i>a posteriori</i> para α e β^* com diferentes distribuições <i>a priori</i> . Aplicação I.	65
A.13	Percentis das estimativas dos parâmetros δ para P_4 e P_7 . Aplicação I. . .	66
A.14	Itens ordenados pela mediana <i>a posteriori</i> de δ usando diferentes distribuições <i>a priori</i> . Aplicação I.	66

A.15	Itens ordenados pela mediana <i>a posteriori</i> de α usando diferentes distribuições <i>a priori</i> . Aplicação I.	67
A.16	Itens ordenados pela mediana <i>a posteriori</i> de β^* usando diferentes distribuições <i>a priori</i> . Aplicação I.	67
B.1	Médias <i>a posteriori</i> para diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 5000. Aplicação II.	69
B.2	Médias <i>a posteriori</i> para diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 10000. Aplicação II.	70
B.3	Relações entre SD e MC Error para os dados da Aplicação II.	71
B.4	Média <i>a posteriori</i> considerando diferentes distribuições <i>a priori</i> . Aplicação II.	72
B.5	Mediana <i>a posteriori</i> considerando diferentes distribuições <i>a priori</i> . Aplicação II.	73
B.6	Desvio padrão <i>a posteriori</i> para diferentes distribuições <i>a priori</i> . Aplicação II.	74
B.7	Percentis das estimativas dos parâmetros δ para P_4 e P_1 . Aplicação II.	75
B.8	Itens ordenados pela mediana <i>a posteriori</i> de δ usando diferentes distribuições <i>a priori</i> , aplicação II.	75
B.9	Itens ordenados pela mediana <i>a posteriori</i> de α usando diferentes distribuições <i>a priori</i> . Aplicação II.	76

B.10	Itens ordenados pela mediana <i>a posteriori</i> de β_* usando diferentes distribuições <i>a priori</i> . Aplicação II.	76
C.1	Médias <i>a posteriori</i> para α_{js} , β_{js} e β_{*js} , ordenadas. Aplicação III . .	78

Capítulo 1

Introdução

Atualmente, por regra de escolas de muitos países, inclusive do Brasil, trabalhos em grupo, entregas de pesquisas de temas específicos fornecidos por professores e provas são fatores contribuintes para as notas finais dos alunos. Estas últimas são os principais instrumentos de medida de conhecimento. Uma prova é formada por um conjunto de itens que são respondidos individualmente.

A aplicação de uma prova tem por finalidade, por exemplo: ranquear um grupo de pessoas, de acordo com seus conhecimentos no tema abordado; selecionar capacitados a iniciarem um curso superior, exercer determinado cargo público, ou simplesmente medir a eficiência do ensino em determinado curso. As provas devem estar adequadas a seus propósitos, a fim de fornecerem com justiça e segurança os resultados pretendidos.

Entre as metodologias utilizadas para analisar os itens que compõe uma prova destacam-se a Teoria Clássica dos Testes (TCT) e a Teoria da Resposta ao Item (TRI). A TCT, também conhecida como Teoria Clássica das Medidas é baseada em resultados obtidos em provas através de escores brutos ou padronizados. Sua principal

característica é que as análises e interpretações estão sempre associadas à prova como um todo, ou seja, dependem apenas do conjunto de itens que a compõe. Por este motivo sua aplicabilidade fica limitada, já que não se pode comparar indivíduos que tenham sido submetidos a provas diferentes.

A TRI é uma metodologia mais sofisticada e precisa que permite não só a avaliação pontual, mas sobretudo, a construção de escalas de habilidades que possam levar a um acompanhamento do progresso do conhecimento adquirido pelos alunos ao longo do tempo. Esta metodologia vem sendo progressivamente introduzida em nosso meio e é um instrumento poderoso nos processos quantitativos de avaliação educacional. O que ela sugere são formas de representar a relação entre a probabilidade de um indivíduo dar uma certa resposta a um item e seus traços latentes, proficiências ou habilidades na área de conhecimento avaliada.

Contudo, a aplicabilidade da TRI ainda encontra algumas dificuldades no campo da estimação dos parâmetros do modelo, tanto do ponto de vista teórico quanto do ponto de vista computacional. Alguns desses aspectos serão discutidos neste trabalho.

Uma das grandes vantagens da TRI sobre a TCT é que ela permite a comparação entre populações, desde que submetidas a provas que tenham alguns itens comuns, ou ainda, a comparação entre indivíduos da mesma população submetidos a provas totalmente diferentes. Isto porque uma das principais características da TRI é que ela tem como elementos centrais os itens, e não a prova como um todo, como é o caso da TCT. Pode-se então, por exemplo, avaliar o desenvolvimento de uma determinada série de um ano para outro ou comparar o desempenho entre escolas públicas e privadas, ver Andrade & Vale (2000).

Os primeiros modelos de resposta ao item são da década de 50. Modelos em que se considerava que uma única habilidade, de um único grupo, estava sendo medida

por um teste onde os itens eram corrigidos de maneira binária (certo ou errado).

O primeiro modelo TRI, introduzido formalmente por Lord (1952), foi o modelo de ogiva normal com respostas binárias. Nesse modelo o acerto a um determinado item foi relacionado de forma funcional à habilidade do indivíduo, esta por sua vez, considerada uma quantidade latente não observável unidimensional. Além disso, a probabilidade de acerto de um item também foi modelada como função de parâmetros associados aos itens.

Os modelos da TRI se diferenciam pelo número de parâmetros que utilizam para descrever o item. Eles são conhecidos como modelos de 1, 2 e 3 parâmetros, os quais consideram, respectivamente:

- I. somente a dificuldade do item;
- II. a dificuldade e a discriminação;
- III. a dificuldade, a discriminação e a probabilidade de resposta correta dada por indivíduos de baixa habilidade.

Assim como para o modelo logístico, ver Andrade & Vale (2000), o modelo de ogiva normal de dois parâmetros considera p_{ij} a probabilidade do i -ésimo indivíduo acertar o j -ésimo item como:

$$p_{ij}(\cdot) = \Phi(\eta_{ij}), i = 1, \dots, n, j = 1, \dots, k, \quad (1.1)$$

onde Φ é a função de distribuição de probabilidades acumulada na normal padrão e $\eta_{ij} = \alpha_j(\theta_i - \beta_j)$ é uma função das habilidades do indivíduo e depende também de parâmetros associados aos itens.

Birnbaum (1968) considerou a seguinte modelagem alternativa:

$$p_{ij} = e^{\eta_{ij}} / (1 + e^{\eta_{ij}}). \quad (1.2)$$

A expressão (1.2) corresponde a função acumulada da distribuição logística e portanto, este é denominado modelo logístico.

Outra possível variação dos modelos TRI é o denotado por Andrade & Vale (2000) como modelos para itens não dicotômicos. Nesses casos a resposta ao item não necessita ser binária, ela pode ser ordinal. É comum obtermos na correção de uma questão a seguinte classificação: certo, errado e parcialmente certo. Esta última categoria pode ainda ser subdividida em subcategorias. Respostas ordinais em geral aparecem em provas com questões dissertativas.

Os modelos considerados neste trabalho serão os modelos de três parâmetros com respostas binárias e ordinais. A metodologia de inferência utilizada será a bayesiana.

No capítulo 2 descrevemos modelos para dados binários de três parâmetros e discutimos como conduzir o processo inferencial sob a abordagem bayesiana. No capítulo 3, apresentamos uma análise de sensibilidade considerando dois conjuntos de dados, resultantes de provas de matemática. O primeiro obtido de exame com 14 questões e 4 alternativas cada, aplicado a 131 alunos de sexta série de escolas peruanas (dados obtidos de Bazan 2005). O segundo, prova de um simulado proveniente do cursinho pré-vestibular Ângulo de São Paulo, contendo 10 questões de múltipla-escolha, com 5 alternativas cada e 277 alunos respondentes (dados obtidos de Rios, 2007).

No capítulo 4 apresentamos um estudo do modelo de 3 parâmetros de ogiva normal considerando respostas ordinais agrupadas nas categorias: certo, parcialmente certo e errado. Analisamos um conjunto de dados obtidos da prova dissertativa, aplicada na

disciplina MAE0116, Noções de Estatística, cursada por alunos do curso de biologia da Universidade de São Paulo (USP).

Finalmente, no capítulo 5 nossas considerações a respeito das teorias e modelos, bem como suas aplicações estudadas ao longo deste trabalho.

Capítulo 2

O modelo de três parâmetros com respostas binárias

Considere n indivíduos respondendo k itens de uma prova. Seja $y_{i,j} = 1$ se o i -ésimo indivíduo acerta o j -ésimo item e $y_{i,j} = 0$, caso contrário, $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, k$. Assumimos que $y_{i,j} \sim \text{Bernoulli}(p_{i,j})$ e que a probabilidade do i -ésimo indivíduo acertar o j -ésimo item é:

$$p_{ij} = \delta_j + (1 - \delta_j)F(\eta_{i,j}) \quad (2.1)$$

$$\eta_{ij} = \alpha_j(\theta_i - \beta_j).$$

F é alguma função de distribuição de probabilidades. Além disso:

- θ_i é uma variável latente associada ao indivíduo i , representando sua habilidade ou traço latente, medidos pela prova;

- β_j é o parâmetro de dificuldade (ou de posição) do item j , medido na mesma escala da habilidade;
- α_j é o parâmetro de discriminação (ou de inclinação) do item j , com valor proporcional à inclinação da curva característica do item CCI, no ponto β_j . É a medida que indica o quanto um item consegue distinguir os indivíduos em relação às suas habilidades;
- δ_j é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente ao item j (muitas vezes referido como a probabilidade de acerto casual).

Uma suposição adicional de independência é usualmente considerada, isto é, condicional a sua habilidade as respostas dos indivíduos aos itens são independentes.

Fixado j , a probabilidade de acerto ao item é uma função das habilidades. Tal função é denominada Curva Característica do Item (CCI), usualmente uma função monótona crescente. Observe que quanto maior for a habilidade do respondente, maior será a probabilidade de responder corretamente ao item.

O modelo proposto baseia-se no fato de que indivíduos com maior habilidade possuem maior probabilidade de acertar o item e que esta relação não é linear. Pode-se perceber a partir do gráfico que a CCI tem forma de “S”. Uma descrição detalhada sobre a CCI pode ser encontrada em Andrade, Tavares & Valle (2000).

A escala usada para medir as habilidades é arbitrária. O importante são as relações de ordem existentes entre seus pontos e não necessariamente sua magnitude. O parâmetro β é medido na mesma unidade das habilidades, por exemplo, se habilidade é suposta ter uma distribuição normal padrão, então os valores para β_j com alta probabilidade estão entre -3 e 3. Os parâmetros β_j também representam a habilidade

necessária para uma probabilidade de acerto igual a $(1 + \delta_j)/2$; quanto maior o valor de β_j , mais difícil é o item, e vice-versa.

O parâmetro δ_j não depende da escala, pois se trata de uma probabilidade, e assim assume sempre valores entre 0 e 1. Ele representa a probabilidade de um aluno com baixa habilidade responder corretamente ao item **j** e é muitas vezes referido como a probabilidade de acerto ao acaso. Então, quando não é possível acertar sem o mínimo conhecimento do conteúdo, δ_j é igual a 0 e β_j representa o ponto na escala das habilidades onde a probabilidade de acertar o item é 0,5. Isso ocorre em itens dissertativos. Em itens de múltipla-escolha é esperado que esse parâmetro seja diferente de zero.

O parâmetro α_j é proporcional à derivada da tangente da curva característica no ponto de inflexão. Assim, itens com α_j negativo não são esperados sob esse modelo, uma vez que indicariam que a probabilidade de responder corretamente ao item diminui com o aumento da habilidade.

Baixos valores de α_j indicam o pouco poder de discriminação do item, ou seja, alunos com habilidades bastante diferentes têm aproximadamente a mesma probabilidade de responder corretamente ao item. Ao contrário, valores elevados indicam alto poder de discriminação do item. Estes dividem os alunos basicamente em dois grupos: os que possuem habilidades abaixo do valor do parâmetro β e os que possuem habilidades acima do valor do parâmetro β_j .

O modelo de ogiva normal, também denominado modelo Probito-Normal, com três parâmetros será denotado por 3PN. Neste caso

$$p_{ij} = \delta_j + (1 - \delta_j)\Phi(\eta_{ij}) \quad (2.2)$$

O modelo Logito-Normal de três parâmetros será denotado por 3LN e dado por

$$p_{ij} = \delta_j + (1 - \delta_j) \frac{e^{\eta_{ij}}}{1 + e^{\eta_{ij}}}. \quad (2.3)$$

Devido à semelhança nos resultados para os dois modelos de três parâmetros com respostas binárias, e considerando o fato de que o modelo 3PN será mais facilmente generalizado para respostas ordinais (ver capítulo 4), optamos exclusivamente pelo modelo dado em (2.2).

2.1 Estimação dos Parâmetros

A estimação dos parâmetros é uma das etapas mais importantes da TRI. A probabilidade de uma resposta correta a um determinado item depende tanto das habilidades dos indivíduos, quanto dos parâmetros que caracterizam os ítems. Em geral, ambos são desconhecidos. Apenas as respostas dos indivíduos aos itens do teste são conhecidas.

Neste modelo o problema está em estimar dois tipos de parâmetros, as habilidades e os parâmetros dos itens. Na TRI, o processo de estimação dos parâmetros dos itens é conhecido como calibração, Andrade, Tavares & Valle (2000).

No processo de estimação podemos encontrar diferentes situações do ponto de vista teórico. Segundo Andrade, Tavares & Valle (2000), a situação mais comum é aquela em que desejamos estimar os parâmetros dos itens e as habilidades dos indivíduos simultaneamente. Esta será a abordagem adotada neste trabalho, embora, o principal interesse seja a estimação dos parâmetros dos itens.

Inicialmente, a estimação era feita através do método da máxima verossimilhança

conjunta, que envolve um número muito grande de parâmetros a serem estimados simultaneamente e, por conseqüência, grandes problemas computacionais. Em 1970, Bock & Lieberman introduziram o método da máxima verossimilhança marginal para a estimação dos parâmetros em duas etapas. Na primeira etapa estimam-se os parâmetros dos itens, assumindo-se uma certa distribuição para as habilidades, e na segunda etapa, assumindo os parâmetros dos itens conhecidos, estimam-se as habilidades. Apesar do avanço que esse método trouxe para o problema, ele requeria que todos os parâmetros dos itens fossem estimados simultaneamente. Em 1981, Bock & Aitkin propuseram uma modificação no método acima, utilizando o algoritmo EM de Dempster, Laird & Rubin (1977), de modo a permitir que os itens pudessem ter seus parâmetros estimados em separado, facilitando em muito o aspecto computacional do processo de estimação clássico.

Mais recentemente, métodos bayesianos foram propostos para contornar problemas como a estimação de itens respondidos corretamente ou incorretamente por todos os indivíduos. Além disso, há a possibilidade de que as estimativas clássicas dos parâmetros dos itens caiam fora do intervalo esperado, tal como valores de α_j negativos, ou valores de δ_j fora do intervalo $[0, 1]$, Andrade, Tavares & Valle (2000).

A função de verossimilhança associada ao modelo TRI com respostas binárias é dado por:

$$L(Y, (\alpha, \beta, \delta) \mid D_{obs}) = \prod_{j=1}^k \prod_{i=1}^n p_{ij}^{y_{ij}} (1 - p_{ij})^{1-y_{ij}}. \quad (2.4)$$

Com p_{ij} dado por (2.2) ou (2.3) dependendo da CCI considerada.

Note que dois valores diferentes do parâmetro podem levar à mesma verossimilhança, por exemplo, considere os vetores de parâmetros

$$(\theta_1, \alpha_1, \beta_1, \delta_1) = (2, a, 1, d) \text{ e } (\theta_2, \alpha_2, \beta_2, \delta_2) = (3, a, 2, d).$$

Note que nesse caso $\alpha_1(\theta_1 - \beta_1) = \alpha_2(\theta_2 - \beta_2)$, como $\delta_1 = \delta_2$, temos que as probabilidades especificadas em (2.2) e (2.3) são iguais para os dois vetores de parâmetros. Concluindo que as funções de verossimilhanças são iguais, isto é,

$$L(2, a, 1, d) = L(3, a, 2, d), \forall b, \forall d.$$

Neste caso o modelo é denominado não identificável. Modelos não identificáveis dificultam o uso da metodologia de máxima verossimilhança na estimação dos parâmetros. Suposições adicionais são necessárias para resolver o problema de identificabilidade do modelo. Na TRI esse problema pode ser resolvido através da especificação de uma distribuição de probabilidades para as habilidades.

Os métodos de estimação na TRI podem ser agrupados em três categorias:

I. Estimação freqüencista (ou clássica), que tem sido dominada pelo uso da estimação de máxima verossimilhança;

II. Estimação parcialmente bayesiana, que é feita através da estimação bayesiana marginal, proposta por Mislevy (1986a), uma generalização da proposta de Bock & Aitkin (1981). Consiste em estabelecer distribuições *a priori* para os parâmetros de interesse, construir uma nova função denominada distribuição *a posteriori* e estimar os parâmetros de interesse com base em alguma característica dessa distribuição.

III. Estimação bayesiana e o uso de métodos de Monte Carlo baseado em cadeias de Markov (MCCM), inicialmente proposta por Albert (1992).

Para o desenvolvimento dos métodos bayesianos, além da especificação da função

de verossimilhança, é necessário atribuir distribuições de probabilidades *a priori* para todos os parâmetros do modelo. O uso da fórmula de Bayes nos permite obter a distribuição *a posteriori*, a partir da combinação das informações das distribuições *a priori* e da função de verossimilhança. Seja $\lambda = (\alpha, \beta, \delta, \theta)$, então a função densidade de probabilidades *a posteriori* é dada por

$$h(\lambda \mid y) = \frac{L(\lambda : y)f(\lambda)}{c}$$

onde c é uma constante numérica (a integral da função de verossimilhança multiplicada pela distribuição *a priori* dos parâmetros) e $f(\lambda)$ é uma função de densidade de probabilidade *a priori*.

2.2 A escolha das distribuições *a priori*

A Inferência Bayesiana tem entre seus fundamentos o fato de possuir um mecanismo formal para agregar informação externa ao experimento, chamada informação *a priori*. Esta informação está em poder de um especialista do problema, que pode ser um estatístico, pesquisador ou qualquer pessoa interessada em fazer afirmações sobre entidades desconhecidas de maneira racional. Boa parte desta informação é de caráter subjetivo, e pode estar radicada em fontes objetivas (dados históricos, literaturas) ou em conclusões pessoais tiradas pelo especialista (experiência, conhecimento tácito).

A quantificação probabilística destas crenças *a priori*, especificando a distribuição *a priori*, recorre a conceitos de probabilidade estranhos à Inferência Frequentista.

Um dos pontos centrais do revés entre as escolas bayesiana e clássica é o fato

de os bayesianos aceitarem incluir informações subjetivas ao abordar um problema, fato repudiado pelos freqüencistas. Lindley (1990) sugere que se um clássico diz a um bayesiano “*where did you get that prior?*“, o bayesiano deve responder “*where did you get that sample space?*“, (ver Paulino C. D., Turkman M. A. A. & Murteira B. , 2003), convidando de certa maneira os freqüencistas a aceitarem a utilização de informações subjetivas. Segundo Gelman A., Carlin J. B., Stern H. S. & Rubin D. B. (2004), , “Todos os métodos estatísticos que utilizam probabilidade são subjetivos no sentido de dependerem de idealizações matemáticas do mundo.“.

Entretanto não só as distribuições *a priori* informativas se fazem presentes. Em diversos casos não se tem uma opinião subjetiva sobre um parâmetro ou vetor de parâmetros, situação conhecida como “ignorância *a priori* (Paulino C. D., Turkman M. A. A. & Murteira B., 2003). Distribuições que conseguem representar a “ignorância“ em algum sentido, são chamadas de *prioris* vagas, difusas ou **não-informativas**. Poderia-se dizer que o pesquisador ou o estatístico deixa os dados falarem por si mesmos, buscando a situação em que a distribuição *a posteriori* seja afetada minimamente por informações externas aos dados (Gelman A., Carlin J. B., Stern H. S. & Rubin D. B., 2004). Distribuições *a priori* não-informativas são freqüentemente úteis.

Além de as distribuições *a priori* dos parâmetros do modelo fornecerem informação ou não, também podem ser próprias ou impróprias, características importantes sobre as quais falaremos a seguir.

Uma distribuição de probabilidade é dita **própria** quando não depende dos dados e sua integral sobre o espaço paramétrico é 1. Se ela integra qualquer valor finito positivo, é chamada **densidade não normalizada**, podendo ser renormalizada, ou seja, multiplicada por uma constante apropriada para integrar 1. Se uma função de

probabilidade integra infinito, ela é chamada de **imprópria**.

Distribuições *a priori* próprias garantem que as distribuições *a posteriori* para o modelo sejam próprias. Segundo Albert & Ghosh (2000), a escolha de uma distribuição *a priori* própria para o traço latente resolve os problemas de identificação.

Distribuições *a priori* impróprias podem resultar ou não em distribuições *a posteriori* próprias. Há problema quando a distribuição *a posteriori* do parâmetro de interesse é imprópria, pois inviabiliza naturalmente a realização de inferências.

Distribuições *a priori* não informativas podem ser impróprias ou próprias. Estas levam a distribuições *a posteriori* próprias, mas aquelas nem sempre o fazem.

Ainda que se tenha distribuições *a priori* próprias para os parâmetros, em modelos mais complexos temos unicamente o núcleo da distribuição conjunta. Há um grande problema de integração, dificuldades na construção das marginais e, conseqüentemente na obtenção da distribuição *a posteriori*.

Para solucionar este problema existem métodos computacionais bastante eficientes, como por exemplo, os métodos MCCM, sobre os quais tratamos na próxima seção.

Para a variável latente θ_i consideramos a distribuição normal padrão como distribuição *a priori*. Do ponto de vista clássico também é usada a suposição de normalidade para os θ_i . No entanto, com interpretação diferente, esta não é considerada uma distribuição *a priori* e é necessário encontrar uma outra justificativa para a aleatoriedade dos traços latentes.

No contexto puramente bayesiano é necessário também atribuir distribuições de probabilidade *a priori* para os parâmetros dos itens. Vamos denotá-las por $f(\alpha, \beta, \delta) = \prod_{j=1}^k h(\alpha_j, \beta_j, \delta_j)$. Supondo independência *a priori* entre os parâmetros dos itens:

$$h(\alpha_j, \beta_j, \delta_j) = h_1(\alpha_j)h_2(\beta_j)h_3(\delta_j)$$

Assim,

$$h(\theta, \alpha, \beta, \delta) = \prod_{i=1}^n \phi(\theta_i) \prod_{j=1}^k h_1(\alpha_j)h_2(\beta_j)h_3(\delta_j)$$

Como há uma grande dificuldade para encontrarmos os resultados pretendidos analiticamente, devido à complexidade do modelo, sobretudo pelo número de parâmetros a serem estimados, vamos utilizar algoritmos de simulação do método MCMC. Tais algoritmos estão implementados no aplicativo bayesiano WinBUGS, de uso livre e já bastante utilizado nos meios acadêmicos, o qual pode ser uma boa alternativa aos programas clássicos usualmente considerados para a TRI.

A versão do WinBUGS utilizada foi 1.4, e os programas foram executados em computador com processador core 2 duo 1.0 gb e 1.73 ghz.

2.3 Inferência Bayesiana usando MCMC para o modelo 3PN

O que denominamos MCMC é um conjunto de algoritmos de simulação de uma distribuição multivariada usualmente desconhecida. Tais métodos são baseados na construção de uma Cadeia de Markov cuja distribuição estacionária da cadeia é a distribuição multivariada de interesse. No contexto da Inferência Bayesiana, essa distribuição é a distribuição *a posteriori*. Assim, estatísticas da distribuição teórica de interesse (desconhecida) podem ser estimadas através das correspondentes estatísticas

da amostra aleatória simulada.

O algoritmo de Gibbs é um exemplo na classe MCCM, essencialmente um esquema iterativo de amostragem de uma Cadeia de Markov cujo núcleo é formado pelas distribuições condicionais completas, e a distribuição limite é a densidade de interesse. Para mais informações sobre esses algoritmos ver Gamerman e Lopes (2006)

Para construção de um algoritmo de Gibbs muitas vezes considera-se um conjunto de variáveis auxiliares no modelo. Esses algoritmos são conhecidos como Data Augmented Gibbs Sampler (DAGS) ou dados aumentados usando o amostrador de Gibbs (DAAG).

2.4 Modelo de Dados Aumentados para o 3PN

Albert (1992) foi o primeiro trabalho que utilizou o DAAG para o modelo de ogiva normal. Bazan (2005) e Rios (2007) também consideraram o algoritmo de dados aumentados para obtenção das estimativas *a posteriori* no modelo 2PN. Aqui vamos desenvolver o DAAG para o modelo 3PN binário, utilizando a construção feita por Beguin and Glas (2001).

Neste caso, o modelo de dados aumentados é um modelo que considera não só os dados observados, mas também dois conjuntos de variáveis latentes auxiliares. O primeiro, denotado por z é um vetor composto por todos os z_{ij} , $i = 1, \dots, n$ e $j = 1, \dots, k$, tais que:

$$z_{ij} \mid \theta, \alpha, \beta, \delta \sim N(\eta_{ij}, 1). \quad (2.5)$$

O segundo, denotado por w , que também é um vetor composto dos w_{ij} , $i=1, \dots,$

n e $j=1, \dots, k$. $w_{ij} = 1$, se o i -ésimo sujeito conhece a resposta do j -ésimo item, e $w_{ij} = 0$, caso contrário. As probabilidades são dadas por

$$P(w_{ij} = 1) = P(z_{ij} > 0) \quad (2.6)$$

e

$$P(w_{ij} = 0) = P(z_{ij} \leq 0) \quad (2.7)$$

Note que quando $w_{ij} = 1$ necessariamente $y_{ij} = 1$, isto, se o sujeito conhece a resposta então ele acerta o item. No entanto, se $w_{ij} = 0$, y_{ij} pode ser 1 ou 0. Segundo o modelo 3PN a probabilidade do sujeito acertar o item dado que não conhece a resposta é

$$P(y_{ij} = 1 \mid w_{ij} = 0) = \delta_j \quad (2.8)$$

Denotando $\lambda_{ij} = (\alpha_j, \beta_j, \delta_j, \theta_i)$, note que

$$\begin{aligned} P(y_{ij} = 1 \mid \lambda_{ij}) &= P(y_{ij} = 1 \mid \lambda_{ij}, w_{ij} = 0)P(w_{ij} = 0) + P(y_{ij} = 1 \mid \lambda_{ij}, w_{ij} = 1)P(w_{ij} = 1) \\ &= \delta_j P(Z_{ij} > 0) + 1P(Z_{ij} \leq 0) = \delta_j(1 - \Phi(\eta_{ij})) + \Phi(\eta_{ij}), \end{aligned}$$

a qual coincide com a probabilidade dada em (2.2) que caracteriza o modelo probito de três parâmetros.

Vamos agora obter as distribuições condicionais completas, as quais caracterizam o algoritmo DAAG. Usando (2.6), (2.7) e (2.8), obtemos a primeira distribuição a *posteriori* condicional.

$$\begin{aligned}
P(w_{ij} = 1 \mid y_{ij} = 1, \lambda_{ij}) &\propto P(w_{ij} = 1, y_{ij} = 1 \mid \lambda_{ij}) \\
&\propto P(y_{ij} = 1 \mid w_{ij} = 1, \lambda_{ij})P(w_{ij} = 1)
\end{aligned}$$

Das observações feitas anteriormente temos que $P(y_{ij} = 1 \mid w_{ij} = 1, \lambda_{ij}) = 1$ e $P(w_{ij} = 1) = \Phi(\alpha_j \theta_i - \beta_j)$. Assim,

$$P(w_{ij} = 1 \mid y_{ij} = 1, \lambda_{ij}) \propto \Phi(\alpha_j \theta_i - \beta_j)$$

Por outro lado,

$$\begin{aligned}
P(w_{ij} = 0 \mid y_{ij} = 1, \lambda_{ij}) &\propto P(w_{ij} = 0, y_{ij} = 1 \mid \lambda_{ij}) \\
&\propto P(y_{ij} = 1 \mid w_{ij} = 0, \lambda_{ij})P(w_{ij} = 0)
\end{aligned}$$

De (2.7) e (2.8), temos:

$$P(w_{ij} = 0 \mid y_{ij} = 1, \lambda_{ij}) \propto \delta_j(1 - \Phi(\alpha_j \theta_i - \beta_j))$$

Isto é, a distribuição *a posteriori* dos w_{ij} condicional a $y_{ij} = 1$ é uma Bernoulli com probabilidade de sucesso

$$p_{ij} = \frac{\Phi(\eta_{ij})}{(\Phi(\alpha_j \theta_i - \beta_j) + \delta_j(1 - \Phi(\eta_{ij})))} \quad (2.9)$$

Por outro lado, quando $Y_{ij} = 0$ temos que

$$P(w_{ij} = 0 \mid y_{ij} = 0, \eta_{ij}, \delta_j) = 1. \quad (2.10)$$

Note que por construção: $z_{ij} > 0$, se $w_{ij} = 1$ e $z_{ij} < 0$, se $w_{ij} = 0$. Portanto,

$$P(z_{ij} \mid w_{ij}, \lambda_{ij}) \propto \Phi(z_{ij} - \eta_{ij})[I_{(z_{ij} > 0)}I_{(w_{ij}=1)} + I_{(z_{ij} \leq 0)}I_{(w_{ij}=0)}] \quad (2.11)$$

Com as distribuições condicionais derivadas nesta seção é possível a obtenção de uma amostragem usando o DAAG. Isto pode ser feito no programa WinBUGS, porém não de forma direta, requerendo uma correta especificação para as variáveis indicadoras como em Bazan (2005). O mais apropriado seria a construção de um algoritmo próprio usando, por exemplo, o aplicativo R. No entanto, por facilidade, no Capítulo 3 vamos utilizar o WinBugs sem o esquema DAAG.

2.5 Construindo a amostra MCCM

Quando é usado um algoritmo do tipo MCCM, usualmente valores das iterações iniciais da cadeia são desconsiderados por não serem ainda uma amostra da distribuição estacionária da cadeia (distribuição de interesse). Além disso, quando é observada a presença de autocorrelação entre os valores simulados, sugere-se gerar uma grande quantidade de valores e utilizar apenas um subconjunto desses valores. Define-se um valor s , e a amostra final é obtida saltando-se s valores a partir de um ponto inicial. O termo em inglês para s é *thin*, neste trabalho será traduzido por **salto**. O uso dessa estratégia é recomendado por Gamerman e Lopes (2006), para obter uma amostra aproximadamente independente. No entanto, valores altos de s indicam que muitos valores gerados que não serão aproveitados, resultando em um gasto maior de tempo computacional.

Por exemplo, na geração de uma cadeia com 50000 iterações, descartando-se as primeiras 500 iterações, temos para $s = 1, 5$ e 10 os tamanhos de amostra de 49500,

9900 e 4950, respectivamente. Observe que não é desprezível a diferença nos tamanhos de amostra geradas para um mesmo tempo de execução do programa.

Bazán (2005) notou que as amostras geradas pelo algoritmo MCCM para o modelo 2PN são altamente correlacionadas e, portanto, sugere altos valores de salto para compor a amostra final. Essa estratégia acarreta um grande descarte de valores simulados. Nos capítulos que seguem, faremos um estudo de sensibilidade das estimativas *a posteriori* dos parâmetros ao comprimento do salto na construção da amostra final.

Capítulo 3

Aplicações

Nesta seção apresentamos duas aplicações do modelo de três parâmetros com respostas binárias. O primeiro conjunto de dados analisado foi retirado da tese de doutorado de Bazán (2005) e o segundo da dissertação de mestrado de Rios (2007). Para os dois conjuntos de dados estudamos primeiro o efeito do salto nas estimativas *a posteriori* dos parâmetros dos itens. Em seguida faremos uma análise de sensibilidade à escolha das distribuições *a priori* e, finalmente, interpretamos as estimativas obtidas para os parâmetros dos itens.

Por facilidade computacional, consideramos $\beta_{*j} = \beta_j \alpha_j$, assim $\eta_{ij} = \alpha_j \theta_i - \beta_{*j}$.

Ambos os conjuntos de dados são obtidos de provas com itens com respostas de múltipla-escolha. Este fato nos leva a propor o ajuste do modelo que inclui um terceiro parâmetro dos itens: a probabilidade de acerto ao acaso.

3.1 Aplicação I: prova de matemática para o ensino básico

Utilizamos nesta aplicação dados obtidos a partir de uma prova de matemática aplicada a 131 alunos de sexta série de escolas peruanas. A prova é composta de 14 itens de múltipla-escolha, cada um com 4 alternativas. Esse conjunto de dados foi analisado anteriormente por Bazán (2005), e o modelo 2PN foi o ajustado.

A seguir apresentamos uma descrição dos procedimentos adotados para análise e comentários sobre os valores encontrados durante o processo.

3.1.1 Um estudo sobre a influência do salto na construção da amostra MCCM

Em posse dos dados dos 131 estudantes do Peru, utilizando o modelo probito normal partimos para a execução do programa.

O programa WinBUGS 14 define para cada conjunto de dados um número mínimo de valores a serem descartados inicialmente, denominado *burn in*. Para este conjunto de dados a sugestão foi 4000.

O próximo passo foi definir um valor de salto, o qual não trouxesse prejuízo, tanto do ponto de vista do tempo computacional, quanto na confiabilidade das estimativas resultantes. Para isto fixamos dois tamanhos de amostra, $m = 5000$ e 10000 . Executamos o programa variando os valores dos saltos, $s = 1, 2, 5, 10$ e 50 , e anotamos os resultados para cada valor de s .

Para a execução mais longa, quando $s = 50$, precisamos de 504000 iterações. Isso

demorou cerca de 3 horas e 20 minutos, enquanto para $m = 5000$ e $s = 1$, execução mais curta, onde foram necessárias 9000 iterações, o tempo gasto foi 3 minutos e 50 segundos.

As estimativas dos parâmetros (médias *a posteriori*) foram organizadas nas tabelas A.1 e A.2 para $m = 5000$, e tabelas A.3 e A.4 para $m = 10000$. Para cada uma delas consideramos os cinco valores de s . Subtraímos da coluna $s = 50$ as colunas referentes aos outros valores de s , essas diferenças são apresentadas nas quatro últimas colunas das referidas tabelas (ver apêndice).

Notamos que a partir de $s = 2$ os valores dos módulos das diferenças são menores do que 0,08 para os dois tamanhos de amostras. Indicando que as estimativas dos parâmetros são bastante próximas, o que nos leva a questionar a necessidade do uso de um salto maior do que 2. Utilizamos também o MC erro, uma medida de qualidade da amostra simulada (quanto menor o MC erro mais confiantes estamos que a amostra é originária da distribuição de interesse), e fizemos uma comparação entre os valores de MC erro e desvio padrão ($MC\ erro/dp$), apresentados nas tabelas A.5 e A.5 . É fácil notar que quando se aumenta o valor do salto, diminui o valor dessa razão. Além disso, para um mesmo valor de s , a razão diminui com o aumento do tamanho da amostra. A razão $MC\ erro/dp$ foi superior a 5 % em apenas um caso quando o tamanho da amostra fixado foi $m = 10000$ e $s = 2$. Para $m = 5000$ e $s = 2$ encontramos uma quantidade maior de valores superiores a 5 %, obtendo inclusive valores na ordem de 8,5%. Optamos então por trabalhar com a amostra maior, $m = 10000$.

O tempo gasto para a execução com $s = 2$ e tamanho de amostra $m = 10000$ foi aproximadamente 13 minutos e 30 segundos, mais de 3 horas a menos do que o tempo gasto na execução mais longa.

3.1.2 Análise de sensibilidade às distribuições *a priori*.

Fixados $s = 2$ e $m = 10000$, passamos à análise de sensibilidade das estimativas *a posteriori* à escolha de diferentes distribuições *a priori*. Uma análise similar foi feita por Bazan, Bolfarine e Leandro (2007) para o modelo 2PN.

As distribuições *a priori* escolhidas para os parâmetros foram:

- $P_1 : \alpha \sim N(1, 0.5)I(0,)$, $\beta \sim N(0, 2)$, $\delta \sim \text{Beta}(1, 1)$
- $P_2 : \alpha \sim N(1, 0.5)I(0,)$; $\beta \sim N(0, 2)$; $\delta \sim \text{Beta}(0.5, 0.5)$
- $P_3 : \alpha \sim N(1, 0.5)I(0,)$; $\beta \sim N(0, 2)$; $\delta \sim \text{Beta}(1, 3)$
- $P_4 : \alpha \sim N(1, 0.5)I(0,)$; $\beta \sim N(0, 2)$; $\delta \sim \text{Beta}(10, 30)$
- $P_5 : \alpha \sim N(0.1, 0.5)I(0,)$; $\beta \sim N(0, 2)$; $\delta \sim \text{Beta}(0.5, 0.5)$
- $P_6 : \alpha \sim N(2, 0.5)I(0,)$; $\beta \sim N(0, 2)$; $\delta \sim \text{Beta}(0.5, 0.5)$
- $P_7 : \alpha \sim \text{LogNormal}(0, 1)$; $\beta \sim N(0, 2)$; $\delta \sim \text{Beta}(0.5, 0.5)$
- $P_8 : \alpha \sim \text{LogNormal}(1, 0.5)$; $\beta \sim N(0, 2)$; $\delta \sim \text{Beta}(0.5, 0.5)$

A notação $N(a, b)I(0,)$ indica a distribuição normal truncada para valores positivos com média **a** e variância **b**.

Em um primeiro momento fixamos as distribuições *a priori* para os parâmetros α e β , usando as propostas feitas por Sahu (2002) e Patz and Junker (1999): $\alpha \sim N(1, 0.5)I(0,)$ e $\beta \sim N(0, 2)$, e variamos a distribuição *a priori* do parâmetro δ .

As duas primeiras propostas de distribuições *a priori*, P_1 e P_2 , são escolhas objetivas (não informativas). A primeira é a distribuição uniforme no intervalo unitário e a segunda é a distribuição dada pela regra de Jeffreys para o modelo Bernoulli.

As escolhas seguintes, P_3 e P_4 , são subjetivas e baseadas na seguinte idéia: haver 4 alternativas para cada item, sendo que apenas uma delas é correta. Isso nos leva a pensar que a chance de acerto ao acaso é 0,25. Usamos esse número como valor esperado das distribuições Betas propostas. A diferença entre elas está na variabilidade da distribuição. P_3 tem variância maior do que P_4 , com valores 0.0375 e 0.004, respectivamente.

Um segundo conjunto de distribuições *a priori* foi proposto com a intenção de analisar a influência da distribuição *a priori* do parâmetro α . Em P_5 e P_6 modificamos o centro da distribuição *a priori* de α para verificar mudanças expressivas *a posteriori*. Em P_7 e P_8 utilizamos a distribuição *a priori* logNormal para α . A distribuição *a priori* para os parâmetros β foi mantida fixa para as oito propostas.

Após as 8 execuções planilhamos as estatísticas obtidas no programa. Consideramos inicialmente as estimativas de Média, Mediana e Desvio Padrão, (ver tabelas A.7, A.8, A.9, A.10, A.11 e A.12).

A partir dos resultados descritos nas tabelas A.7 à A.10 podemos observar que a distância entre médias e medianas das estimativas são consideravelmente baixas para a maioria dos parâmetros.

Notamos que, não só das estimativas *a posteriori* da média e da mediana, como também do desvio padrão variam bastante conforme a escolha da distribuição *a priori* para todos os parâmetros, com destaque para os parâmetros δ .

Os valores das estimativas das medidas de tendência central, para todos os itens tendem a ser menores sob as distribuições *a priori* subjetivas: P_3 e P_4 , do que sob as não informativas: P_1 e P_2 .

Vemos em P_3 e P_4 , com mais evidência no segundo caso, que os valores de todas

as estimativas *a posteriori* para os δ_j 's são muito próximos de 0,25, o que retrata a grande influência dessa distribuição *a priori*. Enquanto o uso de P_1 e P_2 possibilitou obter uma variabilidade maior de valores para as estimativas.

Observando os resultados para os percentis de 2.5% e 97.5% para os parâmetros δ_j 's, notamos que os intervalos são relativamente grandes, exceto no caso P_4 , em que colocamos uma priori muito informativa e com baixa variabilidade. Dentre as demais propostas, a que possui as maiores distâncias entre os referidos percentis é P_7 , onde o comprimento médio dos intervalos para os 14 parâmetros é 0,75.

Os valores dos percentis, bem como o comprimento do intervalo de credibilidade de 95% considerando P_4 e P_7 , são apresentados na tabela A.13.

As tabelas A.14, A.15 e A.16 apresentam os itens ordenados pela mediana *a posteriori* de δ , α e β^* , respectivamente, usando as diferentes distribuições *a priori*. Notamos que o item 12 aparece como o com menor probabilidade de acerto casual para todas as propostas. Mas isso não significa que os demais itens estejam igualmente ordenados nos 8 casos. Os itens 14, 2, 9, 10 e 5 aparecem com maior probabilidade de acerto ao acaso para a maioria das distribuições *a priori*, com excessão das distribuições *a priori* subjetivas.

O item 14 é o item com maior probabilidade de acerto casual para 6 das 8 propostas de distribuições *a priori*, no entanto, aparece como o 7º e 12º, se olharmos as ordenações de P_3 e P_4 , respectivamente.

Com relação aos parâmetros de dificuldade e discriminação, percebemos bastante similaridade nas ordenações para as 8 propostas.

Agrupamos as propostas de distribuições *a priori* utilizando as ordenações pelas estimativas *a posteriori* da mediana para os parâmetros δ_j , feitas anteriormente com

o objetivo de verificar a similaridade na ordenação dos itens entre as propostas de distribuições *a priori* consideradas.

Para nos auxiliar no cálculo da similaridade utilizamos o Coeficiente de Concordância de Kendall (CCK), ver Siegel (1975) .

Primeiro calculamos o CCK para as 8 propostas de distribuições *a priori* conjuntamente. Em seguida construímos alguns subgrupos afim de verificar um possível aumento da similaridade entre distribuições *a priori* com alguma característica em comum. O primeiro subgrupo foi composto pelas trincas cujas distribuições *a priori* para α e β são fixas, $(P_1, P_2, P_3 \text{ e } P_4)$. O segundo composto pelas trincas cujas alterações ocorrem apenas nas distribuições *a priori* de α , $(P_5, P_6, P_7 \text{ e } P_8)$. O terceiro composto pelas duas distribuições *a priori* pouco informativas para δ e cujas prioris para α e β são as mesmas, $(P_1 \text{ e } P_2)$. Finalmente o quarto subgrupo foi composto pelo que consideramos extremos com relação à informação das distribuições *a priori* para δ : P_2 e P_4 . Pois, P_2 é considerada a menos informativa e P_4 a mais informativa.

Quanto maior o valor do CCK, maior a similaridade entre os resultados da ordenação dos itens.

Na tabela 3.1 são apresentados os valores do CCK para os subgrupos de distribuições *a priori* anteriormente descritos, com relação aos parâmetros dos itens.

Confirmamos que com relação aos parâmetros de dificuldade e discriminação há bastante similaridade nas ordenações, sobretudo por utilizarmos distribuições *a priori* semelhantes para α e iguais para β em todos os casos. A concordância diminui quando comparamos δ , principalmente quando colocamos num mesmo subgrupo distribuições *a priori* informativas e não informativas, neste caso é notável a baixa concordância.

Tabela 3.1: Valores de CCK.

Grupos de prioris	δ	α	β^*
Todas	0,734	0,830	0,819
P_1, P_2, P_3, P_4	0,653	0,861	0,854
P_5, P_6, P_7, P_8	0,953	0,887	0,887
P_1, P_2	0,997	0,993	0,997
P_2, P_4	0,514	0,815	0,837

Pelos dois últimos subgrupos da tabela 3.1 percebemos claramente a similaridade entre a ordenação dos parâmetros δ_j para P_1 e P_2 e a diferença que há entre esta e P_4 , uma distribuição que carrega muita informação *a priori*.

A partir das observações que tivemos durante o estudo apresentado nesta seção, concluímos que este modelo é bastante sensível a escolha de distribuições *a priori*. Escolhemos P_2 como distribuição *a priori* para a análise das estimativas *a posteriori* no modelo de três parâmetros para o conjunto de dados das provas aplicadas no Perú.

3.1.3 Análise das estimativas

De acordo com as estimativas *a posteriori* obtidas a partir da distribuição *a priori* P_2 , $m = 10000$ e $s = 2$, extraímos as informações descritas nas tabelas 3.3 e 3.2. Os itens apresentam-se ordenados pela mediana. Na tabela 3.3 também são apresentados os intervalos de 95% de credibilidade para as estimativas dos parâmetros da probabilidade de acerto ao acaso. Estes intervalos chamam atenção por seus comprimentos.

O item que mais discrimina entre os indivíduos é o item 11 e o que menos discrimina é o item 3. O item 6 é o com maior índice de dificuldade, enquanto o 11

é o estimado menos difícil dentre os 14 itens que compõe a prova. Esses itens são apresentados a seguir.

- **3:** Un metro de tela cuesta SI. 60,50. ¿Cuánto se pagará por 0,5 metros?
- **6:** "El precio de una blusa es S/. 30. Si Ana lo compró con el 20 % de descuento, ¿cuánto pagó por la blusa?"
- **11:** "Luisa, Dora y María compraron tela. Luisa compró medio metro, Dora compró 75 cm. y María compró 50 cm. ¿Quiénes compraron la misma cantidad de tela?"

Com relação à probabilidade de acerto ao acaso, as estimativas de δ para os itens 3 e 14 são as que possuem os valores extremos, respectivamente menor e maior probabilidade dentre os demais itens da prova. Os itens 3 e 12 apresentam estimativas para δ inferiores a 0,2. Os itens 12 e 14 são apresentados a seguir.

- **12:** "Un tanque recibe 4,5 litros de agua por minuto. ¿Cuántos litros de agua tendrá el tanque en una hora y media?"
- **14:** "Observa las siguientes figuras. ¿Cuál es la suma de todos los lados del rombo?"

Todos os itens desta prova podem ser vistos no anexo.

Pela tabela 3.2 vemos que o item 11 é o com maior índice de discriminação e o item com o menor índice de dificuldade.

Tabela 3.2: Estimativas *a posteriori* para os parâmetros α e β , ordenadas pela mediana .

Item	Discriminação (α)		Item	Dificuldade (β)	
	Mediana	Média		Mediana	Média
11	1,536	1,578	6	1,564	1,681
4	1,168	1,231	12	1,471	1,576
7	1,148	1,214	9	0,958	0,7842
8	1,124	1,191	2	0,447	0,357
5	0,866	0,961	3	0,439	0,517
1	0,827	0,928	13	-0,120	-0,121
14	0,823	0,916	14	-0,129	-0,111
12	0,774	0,906	10	-0,220	-0,214
13	0,718	0,838	1	-0,277	-0,331
6	0,696	0,851	5	-0,321	-0,334
10	0,690	0,796	8	-1,026	-0,975
2	0,671	0,809	7	-1,186	-1,098
9	0,632	0,797	4	-1,296	-1,222
3	0,602	0,724	11	-1,521	-1,484

O item 3, apresenta a menor probabilidade de acerto ao acaso, menor índice de discriminação e é o item com o quinto maior índice de dificuldade. Os itens com as mais altas estimativas para δ_j : 14 e 2, apresentam-se em níveis intermediários de dificuldade e discriminação. Os parâmetros discriminação, dificuldade e probabilidade de acerto ao acaso não estão relacionados entre si.

Com relação exclusivamente às estimativas para os parâmetros δ_j 's, em geral os valores das medianas *a posteriori* são superiores a 0.25. Apenas os itens 11, 6, 12 e 3

Tabela 3.3: Estimativas *a posteriori* para os parâmetros de acerto ao acaso (δ), ordenadas pela mediana.

Item	Mediana	Média	Intervalo 95 %
14	0,881	0,820	0,207-0,962
2	0,750	0,706	0,233-0,880
10	0,692	0,622	0,025-0,870
5	0,687	0,621	0,029-0,874
9	0,682	0,640	0,206-0,820
13	0,625	0,566	0,030-0,832
7	0,568	0,516	0,005-0,905
4	0,560	0,509	0,005-0,908
1	0,512	0,467	0,001-0,769
8	0,442	0,417	0,002- 0,817
11	0,268	0,316	0,095-0,827
6	0,213	0,195	0,092-0,398
12	0,189	0,180	0,087-0,374
3	0,167	0,188	0,045-0,492

apresentam estimativas da mediana mais próximas do limite inferior do intervalo de 95%. Note que destes 4 itens, apenas o 11 possui média e mediana superiores a 0.25. As estimativas das probabilidades de acerto ao acaso se concentram em valores altos. Considerando que os itens são compostos de 4 alternativas, sendo apenas uma correta, para que haja uma probabilidade de acerto ao acaso maior que 0,25, é necessário que algumas alternativas forneçam informações que induzam o respondente a preferi-las em detrimento das outras. Mesmo que o aluno conheça muito pouco do assunto, a forma como foi escrita a alternativa, bem como o seu conhecimento extra-curricular,

em conjunto com a sua habilidade são fatores que podem levar à exclusão de algumas das alternativas, contradizendo a idéia de que a escolha de cada uma delas é feita de modo uniforme (igual chance para todas).

3.2 Aplicação II: simulado em um cursinho pré-vestibular

Consideramos agora dados obtidos a partir de um simulado aplicado a 277 alunos do cursinho pré-vestibular Ângulo, em São Paulo. A prova foi composta de 10 itens de múltipla-escolha, cada um com 5 alternativas (Rios 2007). Neste caso, $n = 277$ e $k = 10$.

3.2.1 Um estudo sobre a influência do salto na construção da amostra MCCM

Os procedimentos para a definição do salto e do *burn in* foram os mesmos descritos na subseção anterior. Neste caso o tempo de execução do programa para gerar 504000 iterações foi aproximadamente 5 horas e 26 minutos , já que se trata de uma amostra maior. Para a execução mais rápida, com tamanho de amostra $m = 5000$ e $s = 1$, foram gastos apenas 5 minutos. Veja os resultados das médias para as amostras de tamanho $m = 5000$ e $m = 10000$, e dos diferentes valores de s nas tabelas B.1 e B.2.

Observamos que os resultados para todos os saltos considerados estão muito próximos. Até mesmo para a amostra de tamanho $m = 5000$, $s = 1$, o módulo da diferença para $s = 50$ não passa de 0,05 para nenhum dos parâmetros dos itens, .

As diferenças entre $s = 50$ e os demais valores de salto são bem inferiores às encontradas para os dados da aplicação I, onde encontramos diferenças superiores a 0,08 entre $s = 1$ e $s = 50$. Podemos dizer que a convergência neste caso se dá com saltos e tamanhos de amostra menores.

No entanto, ao considerarmos as relações entre os desvios padrões e os valores de $Mc\ erro$, estes ainda superam os 5% daqueles para uma grande parte dos parâmetros em estudo. Os percentuais para todos os parâmetros encontram-se na tabela B.3.

Com o objetivo de otimizarmos nossos resultados, tentamos aumentar a amostra para 15000 com o mesmo $s = 1$, mas não houve diminuição significativa na razão $Mc\ erro/dp$.

Ao considerarmos $s = 5$ e $m = 5000$, notamos que os percentuais $Mc\ erro/dp$ são relativamente baixos, todos inferiores a 5%. Para $s = 2$ e a amostra maior, $m = 10000$, apenas para um dos parâmetros (β_8), o $Mc\ erro$ chegou a 5% do desvio padrão. Apesar desta situação para o parâmetro β_8 , preferimos adotar estes valores para s e m . Embora, no segundo caso, o tamanho da amostra seja maior o número de iterações necessárias do programa é 24000, contra 29000 no primeiro caso, quando $s = 5$ e $m = 5000$. A economia no tempo de execução é aproximadamente de 4 minutos.

Para $s = 2$ e tamanho de amostra $m = 10000$, o tempo gasto para execução do programa foi aproximadamente 21 minutos. Esta opção nos trouxe resultados muito próximos daqueles obtidos com amostras e s 's maiores.

3.2.2 Análise de sensibilidade às distribuições *a priori*.

Definidos o tamanho de amostra e o salto a serem utilizados passamos à análise de sensibilidade das estimativas *a posteriori* levando em conta a escolha de diferentes distribuições *a priori*.

Consideramos para estes dados o mesmo conjunto de oito propostas de distribuições *a priori*, descritos em 3.1.2, para os parâmetros de discriminação, dificuldade e probabilidade de acerto ao acaso. Apenas foram feitos alguns ajustes em P_3 e P_4 , a saber:

- $P_3^* : \alpha \sim N(1, 0.5)I(0,); \beta \sim N(0, 2); \delta \sim Beta(1, 4)$
- $P_4^* : \alpha \sim N(1, 0.5)I(0,); \beta \sim N(0, 2); \delta \sim Beta(10, 40)$

Neste caso os itens têm 5 alternativas e não 4, como nas provas aplicadas no Peru. Portanto, usando o mesmo raciocínio anterior, a probabilidade de acerto ao acaso esperada é 0,20. Com a escolha da distribuição *a priori* P_4^* , desejamos apenas diminuir a variância (de 0,0266 para 0,0003) *a priori* para verificar possível efeito *a posteriori*.

Nas tabelas B.4, B.5 e B.6 são apresentadas as estimativas *a posteriori* obtidas para cada uma das propostas de distribuições *a priori*. Na tabela B.7, os percentis 2.5% e 97.5%, bem como o comprimento desses intervalos para as propostas com menores e maiores intervalos médios, respectivamente P_4 e P_1 .

Na tabela B.6 notamos que os valores das estimativas dos desvios padrões de grande parte dos parâmetros variam conforme as diferentes propostas de distribuições *a priori*. O mesmo acontece com relação à média e à mediana, cujas estimativas são apresentadas respectivamente nas tabelas B.4 e B.5.

As diferenças nas estimativas, tanto para o desvio padrão, quanto para média e mediana, são maiores quando comparamos os resultados das distribuições *a priori* pouco informativas e muito informativas.

Observando os valores dos quantis 2.5% e 97.5%, encontramos intervalos de confiabilidade mais estreitos para os valores dos δ_j 's, se compararmos aos intervalos encontrados para os mesmos parâmetros na aplicação I.

Neste caso devemos levar em conta que todos os valores das estimativas de média e mediana dos δ_j s são baixas. Isto leva-nos a concluir que as probabilidades de acerto ao acaso para todos os itens são muito pequenas.

Considerando cada priori separadamente, fizemos uma ordenação decrescente das medianas das estimativas obtidas. Em seguida fizemos uma análise para verificar se, mesmo que os valores numéricos não estivessem concordando sob as diferentes prioris, haveria uma preservação na ordem dos itens. Veja as tabelas B.8, B.9 e B.10 que mostram os itens em ordem decrescente, conforme os valores das estimativas das medianas dos parâmetros para cada proposta de distribuição *a priori*. Embora haja certa uniformidade entre os dados, não há concordância completa entre as distribuições *a priori* com relação à ordenação dos itens.

Calculamos o CCK para as oito propostas. Em seguida construímos alguns sub-grupos, seguindo os mesmos critérios descritos na aplicação I, a fim de verificar a similaridade entre distribuições *a priori*.

Diante dos resultados descritos na tabela 3.4, notamos que com relação aos três parâmetros há bastante similaridade nas ordenações dos itens. Todos os valores de CCK superam 0,90. Para α e β^* , os coeficientes são altos, sobretudo por utilizarmos distribuições *a priori* semelhantes para os parâmetros de discriminação e iguais para

Tabela 3.4: Tabela de CCK.

<i>priori</i>	δ	α	β^*
Todas	0,908	0,965	0,973
P_1, P_2, P_3^*, P_4^*	0,938	0,977	0,971
P_5, P_6, P_7, P_8	0,932	0,957	0,980
P_1, P_2	0,939	0,939	0,988
P_2, P_4^*	0,921	0,921	0,976

os de dificuldade. Com relação aos parâmetros de probabilidade de acerto ao acaso, os CCK's são pouco inferiores aos dos demais parâmetros, e também indicam influência mínima da distribuição *a priori* na ordenação das estimativas, mesmo quando comparamos distribuições *a priori* com características bem distintas, como P_2 e P_4^* . O menor valor de CCK para δ é encontrado para o grupo que contém todas as distribuições: 0,908.

Pelos dois últimos subgrupos da tabela 3.1 percebemos claramente a similaridade entre a ordenação dos parâmetros δ_j para P_1 e P_2 e a diferença que há entre esta e P_4^* , uma distribuição que carrega muita informação *a priori*. O maior grau de similaridade para δ encontra-se no subgrupo onde as distribuições *a priori* propostas para os parâmetros probabilidade de acerto ao acaso são pouco informativas.

Com base nesses estudos, escolhemos P_2 para a análise dos parâmetros dos itens. Embora para esse conjunto de dados, a influência da distribuição *a priori* seja menor que anteriormente e, portanto, essa escolha é menos relevante.

3.2.3 Análise das estimativas

Considerando as estimativas *a posteriori* para os parâmetros de item com a proposta de distribuição *a priori* P_2 , cujas estimativas de média e mediana apresentamos nas tabelas 3.5 e 3.6. Os itens estão ordenados pela mediana.

Tabela 3.5: Estimativas *a posteriori* para os parâmetros α e β , ordenadas pela mediana.

Item	Discriminação (α)		Item	Dificuldade (β)	
	Mediana	Média		Mediana	Média
10	1,626	1,671	2	1,391	1,433
6	1,496	1,539	5	0,534	0,558
1	1,442	1,485	9	0,291	0,293
7	1,324	1,376	6	0,105	0,110
9	0,991	1,046	4	-0,015	0,015
3	0,854	0,879	8	-0,014	0,009
5	0,739	0,764	1	-0,0398	-0,032
4	0,735	0,758	7	-0,255	-0,246
8	0,716	0,776	3	-0,266	-0,243
2	0,567	0,582	10	-0,431	-0,421

Pela tabela 3.5 vemos que os itens 10 e 2 são os extremos das classificações das estimativas para os parâmetros α e β . O item 10 é o item que mais discrimina entre os respondentes e o considerado menos difícil, enquanto o 2 ocupa posições extremo-opostas.

Com relação à probabilidade de acerto ao acaso, o item 9 é o que apresenta a maior estimativa e o 2, o que possui menores chances de ser acertado casualmente.

Tabela 3.6: Resumo das estimativas *a posteriori* para os parâmetros de acerto ao acaso (δ), ordenadas pela mediana.

item	Mediana	Média	Intervalo 95 %
9	0,130	0,131	0,089-0,293
8	0,123	0,147	0,0238-0,403
7	0,073	0,093	0,030-0,278
3	0,068	0,093	0,000-0,299
10	0,062	0,083	0,000-0,266
4	0,056	0,084	0,0139-0,289
1	0,046	0,061	0,0144-0,191
6	0,045	0,054	0,0017-0,164
5	0,035	0,054	0,009-0,195
2	0,017	0,031	0,004-0,131

Notamos para todos os itens valores de δ muito baixos para as estimativas de média e mediana. Apenas dois itens (8 e 9) utrapassam 0,1. Lembrando que os itens são de múltipla-escolha, com 5 alternativas cada, a probabilidade de acerto ao acaso esperada seria 0,2. Pela tabela 3.6 notamos que os intervalos de confiabilidade para as estimativas dos δ_j 's são consideravelmente estreitos.

Vimos que o item 2, com alto índice de dificuldade, apresenta baixa probabilidade de acerto ao acaso e discrimina pouco entre os indivíduos. Já o item 10, discrimina bem, é considerado fácil e ocupa a posição mediana com relação à probabilidade de acerto ao acaso. Apenas os itens 8 e 9 apresentam probabilidades de acerto ao acaso superiores a 0,1.

3.3 Similaridades e diferenças entre as aplicações I e II

Notamos que para os dados de escolas peruanas as estimativas *a posteriori* do parâmetro de acerto ao acaso são bastante sensíveis à escolha das distribuições *a priori*. No entanto, para os dados do cursinho paulista esta sensibilidade é menos evidente. Isso se deve ao tamanho da amostra, maior neste caso.

Para os dados de escolas peruanas, com itens com 4 alternativas, encontramos estimativas superiores ao esperado para δ supondo alternativas equiprováveis. Já para os dados do cursinho, as estimativas estão muito abaixo do esperado para itens com 5 alternativas. Portanto, não há evidências de que as alternativas de um mesmo item tenham a mesma probabilidade de serem escolhidas pelo respondente.

Os intervalos de 95 % de credibilidade para os parâmetros δ são menores para o conjunto de dados do cursinho paulista, cuja amostra é composta por 277 indivíduos. Já para os dados do Perú, cuja amostra é composta de 131 respondentes, os intervalos são bastante grandes, chegando a ultrapassar 0,9 de comprimento, ou seja, quase não há precisão. Neste caso não se pode confiar muito nas estimativas pontuais.

Os dados da aplicação II possuem mais de duas vezes o número de indivíduos respondentes do que na aplicação I. Para obtermos estimativas mais precisas para o modelo de três parâmetros seria aconselhável trabalhar com um tamanho de amostra maior para a aplicação I.

Capítulo 4

Modelos com respostas ordinais

Além dos itens do tipo múltipla-escolha, os quais os indivíduos podem obter dois resultados possíveis (acerto ou erro), existem itens que permitem conceitos intermediários na correção. São os itens chamados abertos ou dissertativos. As respostas dadas a esses itens podem ser classificadas como certas, erradas ou parcialmente certas. Esta última categoria pode ainda ser dividida em subcategorias, as quais são ordenadas conforme o grau de acerto. Esse modelo de respostas ordinais será discutido nesta seção.

Seja $y_{i,j}$ a resposta dada pelo i -ésimo indivíduo no j -ésimo item. Note que y_{ij} deixa de ser binário para atingir uma dimensão maior, de acordo com os graus de acerto, ou categorias definidos para a correção. Aqui tratamos especificamente do caso com três categorias e utilizamos a ligação probito para construção da CCI.

As três categorias podem ser assim identificadas:

1. Resposta completamente errada.
2. Resposta parcialmente certa.

3. Resposta completamente certa.

Seguindo a proposta de Albert and Chib (1993) de variáveis latentes para dados ordinais, construímos aqui o modelo de dados aumentados na TRI.

Apresentamos o modelo para o caso de três categorias, mas ele é facilmente estendido para o caso mais geral.

O modelo de dados aumentados com três categorias ordinais é caracterizado por:

$$y_{i,j} = \left\{ \begin{array}{l} 1, \text{ se } Z_{ij} \leq 0, \\ 2, \text{ se } 0 < Z_{ij} \leq \gamma_j, \\ 3, \text{ se } Z_{ij} > \gamma_j, \end{array} \right\} \quad (4.1)$$

Onde as quantidades latentes $Z_{ij} \mid \alpha, \beta, \theta \sim N(\eta_{ij}, 1)$ para $i = 1, \dots, n$, $j = 1, \dots, k$ com

$$\eta_{ij} = \alpha_j \theta_i - \beta * _j .$$

Um novo conjunto de parâmetros foi introduzido no modelo com a seguinte restrição $\gamma_j \geq 0$. Como no modelo 2PN, $\alpha_j > 0$.

Do modelo definido em (4.1) obtemos

$p_3 = P(Z_{ij} > \gamma_j) = \Phi(\eta_{ij} - \gamma_j)$, a probabilidade de acerto do item;

$p_2 = P(0 < Z_{ij} \leq \gamma_j) = \Phi(\eta_{ij}) - \Phi(\eta_{ij} - \gamma_j)$, a probabilidade de acerto parcial,

$p_1 = 1 - \Phi(\eta_{ij})$ a probabilidade de erro do item.

As curvas características dos itens para o modelo ordinal são caracterizadas pelas probabilidades p_1 , p_2 e p_3 . Notamos que p_1 e p_3 são funções monótonas decrescentes ou crescentes. No entanto, para p_2 não observamos esse mesmo tipo de comportamento.

Figura 4.1: Curvas Características para o modelo ordinal.

A Figura (4.1) ilustra as três curvas, denominado gráfico da Curva Característica do Item Ordinal.

Outro tipo de gráfico que pode ser analisado no modelo de resposta ordinal é o dos Limites das Curvas Características (ver Baker e Kim, 2004). Para a construção desse gráfico são consideradas as seguintes probabilidades acumuladas

$$p_4^+ = 0$$

$$p_3^+ = p_3$$

$$p_2^+ = p_2 + p_3$$

$$p_1^+ = 1.$$

Note que $p_2^+ = \Phi(\eta_{ij})$ e representa a probabilidade de não errar completamente o item ou, equivalentemente, de obter algum tipo de acerto (total + parcial).

Figura 4.2: Limites das Curva Característica do Item para o modelo ordinal.

Na Figura (4.2) apresentamos essas curvas. Notamos que as probabilidades associadas a cada uma das categorias de resposta são representadas pelas distâncias entre as curvas.

Para facilitar a interpretação dos parâmetros, vamos considerar a seguinte reparametrização

$$\beta_{j3} = (\beta * _j + \gamma_j) / \alpha_j \quad \text{e} \quad \beta_{j2} = \beta * _j / \alpha_j.$$

Com isso, p_3^+ e p_2^+ podem ser reescritas como

$$p_3^+ = \Phi(\alpha_j(\theta_i - \beta_{j3})) \quad \text{e} \quad p_2^+ = \Phi(\alpha_j(\theta_i - \beta_{j2})).$$

Na Figura (4.2), a curva pontilhada (p_3^+) representa a probabilidade de acerto total ao item e a curva contínua (p_2^+) a probabilidade de algum tipo de acerto (parcial mais completo). As duas curvas são igualmente crescentes a medida em que é maior a habilidade do indivíduo, e são equidistantes em todos os pontos, pois o índice de discriminação α_j se mantém. A diferença está no parâmetro de posição, que aumenta quando passamos da probabilidade de acerto completo do item à de algum tipo de

acerto, pois $\beta_{j3} \geq \beta_{j2}$.

Interpretamos β_{j3} e β_{j2} , respectivamente, como os parâmetros de dificuldade de acerto completo do item e de algum tipo de acerto (não errar completamente o item).

A função de verossimilhança aumentada é proporcional à

$$\left\{ \prod_{i=1}^n \prod_{j=1}^k I_{[y_{ij}=1] \cap [z_{ij} < 0]} + I_{[y_{ij}=2] \cap [0 < z_{ij} < \gamma_j]} + I_{[y_{ij}=3] \cap [z_{ij} > \gamma_j]} \right\} \phi(z_{ij} - \eta_{ij})$$

onde Φ e ϕ são as funções acumulada e densidade da normal padrão, respectivamente.

4.1 Estimação dos parâmetros

Para obtenção das estimativas dos parâmetros de interesse na aplicação III, cujos dados resultantes de itens com respostas ordinais são descritos na próxima seção, utilizamos a função **MCMCordfactanal** do programa "R", versão 2.4.1.

MCMCordfactanal utiliza métodos MCMC para análise de dados ordinais. Essa função gera uma amostra da distribuição *a posteriori* dos parâmetros do modelo.

É assumida independência *a priori* para os parâmetros. Distribuições *a priori* normais com média zero e precisão 0.5 são assumidas para α e β . Para θ é considerada a distribuição normal padrão. Para o ponto de corte γ_j é usada uma distribuição *a priori* uniforme imprópria nos reais positivos.

MCMCordfactanal simula uma amostra da distribuição *a posteriori* usando o método de Metropolis-Hasting em Gibbs. O algoritmo implementado é baseado no trabalho de Cowles (1996). Essa amostra retorna como um objeto MCMC, que pode ser analisado subsequente com outras funções. Além disso, a função retorna as

estatísticas descritivas da amostra simulada para os parâmetros α_j , β_{*j} , θ_i , γ_j .

4.2 Aplicação III: Provas do curso de Noções de Estatística

Para uma análise prática do modelo com respostas ordinais tomamos um conjunto de dados proveniente de provas da disciplina MAE0116: Noções de Estatística. O curso foi ministrado no primeiro semestre do ano de 2007 para alunos do curso de biologia da Universidade de São Paulo e a prova, que segue em anexo, foi respondida por 49 alunos.

Para esta análise descartamos a primeira questão, pois é do tipo verdadeiro ou falso e, portanto, as respostas são binárias e não ordinais. A questão 2 apresenta dois itens, $2 - a$ e $2 - b$. As questões 3 e 4 apresentam três itens cada uma: $3 - a$, $3 - b$, $3 - c$, $4 - a$, $4 - b$ e $4 - c$. Finalmente a questão 5, possui quatro itens $5 - a$, $5 - b$, $5 - c$ e $5 - d$. Para este estudo consideramos estes 12 itens dissertativos numerados de 1 a 12, partindo-se do $2 - a$ e seguindo consecutivamente até o $5 - d$.

O objetivo principal foi classificar esses 12 itens com relação as dificuldades (β_{j3} e β_{j2}) e a discriminação (α_j).

Consideramos uma amostra de tamanho $MCCM = 10000$, descartamos as 3000 iterações iniciais e usamos salto igual a 10. Os parâmetros para as distribuições *a priori* para os α_j 's e β_j 's foram normais com média 0 e precisão 0,5.

A partir dos resultados retornados pela função, observamos que não há diferenças relevantes entre as estimativas de média e mediana para os parâmetros. Notamos também que as estimativas do desvio padrão são relativamente altas, o que leva a

intervalos de confiabilidade não tão precisos.

Na tabela C.1, apresentamos as estimativas para as médias dos 3 parâmetros de interesse: α , β_{j3} e β_{j2} . Os itens apresentam-se classificados em ordem decrescente segundo as médias *a posteriori*.

Note que dos três parâmetros considerados, um é interpretado como índice de discriminação e dois como índices de dificuldades dos itens. Como já visto, β_{j3} e β_{j2} são os parâmetros de dificuldades de acerto completo e de não errar completamente o item, respectivamente. A íntegra dos enunciados dos itens dessa prova de MAE116 encontram-se em anexo.

Os itens que mais discriminam entre os respondentes são respectivamente o 11, o 5 e o 12, apresentados a seguir

- **(5)** "(3-c) Se em uma amostra aleatória de 300 pacientes, 200 precisaram freqüentar as seções de fisioterapia por mais de um ano, determine um intervalo de confiança para a verdadeira proporção p a um nível de 99
- **(11)** "(5-c) Determine a região crítica para o teste considerando o nível de significância igual a 0,10."
- **(12)** "(5-d) Calcule a estatística do teste e conclua."

Os itens 4 e 2 são os itens com mais baixo poder de discriminação.

- **(2)** "(2-b) Qual a diferença entre nível de significância e nível descritivo em um teste de hipótese estatístico?"
- **(4)** "(3-b) Sabendo-se que a proporção p não é inferior a 40 %, é possível reduzir o tamanho da amostra? Em caso afirmativo, qual o novo valor de n ?"

Com relação aos índices de dificuldade, encontramos valores extremos para os itens 1 e 4, sendo o item 1 o mais difícil e o item 4 o mais fácil.

- **(1)** ”(2-a) (a) Formule o problema como um teste de hipóteses estatístico (defina hipóteses, erros, estatística do teste e região crítica).”

Estes itens permanecem na mesma posição quando ordenados pelos dois índices de dificuldade considerados. Contudo isso não ocorre para os demais itens. O item 2, por exemplo, é o segundo menos difícil de se acertar parcialmente, no entanto é o segundo com o maior valor estimado para β_3 .

O maior acréscimo no índice de dificuldade é dado ao item 2. A diferença entre as médias de β_{23} e β_{22} é 3,592 para esse item. Enquanto o menor acréscimo é dado ao item 4, com diferença entre as médias dos dois parâmetros de dificuldade estimada em 0,277. Este item permanece o mais fácil nas classificações para os dois índices de dificuldade.

Nenhum dos 49 respondentes acertou parcialmente o item 10.

- (10)** ”(5-b) (b) Supondo que não existe associação entre as variáveis, calcule o número esperado de indivíduos do sexo feminino que tem preferência por sacarina.”

Assim o índice de dificuldade estimado nesse caso é o mesmo que para se acertar completamente o item ($\beta_{102} = \beta_{103}$).

Enquanto os itens 2 e 4 possuem estimativas baixas para os parâmetros α e β_2 , o item 1 possui altos índices de dificuldades (parcial e completa) e está em décimo lugar com relação ao índice de discriminação, e o item 11 é o que mais discrimina, o décimo classificado como dificuldade de acerto parcial e o décimo primeiro em acerto completo. Não encontramos relações entre os índices de discriminação e dificuldades.

Capítulo 5

Considerações finais

Neste trabalho descrevemos a Teoria da resposta ao item aplicada a modelos de três parâmetros com respostas binárias e ordinais. Os principais objetivos foram estimar os parâmetros dos itens: dificuldade, discriminação e probabilidade de acerto ao acaso para o modelo com respostas binárias, e discriminação e dificuldades de acertos parcial e completo, no caso do modelo com respostas ordinais.

Embora haja uma certa dificuldade na estimação dos parâmetros dos modelos aqui descritos, tanto do ponto de vista teórico quanto computacional, nesse trabalho mostramos que o uso de dois *software livres*, o WinBUGS e o R, nos permite realizar de forma adequada essa tarefa. Sem ferramentas específicas não seria possível a implementação dos modelos.

Para o modelo com respostas binárias, o tamanho do salto na geração da cadeia não exerceu grande influência nos resultados das estimativas, contudo, influência a razão Mc erro/dp. Essa é uma medida de qualidade da amostra simulada, quanto menor o Mc erro mais confiantes estamos que a amostra é originária da distribuição de interesse. Nas duas aplicações utilizadas para o estudo desse modelo houve sensi-

bilidade a escolha das distribuições *a priori*, sobretudo com relação aos parâmetros probabilidade de acerto ao acaso. No entanto, notamos que um tamanho de amostra maior ($n=277$) esta sensibilidade é bem menos evidente.

Ainda assim encontramos na aplicação I, estimativas para δ_j acima de $1/a$, onde a é o número de alternativas dos itens considerados. Possivelmente as alternativas, devido à forma em que são escritas, não sejam equiprováveis de serem escolhidas pelos respondentes.

Não encontramos em nenhum dos modelos relações entre os parâmetros de dificuldade e discriminação.

No modelo com respostas ordinais, o terceiro parâmetro representa o acréscimo no índice de dificuldade quando se aumenta o grau de acerto. A análise do conjunto de dados retirados das provas de MAE116 foi facilitado devido ao uso da função *MCMCordfactanal*. Essa função utilizada para a análise no modelo com respostas ordinais identificou o número de graus em que foram classificadas as respostas dos itens, e isso determinou o número de parâmetros γ para cada item. Sendo assim, a mesma função pode ser utilizada no modelo cuja classificação das respostas é binária. Nesse caso o único parâmetro de corte $\gamma_j = 0$, e o DAAG é facilmente identificado.

Assim sendo, o método descrito para o modelo da TRI, com respostas ordinais pode ser estendido para análise de itens cuja correção classifica as respostas em c categorias, onde $c \geq 2$. A mesma função *MCMCordfactanal* pode ser implementada. Vale conferir as CCI, bem como o comportamento da ordenação dos índices de dificuldade para cada grau de acerto, e se há tendência para os sucessivos parâmetros γ_{jc} , a medida em que k aumenta.

Além da sugestão de implementação da função *MCMCordfactanal* para itens com c categorias de respostas, alguns outros pontos levantados neste trabalho oferecem

possibilidades para futuras pesquisas, tais como:

- uma discussão mais profunda sobre a interpretação do parâmetro δ (probabilidade de acerto ao acaso em itens de múltipla-escolha);
- escolha de distribuições *a priori* assimétricas para o parâmetro das habilidades. Os resultados *a posteriori* estimados para essa variável latente nos modelos com o terceiro parâmetro;
- implementação no programa WinBUGS para os modelos de três parâmetros com respostas binárias e ordinais, utilizando o DAAG.
- escolha das distribuições *a priori* para o modelo de respostas ordinais.

Referências

- [1] Albert, J. (1992). Bayesian estimation of normal ogive item response curves using Gibbs sampling. *Journal of Educational Statistics*, **17**, 251-269.
- [2] Albert, J. and Chib, S. (1993). Bayesian analysis fo binary and polychotomous response data. *JASA*, **88**, 669-679.
- [3] Albert, J.H. and Ghosh, M. (2000). Item response modeling. In Generalized Linear Models: A Bayesian Perspective(D. Dey, S. Ghosh & Mallick, eds.), Marcel-Dekker, New York, 173-193.
- [4] Andrade, D.F., Tavares, H.R. e Valle, R.C. (2000). *Teoria da Resposta ao Item: conceitos e aplicações*. São Paulo: Associação Brasileira de Estatística.
- [5] Bazan, J.L (2005). *Uma família de modelos de resposta ao item normal assimétrica*. Tese de doutorado. IME-USP.
- [6] Bazan, J.L., Bolfarine, H. and Leandro, R.A. (2007). Sensitivity analysis of prior specification for the probit-normal IRT model: An empirical study. Submetido.
- [7] Beguin, A.A. and Glas, C.A. (2001). MCMC estimation and some model-fit analysis of multidimensional IRT models. *Psychometrika*, **66**, 541-562 .
- [8] Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Infering an Examinees Ability. In F. M. Lord & M. R. Novick. Statistical Theories of Mental Test Scores. Reading, MA : Addison-Wesley.
- [9] Bock, R. D. and Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, 64,153-168.
- [10] Cowles M. K (1996). Accelerating Monte Carlo Markov Chain Convergence

for Cumulative-link Generalized Linear Models. *Statistics and Computing*. 6: 101-110.

[11] Gamerman, D. and Lopes, H.L. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Second Edition. Taylor & Francis, New York.

[12] Gelman A. and Carlin J. B. and Stern H. S. and Rubin D. B (2004). *Bayesian Data Analysis*. Chapman & Hall/CRC. Boca Raton.

[13] Lord, F. M. (1952). *A theory of test scores*. New York: Psychometric Society.

[14] Mislevy, R.J. (1986). Bayes modal estimation in item response models. *Psychometrika*, 51, 177-195.

[15] Paulino C. D. and Turkman M. A. A. and Murteira B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.

[16] Rios, C. (2007). *Uma aplicação da TRI assimétrica*. Dissertação de mestrado. IME-USP.

[17] Siegel, S. (1975). *Estatística não-paramétrica*. McGraw-Hill.

Apêndice A

Tabelas: Aplicação I

Tabela A.1: Médias *a posteriori* para α e β^* com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 5000. Aplicação I.

Parâmetro	1	2	5	10	50	50-1	50-2	50-5	50-10
α_1	1,080	1,117	1,107	1,112	1,124	0,044	0,007	0,017	0,012
α_2	0,911	0,903	0,876	0,885	0,868	-0,044	-0,036	-0,009	-0,018
α_3	0,899	0,860	0,859	0,889	0,916	0,017	0,056	0,057	0,027
α_4	1,248	1,258	1,230	1,233	1,244	-0,004	-0,014	0,014	0,011
α_5	1,082	1,070	1,088	1,108	1,095	0,013	0,025	0,007	-0,013
α_6	0,720	0,718	0,722	0,725	0,722	0,002	0,004	0,000	-0,003
α_7	1,178	1,171	1,182	1,170	1,166	-0,012	-0,005	-0,016	-0,004
α_8	1,339	1,324	1,303	1,293	1,279	-0,060	-0,045	-0,024	-0,014
α_9	0,689	0,737	0,720	0,715	0,719	0,029	-0,018	-0,001	0,004
α_{10}	0,857	0,878	0,898	0,900	0,924	0,067	0,046	0,026	0,024
α_{11}	1,448	1,468	1,484	1,482	1,475	0,027	0,007	-0,009	-0,007
α_{12}	0,799	0,781	0,769	0,769	0,761	-0,039	-0,020	-0,008	-0,008
α_{13}	0,980	0,977	0,975	0,968	0,967	-0,013	-0,009	-0,008	-0,001
α_{14}	0,902	0,928	0,935	0,927	0,905	0,004	-0,022	-0,030	-0,022
β^*_{11}	-0,131	-0,092	-0,086	-0,061	-0,043	0,088	0,048	0,043	0,018
β^*_{12}	0,580	0,571	0,445	0,530	0,507	-0,073	-0,064	0,062	-0,023
β^*_{13}	0,967	0,846	0,811	0,841	0,865	-0,102	0,019	0,054	0,024
β^*_{14}	-1,398	-1,424	-1,367	-1,338	-1,360	0,038	0,064	0,007	-0,022
β^*_{15}	-0,078	-0,117	-0,195	-0,173	-0,180	-0,103	-0,063	0,014	-0,007
β^*_{16}	1,562	1,594	1,595	1,599	1,583	0,021	-0,011	-0,012	-0,016
β^*_{17}	-1,122	-1,119	-1,093	-1,092	-1,102	0,020	0,017	-0,009	-0,010
β^*_{18}	-1,055	-1,023	-0,954	-0,955	-0,971	0,084	0,052	-0,018	-0,016
β^*_{19}	1,083	1,164	1,108	1,088	1,076	-0,007	-0,088	-0,032	-0,012
β^*_{20}	-0,098	-0,049	-0,008	0,005	0,030	0,128	0,079	0,039	0,025
β^*_{21}	-1,763	-1,872	-1,921	-1,910	-1,906	-0,143	-0,034	0,015	0,004
β^*_{22}	1,613	1,533	1,526	1,504	1,524	-0,089	-0,009	-0,002	0,020
β^*_{23}	0,054	0,120	0,074	0,126	0,140	0,087	0,021	0,066	0,014
β^*_{24}	-0,549	-0,477	-0,457	-0,457	-0,473	0,076	0,004	-0,017	-0,017

Tabela A.2: Médias *a posteriori* para δ com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 5000. Aplicação I.

Parâmetro	1	2	5	10	50	50-1	50-2	50-5	50-10
δ_1	0,474	0,494	0,492	0,501	0,500	0,026	0,007	0,008	-0,001
δ_2	0,714	0,712	0,677	0,694	0,687	-0,027	-0,026	0,009	-0,007
δ_3	0,319	0,295	0,290	0,295	0,298	-0,021	0,003	0,008	0,003
δ_4	0,532	0,526	0,534	0,539	0,538	0,005	0,012	0,003	-0,002
δ_5	0,642	0,627	0,614	0,624	0,622	-0,020	-0,005	0,008	-0,002
δ_6	0,258	0,262	0,261	0,262	0,259	0,000	-0,004	-0,002	-0,003
δ_7	0,556	0,556	0,560	0,561	0,555	0,000	0,000	-0,005	-0,005
δ_8	0,455	0,462	0,477	0,475	0,468	0,013	0,006	-0,009	-0,007
δ_9	0,653	0,672	0,662	0,657	0,656	0,002	-0,016	-0,006	-0,001
δ_{10}	0,585	0,603	0,616	0,620	0,625	0,039	0,022	0,009	0,005
δ_{11}	0,462	0,434	0,424	0,432	0,425	-0,037	-0,010	0,001	-0,008
δ_{12}	0,244	0,234	0,233	0,230	0,235	-0,009	0,001	0,002	0,005
δ_{13}	0,571	0,577	0,566	0,577	0,582	0,011	0,005	0,016	0,005
δ_{14}	0,707	0,722	0,725	0,719	0,715	0,008	-0,006	-0,010	-0,003

Tabela A.3: Médias *a posteriori* para α e β^* com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 10000. Aplicação I.

Parâmetro	1	2	5	10	50	50-1	50-2	50-5	50-10
α_1	1,119	1,111	1,109	1,109	1,122	0,003	0,011	0,013	0,013
α_2	0,905	0,870	0,890	0,888	0,871	-0,034	0,001	-0,019	-0,017
α_3	0,858	0,864	0,889	0,905	0,908	0,050	0,044	0,019	0,003
α_4	1,260	1,230	1,231	1,235	1,240	-0,020	0,010	0,009	0,005
α_5	1,066	1,077	1,104	1,102	1,099	0,033	0,022	-0,005	-0,003
α_6	0,714	0,731	0,721	0,716	0,715	0,001	-0,016	-0,006	-0,001
α_7	1,171	1,180	1,174	1,168	1,166	-0,005	-0,014	-0,008	-0,002
α_8	1,323	1,303	1,292	1,294	1,284	-0,039	-0,019	-0,008	-0,010
α_9	0,738	0,718	0,714	0,708	0,720	-0,019	0,002	0,006	0,011
α_{10}	0,884	0,884	0,900	0,911	0,914	0,031	0,030	0,014	0,003
α_{11}	1,468	1,480	1,482	1,480	1,480	0,012	0,000	-0,002	0,000
α_{12}	0,777	0,763	0,765	0,771	0,757	-0,019	-0,006	-0,007	-0,014
α_{13}	0,981	0,977	0,971	0,979	0,969	-0,012	-0,009	-0,002	-0,011
α_{14}	0,932	0,937	0,925	0,931	0,910	-0,023	-0,027	-0,015	-0,021
β^*_{11}	-0,091	-0,077	-0,060	-0,054	-0,045	0,046	0,032	0,015	0,009
β^*_{12}	0,571	0,436	0,527	0,531	0,490	-0,081	0,054	-0,037	-0,041
β^*_{13}	0,848	0,820	0,843	0,853	0,853	0,005	0,033	0,009	0,000
β^*_{14}	-1,424	-1,375	-1,332	-1,348	-1,366	0,058	0,009	-0,034	-0,018
β^*_{15}	-0,117	-0,174	-0,173	-0,159	-0,177	-0,060	-0,003	-0,004	-0,018
β^*_{16}	1,593	1,616	1,595	1,577	1,572	-0,021	-0,044	-0,023	-0,005
β^*_{17}	-1,120	-1,118	-1,095	-1,096	-1,098	0,022	0,020	-0,003	-0,002
β^*_{18}	-1,022	-0,937	-0,951	-0,962	-0,986	0,036	-0,050	-0,036	-0,024
β^*_{19}	1,166	1,108	1,086	1,075	1,078	-0,088	-0,030	-0,008	0,003
β^*_{20}	-0,048	-0,024	0,005	0,033	0,008	0,055	0,032	0,002	-0,025
β^*_{21}	-1,871	-1,905	-1,908	-1,913	-1,913	-0,042	-0,008	-0,005	0,000
β^*_{22}	1,534	1,498	1,503	1,535	1,522	-0,012	0,024	0,019	-0,013
β^*_{23}	0,120	0,070	0,126	0,137	0,139	0,019	0,070	0,013	0,003
β^*_{24}	-0,478	-0,461	-0,457	-0,471	-0,450	0,027	0,010	0,007	0,021

Tabela A.4: Médias *a posteriori* para δ com diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 10000. Aplicação I.

Parâmetro	1	2	5	10	50	50-1	50-2	50-5	50-10
δ_1	0,493	0,497	0,500	0,500	0,502	0,008	0,004	0,002	0,002
δ_2	0,711	0,678	0,692	0,694	0,685	-0,026	0,007	-0,007	-0,009
δ_3	0,297	0,292	0,296	0,297	0,296	0,000	0,004	0,001	-0,001
δ_4	0,526	0,530	0,542	0,538	0,536	0,010	0,006	-0,006	-0,002
δ_5	0,628	0,615	0,622	0,625	0,622	-0,007	0,006	0,000	-0,004
δ_6	0,262	0,263	0,262	0,260	0,257	-0,005	-0,006	-0,004	-0,003
δ_7	0,556	0,552	0,561	0,560	0,558	0,000	0,007	-0,003	-0,001
δ_8	0,463	0,482	0,475	0,473	0,465	0,002	-0,017	-0,010	-0,008
δ_9	0,671	0,666	0,658	0,656	0,656	-0,015	-0,010	-0,002	0,001
δ_{10}	0,604	0,610	0,619	0,625	0,619	0,015	0,009	0,000	-0,006
δ_{11}	0,435	0,428	0,428	0,430	0,426	-0,008	-0,002	-0,002	-0,004
δ_{12}	0,234	0,230	0,231	0,235	0,234	0,000	0,004	0,004	-0,001
δ_{13}	0,577	0,567	0,577	0,582	0,580	0,003	0,013	0,003	-0,002
δ_{14}	0,721	0,725	0,721	0,716	0,720	-0,001	-0,006	-0,001	0,004

Tabela A.5: Relações entre SD e MC Error para δ com os dados da aplicação I.

Amostra	5 000					10 000				
s	1	2	5	10	50	1	2	5	10	50
δ_1	7,9%	5,9%	4,1%	2,6%	1,6%	5,7%	4,1%	2,9%	2,2%	1,1%
δ_2	7,8%	6,0%	4,5%	3,7%	1,8%	6,0%	5,0%	3,5%	2,7%	1,3%
δ_3	7,8%	6,1%	4,0%	3,2%	1,5%	6,0%	4,7%	3,2%	2,2%	1,3%
δ_4	6,4%	4,5%	3,2%	2,4%	1,4%	4,5%	3,4%	2,3%	1,7%	1,0%
δ_5	7,5%	6,2%	4,1%	2,6%	1,6%	5,9%	4,6%	2,7%	1,9%	1,1%
δ_6	7,0%	5,7%	3,5%	2,4%	1,5%	5,3%	4,0%	2,4%	1,9%	1,0%
δ_7	6,9%	5,2%	3,3%	2,6%	1,3%	5,2%	3,8%	2,6%	2,0%	1,0%
δ_8	5,9%	4,7%	3,7%	2,7%	1,6%	5,0%	3,7%	2,5%	1,8%	1,1%
δ_9	10,3%	8,6%	5,6%	4,5%	2,0%	7,6%	6,1%	4,7%	3,2%	1,6%
δ_{10}	8,2%	6,6%	4,3%	3,1%	1,6%	5,6%	4,7%	2,9%	2,2%	1,0%
δ_{11}	5,9%	4,5%	3,1%	2,3%	1,4%	4,6%	3,2%	2,1%	1,5%	1,0%
δ_{12}	7,5%	6,4%	4,3%	3,1%	1,5%	5,9%	4,6%	2,9%	2,0%	1,2%
δ_{13}	8,0%	7,2%	4,9%	3,3%	1,7%	6,6%	5,0%	3,4%	2,6%	1,1%
δ_{14}	7,1%	5,6%	4,1%	3,0%	1,6%	5,5%	4,1%	2,8%	2,2%	1,1%

Tabela A.6: Relações entre SD e MC Error para α e β^* dos dados da aplicação I.

Amostra	5 000					10 000				
s	1	2	5	10	50	1	2	5	10	50
α_1	5,1%	3,8%	2,6%	2,0%	1,4%	3,6%	2,8%	1,9%	1,4%	1,0%
α_2	4,6%	3,8%	2,7%	2,1%	1,2%	3,3%	2,8%	1,9%	1,4%	1,0%
α_3	5,8%	4,6%	3,1%	2,4%	1,3%	4,3%	3,5%	2,3%	1,8%	1,0%
α_4	3,9%	2,9%	1,9%	1,6%	1,5%	2,6%	1,8%	1,4%	1,4%	1,0%
α_5	4,0%	3,2%	2,2%	1,7%	1,4%	3,4%	2,3%	1,5%	1,4%	1,0%
α_6	5,5%	3,9%	2,7%	1,8%	1,2%	3,9%	2,7%	1,9%	1,3%	0,9%
α_7	4,3%	2,8%	1,8%	1,7%	1,5%	2,6%	1,9%	1,4%	1,2%	0,9%
α_8	4,7%	3,4%	2,1%	1,9%	1,4%	3,3%	2,3%	1,6%	1,3%	1,0%
α_9	5,7%	4,1%	2,9%	2,4%	1,6%	3,8%	3,1%	2,3%	1,8%	1,1%
α_{10}	4,5%	3,4%	2,2%	1,9%	1,5%	3,3%	2,3%	1,6%	1,4%	1,0%
α_{11}	3,2%	3,2%	2,0%	1,5%	1,4%	3,0%	2,4%	1,6%	1,2%	1,0%
α_{12}	5,3%	4,2%	2,9%	2,5%	1,7%	4,0%	3,3%	2,1%	1,7%	1,1%
α_{13}	5,0%	3,5%	2,9%	1,8%	1,5%	3,5%	2,9%	1,8%	1,5%	1,0%
α_{14}	3,8%	3,0%	1,9%	1,6%	1,3%	3,0%	2,1%	1,4%	1,2%	0,9%
β^*_{11}	8,2%	5,9%	4,3%	2,8%	1,6%	6,0%	4,3%	3,0%	2,2%	1,2%
β^*_{12}	7,8%	6,0%	4,1%	3,3%	1,7%	5,7%	4,7%	3,3%	2,3%	1,2%
β^*_{13}	7,9%	5,9%	4,1%	3,3%	1,5%	5,8%	4,7%	3,2%	2,3%	1,2%
β^*_{14}	7,7%	5,4%	3,9%	2,4%	1,3%	5,7%	4,0%	2,5%	1,8%	1,1%
β^*_{15}	8,1%	6,5%	4,2%	2,7%	1,5%	6,2%	4,8%	2,6%	1,9%	1,1%
β^*_{16}	7,0%	5,6%	3,6%	2,3%	1,4%	5,2%	3,9%	2,5%	1,8%	1,1%
β^*_{17}	8,3%	6,2%	4,2%	3,1%	1,5%	6,2%	4,5%	3,3%	2,3%	1,1%
β^*_{18}	7,7%	5,8%	4,5%	3,0%	1,7%	6,1%	4,7%	3,0%	2,1%	1,2%
β^*_{19}	8,6%	6,6%	4,4%	3,3%	1,8%	6,1%	4,9%	3,7%	2,6%	1,4%
β^*_{20}	8,3%	6,6%	4,1%	2,9%	1,5%	5,8%	4,5%	2,8%	2,1%	1,0%
β^*_{21}	7,3%	5,7%	4,0%	2,8%	1,3%	5,9%	4,1%	2,9%	2,0%	1,0%
β^*_{22}	7,7%	6,4%	4,3%	3,2%	1,4%	6,0%	4,5%	3,0%	2,1%	1,3%
β^*_{23}	8,2%	7,2%	5,0%	3,3%	1,7%	6,7%	5,1%	3,5%	2,4%	1,1%
β^*_{24}	7,9%	6,0%	4,4%	2,9%	1,6%	5,9%	4,5%	3,1%	2,1%	1,2%
Máximo	10,3%	8,6%	5,6%	4,5%	2,0%	7,6%	6,1%	4,7%	3,2%	1,6%

Tabela A.7: Médias *a posteriori* para δ considerando diferentes distribuições *a priori*.
Aplicação I.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
δ_1	0,497	0,454	0,375	0,271	0,414	0,543	0,451	0,552
δ_2	0,678	0,682	0,445	0,261	0,608	0,764	0,670	0,758
δ_3	0,292	0,260	0,243	0,262	0,200	0,363	0,218	0,349
δ_4	0,530	0,504	0,311	0,257	0,451	0,622	0,545	0,643
δ_5	0,615	0,585	0,414	0,265	0,540	0,705	0,588	0,716
δ_6	0,263	0,231	0,230	0,259	0,181	0,288	0,197	0,288
δ_7	0,552	0,503	0,336	0,259	0,447	0,645	0,536	0,710
δ_8	0,482	0,430	0,304	0,259	0,332	0,544	0,436	0,564
δ_9	0,666	0,656	0,448	0,260	0,535	0,733	0,615	0,734
δ_{10}	0,610	0,584	0,396	0,261	0,512	0,720	0,552	0,726
δ_{11}	0,428	0,359	0,252	0,253	0,330	0,499	0,431	0,561
δ_{12}	0,230	0,208	0,209	0,248	0,167	0,263	0,189	0,276
δ_{13}	0,567	0,558	0,391	0,263	0,467	0,658	0,528	0,654
δ_{14}	0,725	0,717	0,373	0,257	0,694	0,833	0,739	0,826

Tabela A.8: Médias *a posteriori* para α e β^* considerando diferentes distribuições *a priori*. Aplicação I.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
α_1	1,111	1,064	0,959	0,809	0,684	1,843	1,630	2,931
α_2	0,870	0,876	0,596	0,341	0,483	1,699	0,946	2,311
α_3	0,864	0,878	0,847	0,882	0,641	1,440	0,689	1,323
α_4	1,230	1,197	1,106	1,058	0,840	1,945	1,688	2,870
α_5	1,077	1,071	0,848	0,646	0,676	1,869	1,447	2,936
α_6	0,731	0,682	0,674	0,776	0,441	1,201	0,541	1,110
α_7	1,180	1,164	1,001	0,908	0,822	1,845	1,488	2,775
α_8	1,303	1,251	1,141	1,122	0,894	1,974	1,810	3,136
α_9	0,718	0,741	0,470	0,226	0,381	1,464	0,619	1,433
α_{10}	0,884	0,884	0,697	0,533	0,585	1,638	1,013	2,255
α_{11}	1,480	1,465	1,470	1,474	1,095	2,023	1,888	2,862
α_{12}	0,763	0,733	0,731	0,836	0,522	1,202	0,625	1,132
α_{13}	0,977	0,938	0,744	0,565	0,556	1,808	1,160	2,714
α_{14}	0,937	0,919	0,612	0,523	0,523	1,889	1,173	3,003
β^*_{*1}	-0,077	-0,123	-0,375	-0,652	-0,179	0,057	-0,144	0,126
β^*_{*2}	0,436	0,520	-0,344	-0,871	0,182	1,084	0,490	1,011
β^*_{*3}	0,820	0,757	0,626	0,621	0,455	1,445	0,533	1,318
β^*_{*4}	-1,375	-1,324	-1,744	-1,808	-1,249	-1,358	-1,282	-1,288
β^*_{*5}	-0,174	-0,220	-0,726	-0,001	-0,307	0,206	-0,181	0,311
β^*_{*6}	1,616	1,420	1,392	1,527	1,043	2,139	1,182	2,059
β^*_{*7}	-1,118	-1,164	-1,508	-1,611	-1,158	-0,918	-1,042	-0,661
β^*_{*8}	-0,937	-0,983	-1,241	-1,354	-1,082	-0,948	-1,047	-0,967
β^*_{*9}	1,108	1,135	0,173	-0,546	0,473	1,928	0,930	1,909
β^*_{*10}	-0,024	-0,034	-0,640	-0,966	-0,267	0,602	-0,113	0,648
β^*_{*11}	-1,905	-1,990	-2,204	-2,213	-1,763	-2,001	-1,905	-1,861
β^*_{*12}	1,498	1,399	1,355	1,572	1,070	2,066	1,252	2,090
β^*_{*13}	0,070	0,135	-0,419	-0,765	-0,121	0,539	0,052	0,553
β^*_{*14}	-0,461	-0,316	-1,382	-1,611	-0,347	-0,040	-0,202	-0,104

Tabela A.9: Medianas *a posteriori* para δ considerando diferentes distribuições *a priori*. Aplicação I.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
δ_1	0,533	0,505	0,389	0,268	0,452	0,586	0,507	0,592
δ_2	0,746	0,764	0,460	0,257	0,710	0,790	0,751	0,781
δ_3	0,304	0,279	0,247	0,259	0,185	0,393	0,212	0,383
δ_4	0,570	0,563	0,285	0,253	0,470	0,707	0,626	0,745
δ_5	0,675	0,672	0,426	0,261	0,619	0,751	0,678	0,755
δ_6	0,280	0,253	0,245	0,260	0,184	0,301	0,209	0,303
δ_7	0,595	0,564	0,310	0,255	0,466	0,736	0,625	0,791
δ_8	0,510	0,465	0,285	0,255	0,303	0,606	0,481	0,638
δ_9	0,719	0,725	0,482	0,257	0,643	0,748	0,711	0,750
δ_{10}	0,675	0,677	0,391	0,255	0,587	0,770	0,645	0,772
δ_{11}	0,434	0,344	0,219	0,249	0,285	0,559	0,445	0,650
δ_{12}	0,243	0,225	0,218	0,248	0,164	0,281	0,203	0,289
δ_{13}	0,624	0,635	0,393	0,259	0,528	0,700	0,618	0,695
δ_{14}	0,817	0,843	0,334	0,253	0,825	0,885	0,860	0,881

Tabela A.10: Medianas *a posteriori* para α e β^* considerando diferentes distribuições *a priori*. Aplicação I.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
α_1	1,035	0,980	0,851	0,732	0,613	1,833	0,921	2,150
α_2	0,778	0,778	0,445	0,307	0,387	1,695	0,598	1,707
α_3	0,777	0,773	0,753	0,796	0,575	1,408	0,562	1,146
α_4	1,180	1,135	1,047	1,016	0,789	1,924	1,170	2,165
α_5	1,003	0,988	0,730	0,598	0,606	1,880	0,871	2,207
α_6	0,656	0,583	0,583	0,713	0,374	1,169	0,429	1,003
α_7	1,120	1,111	0,936	0,866	0,773	1,820	1,010	2,118
α_8	1,247	1,182	1,077	1,069	0,841	1,956	1,250	2,407
α_9	0,609	0,638	0,311	0,191	0,275	1,449	0,447	1,250
α_{10}	0,776	0,777	0,594	0,493	0,518	1,623	0,623	1,641
α_{11}	1,442	1,427	1,439	1,445	1,061	1,988	1,523	2,270
α_{12}	0,687	0,643	0,652	0,776	0,456	1,164	0,499	1,037
α_{13}	0,882	0,839	0,625	0,523	0,481	1,789	0,675	2,012
α_{14}	0,837	0,817	0,522	0,473	0,436	1,894	0,673	2,263
β^*_{*1}	-0,190	-0,281	-0,493	-0,651	-0,353	-0,014	-0,312	0,016
β^*_{*2}	0,343	0,471	-0,604	-0,875	0,037	1,015	0,365	0,933
β^*_{*3}	0,669	0,570	0,491	0,534	0,300	1,418	0,364	1,219
β^*_{*4}	-1,456	-1,424	-1,737	-1,759	-1,408	-1,442	-1,382	-1,402
β^*_{*5}	-0,314	-0,349	-0,858	-1,031	-0,497	0,146	-0,304	0,229
β^*_{*6}	1,487	1,294	1,249	1,399	0,862	2,091	0,967	1,997
β^*_{*7}	-1,264	-1,343	-0,002	-0,002	-1,316	-1,063	-1,215	-0,788
β^*_{*8}	-1,014	-1,093	-1,268	-0,001	-1,169	-0,001	-1,138	-1,046
β^*_{*9}	1,072	1,142	-0,183	-0,553	0,308	1,905	0,908	1,896
β^*_{*10}	-0,220	-0,225	-0,806	-0,961	-0,483	0,521	-0,328	0,585
β^*_{*11}	-1,928	-2,016	-2,166	-2,162	-1,799	-2,062	-1,934	-1,956
β^*_{*12}	1,384	1,259	1,224	1,454	0,890	2,034	1,090	2,025
β^*_{*13}	-0,074	-0,025	-0,602	-0,767	-0,359	0,450	-0,109	0,477
β^*_{*14}	-0,635	-0,509	-1,483	-1,591	-0,555	-0,068	-0,385	-0,167

Tabela A.11: Desvio padrão *a posteriori* para δ com diferentes distribuições *a priori*.
Aplicação I.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
δ_1	0,188	0,230	0,198	0,072	0,248	0,178	0,227	0,171
δ_2	0,197	0,213	0,255	0,070	0,254	0,115	0,218	0,111
δ_3	0,137	0,157	0,132	0,067	0,152	0,128	0,155	0,137
δ_4	0,245	0,287	0,212	0,070	0,299	0,255	0,286	0,260
δ_5	0,209	0,249	0,234	0,071	0,267	0,165	0,251	0,155
δ_6	0,095	0,112	0,102	0,055	0,120	0,084	0,122	0,085
δ_7	0,246	0,293	0,220	0,069	0,293	0,254	0,299	0,222
δ_8	0,217	0,253	0,197	0,069	0,253	0,226	0,259	0,234
δ_9	0,158	0,194	0,248	0,069	0,257	0,087	0,228	0,088
δ_{10}	0,219	0,257	0,241	0,070	0,279	0,168	0,270	0,162
δ_{11}	0,236	0,266	0,182	0,067	0,263	0,273	0,287	0,277
δ_{12}	0,096	0,109	0,095	0,055	0,112	0,092	0,115	0,082
δ_{13}	0,205	0,239	0,227	0,071	0,270	0,157	0,252	0,160
δ_{14}	0,234	0,278	0,256	0,069	0,287	0,160	0,267	0,175

Tabela A.12: Desvio padrão *a posteriori* para α e β^* com diferentes distribuições *a priori*. Aplicação I.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
α_1	0,552	0,552	0,522	0,419	0,402	0,731	2,237	2,642
α_2	0,572	0,580	0,499	0,231	0,381	0,757	1,117	2,076
α_3	0,477	0,501	0,468	0,468	0,369	0,693	0,506	0,764
α_4	0,518	0,515	0,472	0,426	0,409	0,671	1,819	2,433
α_5	0,583	0,593	0,521	0,341	0,413	0,745	1,846	2,510
α_6	0,457	0,462	0,446	0,462	0,317	0,624	0,401	0,569
α_7	0,526	0,516	0,459	0,386	0,401	0,689	1,668	2,271
α_8	0,539	0,530	0,467	0,438	0,401	0,689	1,954	2,506
α_9	0,530	0,538	0,464	0,169	0,346	0,683	0,596	0,955
α_{10}	0,547	0,556	0,467	0,297	0,375	0,764	1,354	2,077
α_{11}	0,466	0,474	0,451	0,435	0,392	0,589	1,586	2,153
α_{12}	0,465	0,468	0,449	0,470	0,343	0,625	0,450	0,565
α_{13}	0,576	0,558	0,493	0,313	0,379	0,738	1,703	2,427
α_{14}	0,578	0,575	0,427	0,322	0,395	0,728	1,681	2,535
β^*_{*1}	0,687	0,778	0,541	0,228	0,739	0,812	0,818	0,953
β^*_{*2}	1,008	1,013	0,823	0,182	0,989	1,040	1,060	1,040
β^*_{*3}	0,705	0,751	0,593	0,411	0,556	0,898	0,632	0,932
β^*_{*4}	0,747	0,811	0,512	0,396	0,764	1,042	0,912	1,125
β^*_{*5}	0,861	0,887	0,559	0,220	0,839	0,988	0,916	1,082
β^*_{*6}	0,819	0,827	0,777	0,644	0,690	0,952	0,793	0,941
β^*_{*7}	0,823	0,877	0,471	0,333	0,770	1,188	0,980	1,239
β^*_{*8}	0,680	0,708	0,453	0,322	0,550	0,863	0,800	1,064
β^*_{*9}	1,053	1,090	0,987	0,170	1,052	1,010	1,150	1,038
β^*_{*10}	0,923	0,977	0,609	0,207	0,887	1,082	1,007	1,108
β^*_{*11}	0,674	0,697	0,538	0,486	0,597	0,886	0,819	1,035
β^*_{*12}	0,775	0,794	0,719	0,655	0,662	0,968	0,773	0,911
β^*_{*13}	0,856	0,948	0,634	0,196	0,905	0,997	0,927	1,040
β^*_{*14}	1,119	1,242	0,600	0,275	1,141	1,241	1,217	1,375

Tabela A.13: Percentis das estimativas dos parâmetros δ para P_4 e P_7 . Aplicação I.

Parâmetro	P_4			P_7		
	2.5%	97.5%	Tam (95%)	2.5%	97.5%	Tam (95%)
δ_1	0,142	0,426	0,284	0,005	0,771	0,766
δ_2	0,135	0,409	0,274	0,039	0,879	0,840
δ_3	0,140	0,399	0,259	0,120	0,506	0,386
δ_4	0,131	0,403	0,272	0,007	0,910	0,904
δ_5	0,140	0,416	0,276	0,010	0,872	0,862
δ_6	0,150	0,365	0,215	0,001	0,399	0,398
δ_7	0,137	0,404	0,267	0,002	0,913	0,912
δ_8	0,134	0,403	0,270	0,003	0,826	0,823
δ_9	0,136	0,405	0,269	0,014	0,826	0,812
δ_{10}	0,138	0,410	0,273	0,005	0,872	0,867
δ_{11}	0,133	0,396	0,263	0,002	0,876	0,874
δ_{12}	0,141	0,353	0,212	0,001	0,380	0,378
δ_{13}	0,138	0,414	0,275	0,006	0,831	0,825
δ_{14}	0,134	0,403	0,269	0,017	0,961	0,944

Tabela A.14: Itens ordenados pela mediana *a posteriori* de δ usando diferentes distribuições *a priori*. Aplicação I.

Proposta	>													<
P_2	14	2	9	10	5	13	7	4	1	8	11	3	6	12
P_1	14	2	9	5	10	13	7	4	1	8	11	3	6	12
P_3	9	2	5	13	10	1	14	7	4	8	3	6	11	12
P_4	1	5	6	3	13	2	9	7	8	10	4	14	11	12
P_5	14	2	9	5	10	13	4	7	1	8	11	3	6	12
P_6	14	2	10	5	9	7	4	13	8	1	11	3	6	12
P_7	14	2	9	5	10	4	7	13	1	8	11	3	6	12
P_8	14	7	2	10	5	9	4	13	11	8	1	3	6	12

Tabela A.15: Itens ordenados pela mediana *a posteriori* de α usando diferentes distribuições *a priori*. Aplicação I.

Proposta	>	<
P_2	11 8 4 7 5 1 13 14 2 10 3 12 9	6
P_1	11 8 4 7 1 5 13 14 2 3 10 12 6	9
P_3	11 8 4 7 1 3 5 12 13 10 6 14 2	9
P_4	11 8 4 7 3 12 1 6 5 13 10 14 2	9
P_5	11 8 4 7 1 5 3 10 13 12 14 2 6	9
P_6	11 8 4 14 5 1 7 13 2 10 9 3 6	12
P_7	11 8 4 7 1 5 13 14 10 2 3 12 9	6
P_8	8 11 14 5 4 1 7 13 2 10 9 3 12	6

Tabela A.16: Itens ordenados pela mediana *a posteriori* de β^* usando diferentes distribuições *a priori*. Aplicação I.

Proposta	>	<
P_2	6 12 9 3 2 13 10 1 5 14 8 7 4	11
P_1	6 2 9 3 2 13 1 10 5 14 8 7 4	11
P_3	6 12 3 7 9 1 13 2 10 5 8 14 4	11
P_4	12 6 3 8 7 9 1 13 2 10 5 14 4	11
P_5	9 6 2 14 12 13 10 3 5 1 7 4 8	11
P_6	12 6 3 9 10 2 13 7 1 5 14 4 8	11
P_7	6 9 12 3 2 10 14 13 5 1 7 4 8	11
P_8	6 12 3 9 10 2 13 7 1 4 5 14 11	8

Apêndice B

Tabelas: Aplicação II

Tabela B.1: Médias *a posteriori* para diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 5000. Aplicação II.

Parâmetro	1	2	5	10	50	50-1	50-2	50-5	50-10
δ_1	0,096	0,094	0,095	0,094	0,092	-0,004	-0,003	-0,003	-0,002
δ_2	0,064	0,062	0,062	0,060	0,060	-0,004	-0,002	-0,002	0,000
δ_3	0,142	0,147	0,148	0,148	0,145	0,002	-0,002	-0,004	-0,004
δ_4	0,148	0,143	0,146	0,144	0,142	-0,006	-0,002	-0,005	-0,002
δ_5	0,088	0,088	0,090	0,092	0,089	0,001	0,000	-0,002	-0,004
δ_6	0,077	0,079	0,077	0,078	0,078	0,000	-0,001	0,000	0,000
δ_7	0,132	0,130	0,132	0,131	0,133	0,001	0,002	0,001	0,002
δ_8	0,228	0,234	0,239	0,237	0,234	0,006	0,000	-0,005	-0,003
δ_9	0,175	0,171	0,166	0,165	0,167	-0,007	-0,004	0,001	0,002
δ_{10}	0,127	0,126	0,127	0,125	0,127	0,000	0,001	0,000	0,002
α_1	1,639	1,637	1,641	1,633	1,626	-0,013	-0,011	-0,015	-0,007
α_2	0,632	0,628	0,624	0,620	0,627	-0,005	-0,001	0,004	0,007
α_3	0,968	0,980	0,979	0,977	0,975	0,007	-0,005	-0,004	-0,002
α_4	0,874	0,866	0,869	0,866	0,864	-0,011	-0,002	-0,005	-0,002
α_5	0,838	0,842	0,849	0,857	0,843	0,005	0,001	-0,006	-0,014
α_6	1,688	1,699	1,680	1,673	1,687	-0,001	-0,012	0,007	0,014
α_7	1,502	1,495	1,508	1,499	1,508	0,006	0,013	0,000	0,009
α_8	0,945	0,967	0,978	0,974	0,968	0,023	0,001	-0,010	-0,006
α_9	1,208	1,200	1,187	1,178	1,185	-0,023	-0,015	-0,002	0,007
α_{10}	1,826	1,823	1,823	1,821	1,827	0,001	0,004	0,004	0,006
β_{*1}	0,071	0,071	0,072	0,059	0,049	-0,022	-0,022	-0,022	-0,010
β_{*2}	0,950	0,940	0,935	0,925	0,925	-0,025	-0,016	-0,010	-0,001
β_{*3}	-0,103	-0,092	-0,090	-0,094	-0,099	0,003	-0,007	-0,010	-0,005
β_{*4}	0,174	0,164	0,169	0,163	0,155	-0,019	-0,009	-0,014	-0,008
β_{*5}	0,523	0,528	0,540	0,546	0,526	0,003	-0,002	-0,014	-0,020
β_{*6}	0,265	0,273	0,266	0,262	0,258	-0,007	-0,015	-0,008	-0,004
β_{*7}	-0,234	-0,239	-0,240	-0,245	-0,239	-0,005	0,001	0,002	0,007
β_{*8}	0,267	0,291	0,306	0,299	0,291	0,024	0,000	-0,016	-0,009
β_{*9}	0,486	0,476	0,461	0,452	0,459	-0,027	-0,017	-0,002	0,007
β_{*10}	-0,607	-0,607	-0,608	-0,617	-0,614	-0,007	-0,007	-0,006	0,003

Tabela B.2: Médias *a posteriori* para diferentes valores de s e diferenças entre as estimativas, tamanho de amostra 10000. Aplicação II.

Parâmetro	1	2	5	10	50	50-1	50-2	50-5	50-10
δ_1	0,094	0,095	0,094	0,092	0,092	-0,002	-0,003	-0,001	0,001
δ_2	0,062	0,062	0,061	0,060	0,060	-0,002	-0,002	0,000	0,001
δ_3	0,146	0,148	0,148	0,147	0,144	-0,003	-0,004	-0,004	-0,003
δ_4	0,144	0,144	0,144	0,144	0,142	-0,002	-0,002	-0,002	-0,002
δ_5	0,088	0,092	0,092	0,091	0,089	0,001	-0,003	-0,003	-0,002
δ_6	0,079	0,076	0,078	0,077	0,078	-0,001	0,001	0,000	0,000
δ_7	0,130	0,130	0,130	0,132	0,132	0,002	0,002	0,001	0,000
δ_8	0,235	0,236	0,238	0,236	0,235	0,000	-0,001	-0,003	-0,002
δ_9	0,172	0,168	0,165	0,166	0,166	-0,005	-0,002	0,002	0,000
δ_{10}	0,125	0,126	0,125	0,127	0,126	0,001	0,001	0,002	0,000
α_1	1,635	1,644	1,635	1,629	1,629	-0,006	-0,015	-0,006	0,000
α_2	0,627	0,628	0,621	0,619	0,623	-0,004	-0,005	0,002	0,004
α_3	0,978	0,978	0,977	0,978	0,974	-0,003	-0,003	-0,003	-0,003
α_4	0,867	0,865	0,867	0,868	0,865	-0,002	-0,001	-0,002	-0,004
α_5	0,841	0,850	0,856	0,852	0,844	0,002	-0,006	-0,013	-0,008
α_6	1,699	1,674	1,676	1,677	1,687	-0,012	0,013	0,011	0,010
α_7	1,495	1,507	1,500	1,503	1,510	0,015	0,003	0,010	0,007
α_8	0,966	0,972	0,973	0,971	0,965	-0,001	-0,007	-0,008	-0,006
α_9	1,200	1,191	1,178	1,182	1,183	-0,017	-0,008	0,005	0,001
α_{10}	1,821	1,820	1,822	1,825	1,822	0,001	0,002	0,000	-0,003
β_{*1}	0,070	0,071	0,060	0,051	0,051	-0,019	-0,020	-0,009	0,000
β_{*2}	0,941	0,938	0,926	0,921	0,921	-0,020	-0,017	-0,005	0,000
β_{*3}	-0,092	-0,093	-0,095	-0,097	-0,105	-0,012	-0,011	-0,010	-0,008
β_{*4}	0,165	0,161	0,164	0,162	0,156	-0,009	-0,005	-0,008	-0,006
β_{*5}	0,529	0,543	0,546	0,540	0,527	-0,002	-0,016	-0,019	-0,013
β_{*6}	0,273	0,259	0,262	0,257	0,258	-0,016	-0,002	-0,004	0,001
β_{*7}	-0,239	-0,245	-0,245	-0,240	-0,242	-0,003	0,004	0,003	-0,001
β_{*8}	0,291	0,299	0,299	0,298	0,292	0,001	-0,007	-0,007	-0,006
β_{*9}	0,476	0,469	0,450	0,453	0,455	-0,021	-0,013	0,006	0,002
β_{*10}	-0,607	-0,611	-0,618	-0,617	-0,616	-0,008	-0,005	0,002	0,001

Tabela B.3: Relações entre SD e MC Error para os dados da Aplicação II.

Tam	m_1					m_2				
s	1	2	5	10	50	1	2	5	10	50
δ_1	5,7%	4,5%	2,9%	1,9%	1,7%	4,5%	3,1%	1,8%	1,3%	1,0%
δ_2	6,4%	4,6%	3,1%	2,2%	1,4%	4,8%	3,2%	2,2%	1,5%	2,5%
δ_3	5,5%	4,2%	2,8%	2,2%	1,4%	4,5%	2,9%	2,0%	1,5%	1,2%
δ_4	6,5%	5,3%	3,4%	2,2%	1,6%	5,7%	3,6%	2,7%	1,9%	1,1%
δ_5	6,4%	4,8%	2,9%	2,3%	1,4%	5,1%	3,3%	2,2%	1,9%	1,0%
δ_6	5,6%	4,4%	3,0%	2,3%	2,0%	4,0%	3,1%	2,2%	1,6%	1,2%
δ_7	5,7%	4,5%	3,1%	2,2%	1,4%	4,2%	2,8%	2,1%	1,6%	1,4%
δ_8	7,8%	6,6%	4,2%	3,1%	1,7%	6,3%	4,6%	3,1%	2,4%	1,0%
δ_9	6,5%	5,0%	3,4%	2,8%	1,5%	4,7%	3,9%	2,7%	2,1%	3,3%
δ_{10}	5,8%	4,0%	2,9%	2,2%	1,4%	4,2%	3,0%	1,9%	1,4%	1,3%
α_1	5,3%	4,2%	2,9%	1,8%	1,3%	3,9%	2,8%	1,8%	1,4%	1,0%
α_2	5,2%	3,8%	2,8%	2,1%	1,2%	4,1%	2,7%	1,9%	1,3%	1,0%
α_3	4,4%	3,4%	2,2%	1,9%	1,3%	3,6%	2,5%	1,6%	1,3%	0,9%
α_4	5,2%	4,5%	2,9%	1,9%	1,6%	4,5%	3,2%	2,2%	1,8%	1,0%
α_5	5,5%	4,6%	3,0%	2,3%	1,3%	4,7%	3,3%	2,1%	1,7%	1,0%
α_6	5,3%	4,2%	3,0%	2,2%	1,6%	3,6%	3,0%	2,1%	1,5%	1,0%
α_7	5,1%	4,1%	3,1%	2,0%	1,3%	3,6%	2,8%	2,1%	1,3%	0,9%
α_8	7,2%	6,2%	4,2%	2,9%	1,7%	6,0%	4,3%	2,8%	2,2%	0,9%
α_9	6,9%	4,9%	3,5%	2,4%	1,6%	5,1%	3,9%	2,4%	1,9%	1,1%
α_{10}	4,1%	3,1%	1,9%	1,5%	1,3%	3,2%	2,2%	1,5%	1,1%	1,0%
β_{*1}	7,2%	5,8%	3,8%	2,4%	1,5%	5,8%	4,0%	2,4%	2,0%	1,1%
β_{*2}	7,4%	5,3%	3,6%	2,3%	1,5%	5,7%	3,8%	2,6%	1,7%	1,0%
β_{*3}	6,1%	4,9%	3,3%	2,5%	1,5%	5,1%	3,7%	2,4%	1,8%	1,1%
β_{*4}	7,1%	5,9%	3,7%	2,5%	1,6%	6,4%	4,3%	3,0%	2,1%	1,2%
β_{*5}	7,3%	6,1%	3,3%	2,7%	1,4%	6,3%	4,1%	2,6%	2,1%	1,1%
β_{*6}	7,5%	6,2%	4,2%	3,1%	1,6%	5,6%	4,5%	2,9%	2,3%	1,1%
β_{*7}	7,1%	5,4%	3,7%	2,5%	1,6%	5,1%	3,7%	2,6%	2,0%	1,2%
β_{*8}	8,2%	7,1%	4,6%	3,4%	1,7%	6,9%	5,1%	3,2%	2,6%	1,1%
β_{*9}	7,5%	5,6%	4,0%	2,8%	1,7%	5,5%	4,5%	2,9%	2,2%	1,1%
β_{*10}	6,7%	5,1%	3,6%	2,7%	1,5%	4,8%	3,7%	2,5%	2,0%	1,1%
Máximo	8,2%	7,1%	4,6%	3,4%	2,0%	6,9%	5,1%	3,2%	2,6%	3,3%

Tabela B.4: Média *a posteriori* considerando diferentes distribuições *a priori*.

Aplicação II.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
δ_1	0,095	0,066	0,083	0,155	0,054	0,085	0,073	0,088
δ_2	0,062	0,035	0,054	0,141	0,031	0,041	0,033	0,044
δ_3	0,148	0,094	0,114	0,183	0,079	0,127	0,090	0,106
δ_4	0,144	0,094	0,118	0,186	0,080	0,131	0,085	0,107
δ_5	0,092	0,061	0,079	0,158	0,050	0,073	0,051	0,061
δ_6	0,076	0,057	0,073	0,140	0,045	0,075	0,065	0,079
δ_7	0,130	0,095	0,112	0,176	0,076	0,138	0,102	0,117
δ_8	0,236	0,183	0,178	0,210	0,127	0,257	0,154	0,226
δ_9	0,168	0,146	0,150	0,188	0,120	0,179	0,131	0,167
δ_{10}	0,126	0,090	0,109	0,175	0,072	0,136	0,136	0,157
α_1	1,644	1,561	1,615	1,794	1,390	1,827	1,678	1,937
α_2	0,628	0,585	0,613	0,753	0,546	0,648	0,553	0,656
α_3	0,978	0,913	0,940	1,004	0,840	1,048	0,894	0,982
α_4	0,865	0,801	0,831	0,911	0,736	0,956	0,770	0,874
α_5	0,850	0,806	0,833	0,967	0,737	0,900	0,760	0,846
α_6	1,674	1,611	1,681	1,871	1,423	1,912	1,790	2,004
α_7	1,507	1,426	1,470	1,610	1,279	1,744	1,507	1,700
α_8	0,972	0,886	0,860	0,862	0,725	1,226	0,821	1,095
α_9	1,191	1,140	1,138	1,205	0,968	1,425	1,083	1,344
α_{10}	1,820	1,731	1,777	1,936	1,539	2,144	2,468	2,770
β_{*1}	0,071	-0,019	0,034	0,215	-0,053	0,080	0,016	0,097
β_{*2}	0,938	0,824	0,898	1,281	0,806	0,868	0,809	0,888
β_{*3}	-0,093	-0,196	-0,156	-0,057	-0,217	-0,115	-0,195	-0,167
β_{*4}	0,161	0,051	0,102	0,232	0,020	0,168	0,033	0,086
β_{*5}	0,543	0,452	0,503	0,739	0,415	0,510	0,417	0,454
β_{*6}	0,259	0,191	0,255	0,443	0,141	0,308	0,257	0,329
β_{*7}	-0,245	-0,313	-0,276	-0,170	-0,344	-0,194	-0,287	-0,245
β_{*8}	0,299	0,158	0,135	0,164	0,013	0,461	0,096	0,331
β_{*9}	0,469	0,392	0,403	0,496	0,296	0,580	0,352	0,511
β_{*10}	-0,611	-0,673	-0,638	-0,555	-0,679	-0,592	-0,619	-0,570

Tabela B.5: Mediana *a posteriori* considerando diferentes distribuições *a priori*.

Aplicação II.

Parâmetro	P1	P2	P3	P4	P5	P6	P7	P8
δ_1	0,092	0,053	0,076	0,154	0,039	0,077	0,060	0,080
δ_2	0,052	0,021	0,044	0,139	0,017	0,025	0,019	0,028
δ_3	0,139	0,067	0,096	0,180	0,054	0,111	0,062	0,081
δ_4	0,135	0,069	0,105	0,183	0,053	0,116	0,054	0,085
δ_5	0,084	0,043	0,070	0,157	0,031	0,057	0,032	0,042
δ_6	0,071	0,048	0,068	0,137	0,035	0,070	0,055	0,074
δ_7	0,126	0,077	0,102	0,174	0,053	0,137	0,084	0,101
δ_8	0,249	0,179	0,169	0,208	0,097	0,284	0,131	0,240
δ_9	0,173	0,151	0,151	0,186	0,118	0,189	0,130	0,176
δ_{10}	0,116	0,071	0,099	0,172	0,051	0,132	0,130	0,155
α_1	1,597	1,511	1,563	1,760	1,350	1,750	1,530	1,718
α_2	0,606	0,576	0,597	0,727	0,535	0,626	0,541	0,632
α_3	0,948	0,884	0,910	0,987	0,820	1,002	0,863	0,944
α_4	0,830	0,771	0,800	0,890	0,715	0,894	0,737	0,834
α_5	0,821	0,777	0,807	0,942	0,719	0,858	0,738	0,816
α_6	1,627	1,565	1,631	1,836	1,390	1,850	1,602	1,815
α_7	1,450	1,371	1,422	1,575	1,244	1,665	1,385	1,537
α_8	0,901	0,814	0,796	0,833	0,684	1,120	0,733	0,955
α_9	1,142	1,085	1,090	1,171	0,924	1,368	0,997	1,249
α_{10}	1,775	1,686	1,735	1,895	1,505	2,089	1,963	2,236
β_{*1}	0,048	-0,053	0,007	0,201	-0,069	0,041	-0,032	0,025
β_{*2}	0,884	0,791	0,861	1,230	0,778	0,811	0,773	0,826
β_{*3}	-0,126	-0,225	-0,187	-0,069	-0,242	-0,159	-0,236	-0,208
β_{*4}	0,110	-0,194	0,058	0,211	-0,020	0,099	-0,014	0,031
β_{*5}	0,502	0,407	0,468	0,711	0,383	0,448	0,383	0,410
β_{*6}	0,234	0,164	0,230	0,421	0,124	0,276	0,197	0,258
β_{*7}	-0,265	-0,341	-0,304	-0,183	-0,367	-0,228	-0,329	-0,303
β_{*8}	0,244	0,069	0,054	0,136	-0,061	0,393	-0,016	0,214
β_{*9}	0,436	0,351	0,367	0,475	0,250	0,533	0,285	0,465
β_{*10}	-0,623	-0,681	-0,649	-0,551	-0,683	-0,614	-0,650	-0,616

Tabela B.6: Desvio padrão *a posteriori* para diferentes distribuições *a priori*.
Aplicação II.

δ_1	0,058	0,058	0,055	0,038	0,051	0,065	0,062	0,067
δ_2	0,047	0,038	0,042	0,035	0,036	0,044	0,038	0,047
δ_3	0,093	0,088	0,085	0,049	0,079	0,103	0,088	0,095
δ_4	0,091	0,088	0,084	0,049	0,080	0,105	0,087	0,093
δ_5	0,062	0,059	0,056	0,040	0,054	0,064	0,054	0,059
δ_6	0,048	0,048	0,046	0,034	0,042	0,052	0,053	0,055
δ_7	0,078	0,082	0,075	0,045	0,074	0,093	0,087	0,094
δ_8	0,116	0,128	0,113	0,056	0,113	0,131	0,126	0,133
δ_9	0,076	0,082	0,073	0,044	0,081	0,083	0,086	0,085
δ_{10}	0,081	0,079	0,074	0,045	0,070	0,092	0,098	0,104
α_1	0,355	0,336	0,345	0,356	0,273	0,447	0,607	0,915
α_2	0,177	0,145	0,159	0,231	0,133	0,180	0,145	0,186
α_3	0,225	0,199	0,206	0,197	0,170	0,265	0,209	0,229
α_4	0,222	0,198	0,204	0,203	0,163	0,294	0,200	0,229
α_5	0,208	0,197	0,201	0,225	0,166	0,243	0,176	0,206
α_6	0,353	0,344	0,359	0,354	0,271	0,453	1,094	0,757
α_7	0,339	0,320	0,319	0,321	0,256	0,463	0,574	0,646
α_8	0,339	0,306	0,287	0,225	0,215	0,502	0,333	0,495
α_9	0,339	0,331	0,318	0,286	0,268	0,454	0,377	0,476
α_{10}	0,377	0,355	0,357	0,376	0,293	0,493	2,108	1,855
β_{*1}	0,227	0,216	0,219	0,217	0,180	0,273	0,256	0,323
β_{*2}	0,255	0,181	0,213	0,311	0,168	0,236	0,188	0,312
β_{*3}	0,219	0,187	0,201	0,161	0,166	0,236	0,200	0,221
β_{*4}	0,256	0,222	0,226	0,190	0,193	0,306	0,219	0,242
β_{*5}	0,239	0,216	0,214	0,224	0,190	0,258	0,189	0,218
β_{*6}	0,230	0,214	0,225	0,221	0,180	0,267	0,370	0,340
β_{*7}	0,217	0,216	0,214	0,197	0,189	0,272	0,247	0,299
β_{*8}	0,385	0,366	0,334	0,208	0,275	0,523	0,366	0,500
β_{*9}	0,301	0,300	0,280	0,236	0,265	0,392	0,330	0,385
β_{*10}	0,230	0,217	0,220	0,219	0,190	0,283	0,358	0,361

Tabela B.7: Percentis das estimativas dos parâmetros δ para P_4 e P_1 . Aplicação II.

Parâmetro	P_4			P_1		
	2.5%	97.5%	Tam (95%)	2.5%	97.5%	Tam (95%)
δ_1	0,08509	0,2337	0,14861	0,005468	0,2189	0,213432
δ_2	0,07765	0,2129	0,13525	0,002573	0,1741	0,171527
δ_3	0,09745	0,286	0,18855	0,006816	0,3374	0,330584
δ_4	0,09845	0,2898	0,19135	0,00716	0,3346	0,32744
δ_5	0,08552	0,2389	0,15338	0,004139	0,2253	0,221161
δ_6	0,07809	0,2094	0,13131	0,004609	0,1819	0,177291
δ_7	0,0943	0,27	0,1757	0,007146	0,2889	0,281754
δ_8	0,1077	0,3257	0,218	0,01722	0,4298	0,41258
δ_9	0,1055	0,2753	0,1698	0,01816	0,3055	0,28734
δ_{10}	0,09249	0,2684	0,17591	0,006566	0,3	0,293434

Tabela B.8: Itens ordenados pela mediana *a posteriori* de δ usando diferentes distribuições *a priori*, aplicação II.

Proposta	>	<
P_2	8 9 7 10 4 3 1 6 5	2
P_1	8 9 3 4 7 10 1 5 6	2
P_3^*	8 9 4 7 10 3 1 5 6	2
P_4^*	8 9 4 3 7 10 5 1 2	6
P_5	9 8 3 7 4 10 1 6 5	2
P_6	8 9 7 10 4 3 1 6 5	2
P_7	8 9 10 7 3 1 6 4 5	2
P_8	8 9 10 7 4 3 1 6 5	2

Tabela B.9: Itens ordenados pela mediana *a posteriori* de α usando diferentes distribuições *a priori*. Aplicação II.

Proposta	>	<
P_2	10 6 1 7 9 3 8 5 4	2
P_1	10 6 1 7 9 3 8 4 5	2
P_3^*	10 6 1 7 9 3 5 4 8	2
P_4^*	10 6 1 7 9 3 5 4 8	2
P_5	10 6 1 7 9 3 5 4 8	2
P_6	10 6 1 7 9 8 3 4 5	2
P_7	10 6 1 7 9 3 5 4 8	2
P_8	10 6 1 7 9 8 3 4 5	2

Tabela B.10: Itens ordenados pela mediana *a posteriori* de β^* usando diferentes distribuições *a priori*. Aplicação II.

Proposta	>	<
P_2	2 5 9 6 8 1 4 3 7	10
P_1	2 5 9 8 6 4 1 3 7	10
P_3^*	2 5 9 6 4 8 1 3 7	10
P_4^*	2 5 9 6 4 1 8 3 7	10
P_5	2 5 9 6 4 8 1 3 7	10
P_6	2 9 5 8 6 4 1 3 7	10
P_7	2 5 9 6 4 8 1 3 7	10
P_8	2 9 5 6 8 4 1 3 7	10

Apêndice C

Tabela: Aplicação III

Tabela C.1: Médias *a posteriori* para α_{js} , β_{js} e β^*_{js} , ordenadas. Aplicação III

α		β_2		β_3	
Item	Média	Item	Média	Item	Média
11	1,319	1	3,216	1	5,708
5	1,140	10	2,174	2	3,628
12	1,096	9	1,860	6	2,736
8	0,933	6	1,749	5	2,339
6	0,644	12	1,069	10	2,174
3	0,605	5	1,031	9	2,143
10	0,586	3	0,955	3	2,009
7	0,578	8	0,767	8	1,778
9	0,530	7	0,362	2	2,341
1	0,515	11	0,260	7	1,129
2	0,403	2	0,036	11	0,847
4	0,389	4	-0,384	4	-0,107