

Universidade Federal de São Carlos  
Centro de Ciências Exatas e de Tecnologia  
Programa de Pós Graduação em Estatística  
Departamento de Estatística

# INFERÊNCIA BAYESIANA PARA TESTE DIAGNÓSTICO

Karolina Felcar Saraiva

Orientador: Prof. Dr. Josemar Rodrigues

Dissertação apresentada ao Departamento de Estatística da Universidade Federal de São Carlos como parte dos requisitos necessários para obtenção do título de Mestre em Estatística

São Carlos

Março - 2004

**Ficha catalográfica elaborada pelo DePT da  
Biblioteca Comunitária/UFSCar**

S243ib

Saraiva, Karolina Felcar.

Inferência bayesiana para teste diagnóstico / Karolina Felcar Saraiva. -- São Carlos : UFSCar, 2004.  
200 p.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2004.

1. Teoria bayesiana de decisão estatística. 2. Gibbs sampling. 3. Variável latente. 4. Especificidade a posteriori. 5. Sensibilidade a posteriori. 6. Prevalência a posteriori. I. Título.

CDD: 519.542 (20<sup>a</sup>)

# Agradecimentos

Agradeço a todos que me ajudaram a concluir esta dissertação, principalmente

A Deus pela vida, pela proteção e pelas oportunidades;

A meus pais, pelo apoio incondicional, confiança, amor e presença durante todos estes anos;

Ao meus irmãos, pelo amor e amizade;

Ao meu namorado, pelo incentivo, apoio e confiança;

Ao professor Dr. Josemar Rodrigues pela paciência, compreensão e dedicada orientação em todas as etapas do trabalho;

A todos os professores e funcionários do Departamento de Estatística da UFSCar, especialmente, Carlos Diniz, José Galvão Leite e Luís Milan, que participaram da minha formação profissional e pessoal;

Aos amigos da pós-graduação: Clécia, Eliza, Érica, Gislene, Hélio, Miriam, Silvia Lamas e Silvia Reyes;

À CAPES pelo apoio financeiro.

Muito obrigada!

*Dedico este trabalho a meus pais,  
José Salvador e Clarice pelo amor,  
compreensão e incentivo dispensados  
ao longo de toda a minha vida.*

# Abstract

The simpler screening tests applied to detect disease instead of the more elaborated, usually result in the risk of incorrect diagnostic. However, these tests are only useful when the risks of misclassifications are known and considered acceptably low.

So, with the purpose of looking for information on the proprieties of screening tests, as well as measuring their error rates, a Bayesian procedure was formulated using a simulation technique (Gibbs Sampling with latent variables) for estimation of the parameters of interest in the absence of a gold standard.

Two applications to real data have been explored. The first one refers to the detection of the infection caused by the strongyloides parasite on 162 refugees from Cambodia that arrived in Montreal, Canada, between July 1982 to February 1983, using data from serologic test and stool examination. The second one has the purpose detecting the obesity rates on males and females school pupils through the information supplied by the anthropometric Must and Cole criteria.

## Resumo

O uso de testes diagnósticos mais simples como substitutos dos mais elaborados para indicar a presença de doença, geralmente resulta em risco de diagnóstico incorreto. Entretanto, estes testes são úteis quando os riscos de erros de classificação são conhecidos e aceitavelmente baixos.

Então, com o objetivo de obter informações das propriedades de testes diagnósticos, assim como medir suas taxas de erro, formulou-se um procedimento bayesiano utilizando uma técnica de simulação (*Gibbs Sampling* com variáveis latentes) para estimação dos parâmetros de interesse na ausência de um padrão ouro.

Duas aplicações com dados reais foram exploradas. A primeira refere-se a detecção da infecção causada pelo parasita *strongyloides* em 162 refugiados do Camboja que chegaram em Montreal, Canadá, entre julho de 1982 a fevereiro de 1983, usando dados do teste sorológico e exame de fezes. A segunda, tem por objetivo detectar as taxas de obesidade em escolares do sexo masculino e feminino, através das informações fornecidas pelos critérios antropométricos Must e Cole.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Propriedades de um Teste Diagnóstico . . . . .	5
1.1.1	Prevalência . . . . .	7
1.1.2	Valor Preditivo . . . . .	8
1.1.3	Testes Múltiplos . . . . .	9
1.2	Apresentação da Tese . . . . .	10
<b>2</b>	<b>Verossimilhança para Teste Diagnóstico</b>	<b>12</b>
2.1	Verossimilhança . . . . .	15
2.2	Análise Bayesiana . . . . .	19
2.3	Distribuições <i>a priori</i> . . . . .	37
<b>3</b>	<b>Inferência Bayesiana para Dados Dicotômicos com Erros de Classificação</b>	<b>38</b>
3.1	Mistura de Modelos . . . . .	40
3.2	Modelo Misto . . . . .	40
3.2.1	Algoritmo <i>Gibbs Sampling</i> com Amostra Ampliada . . . . .	43
3.3	Modelo Misto aplicado a Teste Diagnóstico . . . . .	44
3.3.1	Algoritmo <i>Gibbs Sampling</i> para o Modelo com Erros de Classificação	52
3.4	Algoritmo . . . . .	53

<b>4</b>	<b>Aplicação: Diagnóstico do Parasita <i>Strongyloides</i></b>	<b>61</b>
4.1	Estimação Bayesiana dos Parâmetros de Teste Diagnóstico na ausência de um Padrão Ouro . . . . .	62
4.1.1	Um Teste Diagnóstico . . . . .	63
4.1.2	Dois Testes Diagnósticos . . . . .	72
4.1.3	Distribuições <i>a priori</i> . . . . .	82
4.2	Infecção de <i>Strongyloides</i> em Refugiados do Camboja . . . . .	88
4.3	Resultados . . . . .	89
4.3.1	Teste Sorológico . . . . .	90
4.3.2	Exame de Fezes . . . . .	94
4.3.3	Combinação do Teste Sorológico e Exame de Fezes . . . . .	98
4.4	A Indicação de Convergência . . . . .	104
4.5	Comentários . . . . .	111
<b>5</b>	<b>Aplicação: Diagnóstico de Obesidade</b>	<b>117</b>
5.1	Obesidade e Critérios Antropométricos . . . . .	118
5.2	Obesidade em Escolares . . . . .	119
5.3	Resultados . . . . .	121
5.3.1	Critérios Antropométricos Must e Cole para o Sexo Feminino . . . . .	121
5.3.2	Critérios Antropométricos Must e Cole para o Sexo Masculino . . . . .	129
5.3.3	Combinação dos Critério Must e Cole para ambos os Sexos . . . . .	136
5.4	A Indicação de Convergência . . . . .	148
5.4.1	Critérios Antropométricos Must e Cole para o Sexo Feminino . . . . .	148
5.4.2	Critérios Antropométricos Must e Cole para o Sexo Masculino . . . . .	152
5.4.3	Combinação dos Critério Must e Cole para ambos os Sexos . . . . .	155
5.5	Comentários . . . . .	161
<b>6</b>	<b>Conclusões e Perspectivas Futuras</b>	<b>169</b>

<b>Apêndice</b>	<b>171</b>
A. Transformação de Variável . . . . .	171
B. Gibbs Sampling: Definições e Propriedades . . . . .	173
C. Diagnósticos de Convergência . . . . .	175
D. Diagnóstico de Convergência de Gelman Rubin (1992) . . . . .	176
E. Software WinBUGS . . . . .	176
F. Programas para o cálculo das estimativas via Gibbs Sampling do Teste Sorológico, Exame de Fezes e Combinação dos dois Testes . . . . .	178
G. Programas para o cálculo das estimativas via Gibbs Sampling do Critério de Must, Cole e Combinação dos dois Critérios . . . . .	181
<b>Referências Bibliográficas</b>	<b>189</b>

# Capítulo 1

## Introdução

Segundo Medronho *et al.* (2002) grande parte do trabalho clínico consiste na busca diagnóstica de doenças. Para tal, o médico lança mão de um arsenal preliminar que se inicia com a anamnese (informação sobre o princípio e evolução de uma doença até a primeira observação do médico) e exame físico, podendo ir até exames mais sofisticados como a ressonância magnética e o PCR. Além disso, conscientemente ou não, o médico utiliza-se de um raciocínio probabilístico para realizar o diagnóstico de um paciente: por exemplo, qual a probabilidade de um diagnóstico de meningite em um paciente com febre, vômito e cefaléia? Se o paciente apresentar rigidez de nuca, a probabilidade de se diagnosticar a doença aumenta? Note que o paciente tem ou não a doença, entretanto, o clínico não tem ainda a certeza do diagnóstico e para eliminar ou reduzir ao máximo o seu grau de incerteza, utiliza os testes diagnósticos.

Geralmente, denomina-se de testes diagnósticos àqueles exames realizados em laboratório (hemograma completo, bioquímica do liquor, etc.); todavia, um conjunto de sinais e sintomas também pode ser visto como um teste para o diagnóstico de uma determinada doença.

Como já foi descrito anteriormente, o procedimento diagnóstico é um processo que inclui um razoável grau de incerteza, que é aumentada ou diminuída na dependência

de um bom juízo crítico por parte dos médicos, além de um sólido conhecimento da literatura médica. Assim, a prática médica moderna utiliza-se das leis de probabilidade como um importante auxiliar na interpretação dos testes diagnósticos. Ao solicitar um teste diagnóstico o médico se vê diante de quatro possibilidades: o exame resultar positivo na presença da doença (verdadeiro positivo), positivo na sua ausência (falso positivo), negativo na ausência da doença (verdadeiro negativo) e negativo na sua presença (falso negativo). A Tabela 1.1 mostra a relação entre essas quatro possibilidades.

**Tabela 1.1.** Possíveis Resultados de um Teste Diagnóstico para Identificar uma Doença.

		DOENÇA		Total
		Presente	Ausente	
TESTE	Positivo	<sup>a</sup> verdadeiro positivo	<sup>b</sup> falso positivo	$a + b$
	Negativo	<sup>c</sup> falso negativo	<sup>d</sup> verdadeiro negativo	$c + d$
Total		$a + c$	$b + d$	$a + b + c + d$

Então, na prática há dois tipos de erros de classificação em teste diagnóstico envolvendo a proporção de indivíduos que possuam ou não determinada doença. Quando os erros de classificação ou erros de codificação são considerados, os dados obtidos muitas vezes não expressam a realidade. Por exemplo, suponha que se tenha interesse na proporção de indivíduos que possuam uma certa doença. Dois tipos de erros são possíveis: um indivíduo que não tem a doença, e após um teste diagnóstico é considerado como doente e um indivíduo que tem a doença, e após um teste diagnóstico é considerado como não doente. O problema fundamental é como utilizar este tipo de dado para fazer inferência a respeito da proporção real de indivíduos que possuam a referida doença (proporção de interesse).

A avaliação da acurácia de um teste baseia-se na sua relação com algum meio de saber se a doença está ou não realmente presente - um indicador mais fiel da verdade, geralmente é referido como “padrão-ouro” (*gold standard*). O que acontece é que este “padrão-ouro” é freqüentemente difícil de ser encontrado. Algumas vezes o “padrão-ouro”

é, por si só, um teste simples e barato, por exemplo, um teste de anticorpo para o vírus da imunodeficiência humana. Contudo, usualmente este não é o caso. Mais freqüentemente, para se estar certo de que a doença está realmente presente ou ausente, deve-se lançar mão de testes relativamente elaborados, caros ou arriscados. Entre estes, estão a biópsia, a exploração cirúrgica e, certamente, a autópsia.

Como esses meios mais acurados de estabelecer um diagnóstico correto são quase sempre mais caros e invasivos, clínicos e pacientes preferem testes mais simples do que um teste-padrão rigoroso, pelo menos inicialmente. Por exemplo, eletrocardiogramas e enzimas séricas são habitualmente usados para estabelecer o diagnóstico de infarto agudo do miocárdio, em vez de cateterismo ou exames de imagem. O uso de testes mais simples como substitutos dos mais elaborados e mais exatos no estabelecimento da presença de doença é feito com o entendimento de que isso resulta em um certo risco de diagnóstico incorreto. Tal risco é justificado pela segurança e conveniência do teste mais simples. Entretanto, estes testes são úteis somente quando os riscos de erros de classificação são conhecidos e considerados aceitavelmente baixos.

Então, é necessário que se faça alguma tentativa para regularizar ou medir estes tipos de erros e, conseqüentemente, para que as taxas de erros sejam conhecidas. Com isso, a tomada de decisão sobre os resultados obtidos para as proporções de interesse será mais realística. E devido a grande utilização de novos testes, os laboratórios ou clínicas têm a necessidade de buscar informações com maior confiabilidade sobre os dados dos testes, justificando assim a importância do tratamento deste tema.

Este trabalho teve como objetivo formular um procedimento Bayesiano para teste diagnóstico utilizando uma técnica de simulação, *Gibbs Sampling* com variáveis latentes, para estimação dos parâmetros de interesse na ausência de um padrão-ouro. Exemplos envolvendo a aplicação desta metodologia a dados reais são apresentados.

A análise bayesiana consiste na obtenção de medidas resumos ou densidades *a posteriori* para os parâmetros de interesse nos modelos em estudo. Estas medidas ou densidades são obtidas combinando as informações *a priori* sobre os parâmetros de interesse (densi-

dade *a priori*) e informações contidas nos dados da amostra (função de verossimilhança). Quando as distribuições *a priori* e a função de verossimilhança são conjugadas, a obtenção das medidas de interesse, em geral, é imediata. No entanto, nem sempre isso se verifica e as distribuições *a posteriori* resultantes são em geral, analiticamente, intratáveis. Nestes casos, métodos de aproximação numérica precisam ser utilizados.

Nos últimos anos, métodos avançados de aproximação numérica para obtenção de densidades marginais *a posteriori*, têm sido apresentados na literatura. Estes métodos podem ser subdivididos em métodos numéricos analíticos e métodos baseados em amostragem. De acordo com a literatura, embora os métodos analíticos sejam considerados mais precisos, apresentam a desvantagem de serem baseados em resultados assintóticos e na suposição de normalidade, perdendo a tratabilidade à medida em que a dimensão do espaço paramétrico aumenta. Desta forma, a aplicabilidade destes métodos para uma análise bayesiana completa em modelos mais complexos é limitada. Por outro lado, os métodos baseados em amostragem podem ser aplicados mesmo em modelos com estruturas bastante complexas, embora devam ser utilizados somente quando a aplicação dos métodos analíticos seja intratável. Isto porque, como são baseados em geração de números aleatórios, são menos precisos que os analíticos.

Os métodos numéricos baseados em amostragem englobam os métodos de Monte Carlo Simples, os métodos de reamostragem por importância e os métodos de Monte Carlo via Cadeias de Markov. Dentre os métodos de Monte Carlo via Cadeias de Markov os mais citados são o *Gibbs Sampling* (Geman & Geman, 1984) e o algoritmo *Metropolis-Hastings* (Hitchcock, 2003). Neste trabalho, foi utilizado o *Gibbs Sampling* para a obtenção das distribuições *a posteriori*. Este método consiste em um esquema markoviano de atualização, que permite a obtenção de amostras de uma distribuição conjunta, através de amostragens iterativas das distribuições condicionais completas. O grande interesse por esta técnica é devido à sua simplicidade conceitual e facilidade de implementação.

Na ausência de técnicas gerais para se determinar o comprimento da cadeia, *a priori*, análises estatísticas devem ser realizadas, *a posteriori*, para assegurar a convergência da

cadeia. Estes procedimentos são chamados diagnósticos de convergência.

A monitoração informal da convergência é feita através da análise do comportamento da trajetória de uma cadeia ao longo das iterações, da monitoração dos gráficos das densidades *a posteriori*, estimadas ao longo das iterações, e de gráficos das médias ergódicas ao invés de valores gerados.

A verificação formal da convergência é baseada em propriedades estatísticas das cadeias simuladas. Brooks & Roberts (1995) discutiram vários métodos de diagnósticos de convergência, apresentados na literatura, com particular ênfase sobre implementação. Os métodos são comparados em termos de sua interpretabilidade e aplicabilidade com recomendações para particulares classes de problemas. Dentre eles, está o método de Gelman & Rubin (1992) que foi explorado neste trabalho.

O *Software WinBUGS*, versão interativa do *Windows* para o programa BUGS (Bayesian Inference Using Sample) (Spiegelhalter, Thomas, Best e Gilks, 1994), programa este direcionado à aplicação da Inferência Bayesiana, em problemas estatísticos, usando o *Gibbs Sampling*, foi utilizado devido mostrar-se uma ferramenta simples e muito eficiente, pois fornece automaticamente os resumos decorrentes da amostra obtida como média, mediana, desvio padrão, intervalo de credibilidade e, também, gráficos para análise das trajetórias das cadeias geradas, densidades, autocorrelações e diagnósticos de convergência de Gelman & Rubin (1992) de cada parâmetro.

## 1.1 Propriedades de um Teste Diagnóstico

Considere que as variáveis aleatórias  $D$  e  $T$ , correspondam a doença (realidade desconhecida) e teste (observável), respectivamente. Cada uma destas variáveis aleatórias assumem somente os valores 0 e 1. Assim, a partir da Tabela 1.1 é possível calcular as seguintes propriedades de um teste diagnóstico:

- **Sensibilidade ( $\eta$ )**

É a proporção de verdadeiros positivos entre todos os doentes. Expressa a probabilidade de um teste dar positivo na presença da doença, isto é, avalia a capacidade do teste detectar a doença quando ela está de fato presente. É calculada da seguinte forma:

$$\hat{\eta} = \frac{a}{a + c}.$$

Em termos de probabilidade é dada por:

$$\eta = P(T = 1|D = 1).$$

- **Especificidade ( $\theta$ )**

É a proporção de verdadeiros negativos entre todos sadios. Expressa a probabilidade de um teste dar negativo na ausência da doença, isto é, avalia a capacidade de o teste afastar a doença quando ela está ausente. É calculada da seguinte forma:

$$\hat{\theta} = \frac{d}{b + d}.$$

Em termos de probabilidade é dada por:

$$\theta = P(T = 0|D = 0).$$

Essas propriedades são essenciais a todo teste diagnóstico e devem ser consideradas ao solicitá-lo.

Note que um teste muito sensível raramente deixará de diagnosticar indivíduos com a doença e que um teste muito específico raramente classificará como doente um indivíduo sem a doença (Fletcher *et al.*, 1996).

Os testes sensíveis são muito utilizados para realizar o rastreamento de doença em grupos populacionais como, por exemplo, os inquéritos de soro-prevalência para doenças infecciosas e o uso do teste anti-HIV em bancos de sangue.

Os testes específicos são utilizados para confirmar um diagnóstico, já que um teste muito específico raramente resultará positivo na ausência da doença, ou seja, dá poucos resultados falsos positivos. São particularmente necessários quando o resultado falso positivo pode lesar o paciente tanto física, como emocional ou financeiramente, como, por exemplo, o teste anti-HIV.

### 1.1.1 Prevalência

- **Prevalência ( $\pi$ )**

Refere-se a todos os casos da doença existentes previamente à realização de um teste. Dessa forma, ela expressa a probabilidade da doença antes de o teste ser realizado, por isso é também denominada probabilidade pré-teste. É calculada da seguinte forma:

$$\hat{\pi} = \frac{a + c}{a + b + c + d}.$$

Em termos de probabilidade é dada por:

$$\pi = P(D = 1).$$

Note que essa medida não necessariamente significa a prevalência da doença na população em geral e sim naquela determinada população em estudo. Assim, ela pode ser derivada de um determinado estrato populacional, por exemplo. Para uma mesma doença, sua prevalência pode variar enormemente dentro de um país, estado, etc. Além disso, ela pode se modificar após a realização de um teste.

## 1.1.2 Valor Preditivo

Ao solicitar um teste diagnóstico, o médico leva em conta a sensibilidade e a especificidade. Entretanto, dado que o exame foi realizado, qual a probabilidade de um resultado positivo ter realmente identificado a doença? A probabilidade de doença, dados os resultados de um teste diagnóstico é denominada valor preditivo do teste, que pode ser dividido em:

- **Valor Preditivo Positivo ( $\gamma$ )**

É a proporção de verdadeiros positivos entre todos os indivíduos com teste positivo. Expressa a probabilidade de um paciente com o teste positivo ter a doença. É calculado do seguinte modo:

$$\hat{\gamma} = \frac{a}{a + b}.$$

Em termos de probabilidade é dado por:

$$\gamma = P(D = 1|T = 1).$$

- **Valor Preditivo Negativo ( $\delta$ )**

É a proporção de verdadeiros negativos entre todos os indivíduos com teste negativo. Expressa a probabilidade de um paciente com teste negativo não ter a doença. É calculado da seguinte maneira:

$$\hat{\delta} = \frac{d}{c + d}.$$

Em termos de probabilidade é dado por:

$$\delta = P(D = 0|T = 0).$$

## • Determinantes do Valor Preditivo

O valor preditivo não é propriedade apenas do teste, sendo determinado, além de pela sensibilidade e especificidade, pela prevalência da doença na população testada. Nesse caso, a prevalência da doença deve ser interpretada como a probabilidade de um indivíduo ter a doença antes da realização do teste diagnóstico, por isso ela é denominada também probabilidade pré-teste.

Os valores preditivos positivo e negativo são estimativas da mesma probabilidade para aqueles indivíduos que tiveram os testes positivo e negativo, respectivamente. Por isso, o valor preditivo é denominado também probabilidade pós-teste.

Os valores preditivos positivo e negativo de um teste podem ser calculados para qualquer prevalência de doença de acordo com as fórmulas abaixo:

$$\gamma = \frac{\pi\eta}{\pi\eta + (1 - \pi)(1 - \theta)},$$

$$\delta = \frac{(1 - \pi)\theta}{\pi(1 - \eta) + (1 - \pi)\theta}.$$

Quanto mais sensível o teste, melhor seu valor preditivo negativo, isto é, maior a segurança de que um paciente com resultado negativo não tenha a doença.

Quanto mais específico o teste, melhor o seu valor preditivo positivo, ou seja, maior a segurança de que um paciente com resultado positivo tenha a doença.

Para um mesmo teste, quanto maior a prevalência da doença, maior o valor preditivo positivo e menor o valor preditivo negativo.

### 1.1.3 Testes Múltiplos

Na prática, não existe teste diagnóstico com sensibilidade e especificidade iguais a 100%. Assim, o clínico não pode se basear apenas em um teste para diagnosticar uma

doença. Além disso, muitos testes são muito caros ou oferecem risco ou desconforto ao paciente, necessitando a realização prévia de um teste mais barato ou mais seguro, mesmo se o teste apresentar sensibilidade e especificidades inferiores. Frequentemente, é necessário lançar mão de um verdadeiro arsenal para diagnosticar uma determinada doença, solicitando um conjunto de testes diagnósticos.

A utilização de múltiplos testes na prática clínica aumenta a qualidade do diagnóstico, diminuindo o número de resultados falsos.

Quando os resultados dos testes são consistentes, ou seja, todos positivos ou todos negativos, o processo de diagnóstico ou afastamento de uma doença é fácil, o problema é quando os resultados são contraditórios. Neste caso, a interpretação torna-se mais complicada.

Os testes múltiplos podem ser solicitados todos ao mesmo tempo (testes em paralelo), considerando-se o resultado positivo em qualquer um dos testes como evidência de doença, ou serem solicitados seqüencialmente (testes em série), onde novos testes são solicitados em função do resultado do teste anterior. Neste caso, todos os resultados devem ser positivos para que se possa estabelecer o diagnóstico de uma doença.

## 1.2 Apresentação da Tese

A tese está organizada em seis capítulos. Este primeiro capítulo discute a motivação para estudar o assunto, apresenta algumas propriedades sobre teste diagnóstico e expõe uma idéia geral da estrutura da tese. As referências básicas para seguir este capítulo são Medronho *et al.* (2002, capítulo 18) e Fletcher *et al.* (1996).

O segundo capítulo, apresenta a função de verossimilhança para teste diagnóstico, o problema de identificação dos parâmetros do modelo, onde foi introduzida a restrição  $(1 - \eta^*) > \theta^*$  para evitar este problema; a análise bayesiana desenvolvida para obtenção da densidade conjunta *a posteriori* exata e das distribuições marginais *a posteriori* exatas

para  $\pi, \eta^*$  e  $\theta^*$ , e explica o motivo da utilização do *Gibbs Sampling*. A referência para seguir este capítulo é Silveira (1999, capítulo 2).

No terceiro capítulo um breve histórico sobre Misturas de Modelos, definições sobre o Modelo Misto e o algoritmo *Gibbs Sampling* para amostra ampliada, são mostrados. O modelo misto para teste diagnóstico e o algoritmo *Gibbs Sampling* com variável latente para este modelo, foram desenvolvidos utilizando a teoria sobre modelo misto e o algoritmo *Gibbs Sampling* para amostra ampliada. O algoritmo *Gibbs Sampling* para o modelo com erros de classificação, modificado devido a consideração dos resultados de testes positivos e negativos também é dado neste capítulo. As referências centrais para este capítulo são Silveira (1999, capítulo 3), Diebolt & Robert (1994) e Joseph *et al.* (1995).

O Capítulo 4, mostra a estimação bayesiana dos parâmetros de um teste e de dois testes diagnósticos na ausência de um padrão ouro, assim como a aplicação envolvendo esta metodologia para os dados reais sobre a infecção de *strongyloides* em 162 refugiados do Camboja. Para o desenvolvimento da aplicação foi utilizado o *Software WinBUGS*, discutido resumidamente no Apêndice E. Aqui a referência principal foi o artigo de Joseph *et al.* (1995).

O Capítulo 5 trata da aplicação sobre a obesidade em escolares do sexo feminino e masculino, com o objetivo de estimar os parâmetros de um único critério antropométrico e da combinação dos dois critérios (Must e Cole). Esta aplicação considera os métodos bayesianos descritos no Capítulo 4 e o *Software WinBUGS*. As referências básicas utilizadas nesse capítulo foram Sotelo (2003), Cole *et al.* (2000) e Must *et al.* (1991).

No Capítulo 6, algumas considerações são feitas, com base nos resultados obtidos, e as conclusões e perspectivas futuras são apresentadas.

## Capítulo 2

# Verossimilhança para Teste Diagnóstico

Esse capítulo foi desenvolvido baseando-se na Tese de Mestrado de Vanda Donizetti Redondo Silveira, com orientação do Professor Doutor Josemar Rodrigues: Inferência Bayesiana para pesquisa de mercado com erros de resposta utilizando modelos mistos, 1999. Neste capítulo será tratada a verossimilhança desenvolvida para teste diagnóstico, o problema de identificação dos parâmetros para o modelo, análise bayesiana exata e a razão da utilização do *Gibbs Sampling*.

Os dados de testes diagnósticos são coletados para fazer vários tipos de inferência, com o propósito de estudar o comportamento de um paciente ter uma determinada doença. É reconhecido nas literaturas relacionadas ao assunto que tais dados podem estar com erro. Por esse motivo, introduziu-se um procedimento, que leva em consideração dois tipos de erros que podem ocorrer em dados de testes diagnósticos na ausência de um padrão-ouro, como por exemplo: em um teste diagnóstico para uma certa doença, o teste pode ser negativo para um paciente que possua a doença (falso negativo), e um teste pode ser positivo para um paciente que não possua a doença (falso positivo). Portanto, o modelo a ser estudado, contém dois parâmetros de taxa de erro além do parâmetro de

interesse sobre o teste diagnóstico. O modelo probabilístico proposto, revela um problema de identificação dos parâmetros, que pode ser evitado quando uma restrição é introduzida. Para entender melhor o problema, considere o seguinte:

- Seja  $\pi$  (prevalência) representando a proporção de indivíduos que possuem determinada doença. Se as informações são registradas com precisão em um teste diagnóstico, então os dados dicotômicos seguem um processo de Bernoulli com parâmetro  $\pi$ ; porém erros nos dados modificam este processo. Considere:
- **Complemento da Sensibilidade ( $1 - \eta$ )**

É a proporção de falsos negativos entre todos os doentes. Expressa a probabilidade de um teste dar negativo na presença da doença, isto é, avalia a capacidade do teste não detectar a doença quando ela está presente. É calculada da seguinte forma:

$$\hat{\eta}^* = 1 - \hat{\eta} = \frac{c}{a + c}.$$

Em termos de probabilidade é dada por:

$$\eta^* = 1 - \eta = P(T = 0|D = 1).$$

- **Complemento da Especificidade ( $1 - \theta$ )**

É a proporção de falsos positivos entre todos os sadios. Expressa a probabilidade de um teste dar positivo na ausência da doença, isto é, avalia a capacidade de o teste diagnosticar a doença quando ela está ausente. É calculada da seguinte maneira:

$$\hat{\theta}^* = 1 - \hat{\theta} = \frac{b}{b + d}.$$

Em termos de probabilidade é dada por:

$$\theta^* = 1 - \theta = P(T = 1|D = 0).$$

- **Complemento da Prevalência ( $1 - \pi$ )**

Refere-se a todos os casos da doença inexistentes previamente à realização de um teste. Dessa forma, ela expressa a probabilidade da ausência de doença antes de o teste ser realizado. É calculada da seguinte forma:

$$1 - \hat{\pi} = \frac{b + d}{a + b + c + d}.$$

Em termos de probabilidade é dada por:

$$1 - \pi = P(D = 0).$$

Utilizando as propriedades da teoria de probabilidade que trata do cálculo de probabilidades condicionais, pode-se expressar as seguintes probabilidades dos dados sujeitos a erro:

- **Probabilidade do Teste Positivo ( $\tau$ )**

É a probabilidade do indivíduo ser registrado como tendo a doença.

$$\begin{aligned} \tau &= P(T = 1) = P(D = 1, T = 1) + P(D = 0, T = 1) & (2.1) \\ &= P(D = 1)P(T = 1|D = 1) + P(D = 0)P(T = 1|D = 0) \\ &= \pi\eta + (1 - \pi)(1 - \theta) = \pi(1 - \eta^*) + (1 - \pi)\theta^*. \end{aligned}$$

- **Probabilidade do Teste Negativo ( $1 - \tau$ )**

É a probabilidade do indivíduo ser registrado como não tendo doença. Dada pelo complemento da probabilidade do teste positivo.

$$1 - \tau = P(T = 0) = P(D = 1, T = 0) + P(D = 0, T = 0) \quad (2.2)$$

$$\begin{aligned}
&= P(D = 1)P(T = 0|D = 1) + P(D = 0)P(T = 0|D = 0) \\
&= \pi(1 - \eta) + (1 - \pi)\theta = \pi\eta^* + (1 - \pi)(1 - \theta^*).
\end{aligned}$$

Nesse caso, o registro de indivíduos com e sem a doença segue um processo de Bernoulli, com parâmetro  $\tau$  ao invés de  $\pi$ .

## 2.1 Verossimilhança

Considere que numa amostra aleatória de  $n$  indivíduos selecionados um de cada vez e com reposição, são registrados  $a$  como tendo a doença e  $b = n - a$  como não tendo a doença. Então, a função de verossimilhança correspondente será proporcional a distribuição binomial descrita como:

$$\begin{aligned}
L(\pi, \eta^*, \theta^* | n, a, b) &\propto \tau^a (1 - \tau)^b & (2.3) \\
&\propto [\pi(1 - \eta^*) + (1 - \pi)\theta^*]^a [\pi\eta^* + (1 - \pi)(1 - \theta^*)]^b.
\end{aligned}$$

Se as taxas de erros  $\eta^* = (1 - \eta)$  e  $\theta^* = (1 - \theta)$  são conhecidas, então o estimador de máxima verossimilhança para  $\pi$  baseado em (2.3), é dado pelo seguinte resultado:

**Resultado 2.1** - O Estimador de Máxima Verossimilhança para  $\pi$  com  $\eta^*$  e  $\theta^*$  conhecidos é dado por:

$$\hat{\pi} = \begin{cases} 1 & \text{se } [(a/n) - (1 - \theta)]/(\eta + \theta - 1) > 1 \\ [(a/n) - (1 - \theta)]/(\eta + \theta - 1) & \text{se } 0 \leq [(a/n) - (1 - \theta)]/(\eta + \theta - 1) \leq 1 \\ 0 & \text{se } [(a/n) - (1 - \theta)]/(\eta + \theta - 1) < 0 \end{cases} \quad (2.4)$$

**Demonstração:** Aplicando o logaritmo em  $L(\pi, \eta^*, \theta^* | n, a, b)$ , obtém-se:

$$\begin{aligned}\ln L(\pi, \eta^*, \theta^* | n, a, b) &= \ln([\pi(1 - \eta^*) + (1 - \pi)\theta^*]^a [\pi\eta^* + (1 - \pi)(1 - \theta^*)]^b) \\ &= a \ln[\pi(1 - \eta^*) + (1 - \pi)\theta^*] + b \ln[\pi\eta^* + (1 - \pi)(1 - \theta^*)].\end{aligned}$$

Derivando em relação a  $\pi$  e igualando a zero, tem-se:

$$\begin{aligned}\frac{\partial \ln L(\pi, \eta^*, \theta^* | n, a, b)}{\partial \pi} &= 0 \\ \Rightarrow \frac{a[1 - \eta^* - \theta^*]}{\pi(1 - \eta^*) + (1 - \pi)\theta^*} + \frac{b[\eta^* + \theta^* - 1]}{\pi\eta^* + (1 - \pi)(1 - \theta^*)} &= 0 \\ \Rightarrow \frac{a[1 - \eta^* - \theta^*]}{\pi(1 - \eta^*) + (1 - \pi)\theta^*} + \frac{(n - a)[\eta^* + \theta^* - 1]}{\pi\eta^* + (1 - \pi)(1 - \theta^*)} &= 0 \\ \Rightarrow \frac{a[1 - \eta^* - \theta^*]}{\pi(1 - \eta^*) + (1 - \pi)\theta^*} + \frac{n[\eta^* + \theta^* - 1]}{\pi\eta^* + (1 - \pi)(1 - \theta^*)} - \frac{a[\eta^* + \theta^* - 1]}{\pi\eta^* + (1 - \pi)(1 - \theta^*)} &= 0 \\ \Rightarrow \frac{a[1 - \eta^* - \theta^*]}{\pi(1 - \eta^*) + (1 - \pi)\theta^*} + \frac{a[1 - \eta^* - \theta^*]}{\pi\eta^* + (1 - \pi)(1 - \theta^*)} + \frac{n[\eta^* + \theta^* - 1]}{\pi\eta^* + (1 - \pi)(1 - \theta^*)} &= 0 \\ \Rightarrow a(1 - \eta^* - \theta^*)[\pi - \pi\eta^* + \theta^* - \pi\theta^* + \pi\eta^* + 1 - \theta^* - \pi + \pi\theta^*] + n(\eta^* + \theta^* - 1) &= 0 \\ &[\pi - \pi\eta^* + \theta^* - \pi\theta^*] = 0 \\ \Rightarrow a(1 - \eta^* - \theta^*) + n(\eta^* + \theta^* - 1)[\pi(1 - \eta^* - \theta^*) + \theta^*] &= 0 \\ \Rightarrow a(1 - \eta^* - \theta^*) + n\pi(\eta^* + \theta^* - 1)(1 - \eta^* - \theta^*) + n\theta^*(\eta^* + \theta^* - 1) &= 0 \\ \Rightarrow a(1 - \eta^* - \theta^*) + n\pi(\eta^* + \theta^* - 1)(1 - \eta^* - \theta^*) - n\theta^*(1 - \eta^* - \theta^*) &= 0 \\ \Rightarrow (1 - \eta^* - \theta^*)[a + n\pi(\eta^* + \theta^* - 1) - n\theta^*] = 0 \Rightarrow a + n\pi(\eta^* + \theta^* - 1) - n\theta^* &= 0 \\ \Rightarrow n\pi(\eta^* + \theta^* - 1) = -a + n\theta^* \Rightarrow \pi(\eta^* + \theta^* - 1) = -(a/n) + \theta^* \\ \Rightarrow -\pi(1 - \eta^* - \theta^*) = [-(a/n) + (1 - \theta)] \Rightarrow \hat{\pi} = \frac{[(a/n) - \theta^*]}{(1 - \eta^* - \theta^*)}\end{aligned}$$

Substituindo  $\eta^*$  e  $\theta^*$  por  $(1 - \eta)$  e  $(1 - \theta)$  tem-se que:

$$\hat{\pi} = \frac{[(a/n) - (1 - \theta)]}{(1 - (1 - \eta) - (1 - \theta))} = \frac{[(a/n) - (1 - \theta)]}{(1 - 1 + \eta - 1 + \theta)} = \frac{[(a/n) - (1 - \theta)]}{(\eta + \theta - 1)}.$$

Logo,

$$\hat{\pi} = \frac{[(a/n) - (1 - \theta)]}{(\eta + \theta - 1)}. \blacksquare$$

Geralmente  $\eta^* = (1 - \eta)$  e  $\theta^* = (1 - \theta)$  são desconhecidos e neste caso, deve-se considerar a função de verossimilhança como uma função de três parâmetros  $(\pi, \eta^*, \theta^*)$ . A estimativa de máxima verossimilhança de  $(\pi, \eta^*, \theta^*)$  não é única, isto é, diferentes combinações de valores entre  $\pi, \eta^*$  e  $\theta^*$  poderá conduzir ao mesmo valor de  $\tau$  e, conseqüentemente, a mesma função de verossimilhança. Aqui surge um problema de identificação dos parâmetros, e como resultado, a função de máxima verossimilhança atinge seu valor de máximo em todos os pontos  $(\pi, \eta^*, \theta^*)$ , tal que  $\pi(1 - \eta^*) + (1 - \pi)\theta^* = a/n$  (estimativa de  $\tau$ ). Então para evitar a não identificação dos parâmetros no modelo será considerado a restrição onde  $(1 - \eta^*) > \theta^*$ .

Esta restrição é realística, pois, deseja-se que a probabilidade de um indivíduo doente ser registrado corretamente como tal, seja maior, do que a probabilidade de um indivíduo saudável ser registrado erroneamente como doente.

Aplicando o Binômio de Newton na função de verossimilhança dada pela expressão (2.3) é possível reescrevê-la de forma mais complexa com o propósito de obter uma inferência bayesiana exata, dada pelo seguinte resultado:

**Resultado 2.2** - A função de verossimilhança de (2.3) pode ser dada por:

$$L(\pi, \eta^*, \theta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \pi^{n-j-t} (1 - \pi)^{j+t} (\eta^*)^{(b-t)} (1 - \eta^*)^{a-j} (\theta^*)^j (1 - \theta^*)^t \quad (2.5)$$

tal que  $(\pi, \eta^*, \theta^*) \in \{(\pi, \eta^*, \theta^*) : (1 - \eta^*) > \theta^*\}$ .

onde

$n$  = tamanho da amostra,

$j$  = n<sup>o</sup> de indivíduos sem a doença registrados erroneamente como doentes,

$t$  = n<sup>o</sup> de indivíduos sem a doença registrados corretamente como não doentes,

$a$  = n<sup>o</sup> de indivíduos com a doença,

$b$  = n<sup>o</sup> de indivíduos sem a doença.

**Demonstração:** Utilizando o Binômio de Newton,  $(a + b)^n = \sum_{j=0}^n \binom{n}{j} a^{n-j} b^j$ , tem-se que:

$$\begin{aligned}
 L(\pi, \eta^*, \theta^* | n, a, b) &= [\pi(1 - \eta^*) + (1 - \pi)\theta^*]^a [\pi\eta^* + (1 - \pi)(1 - \theta^*)]^b \\
 &= \sum_{j=0}^a \binom{a}{j} [\pi(1 - \eta^*)]^{a-j} [(1 - \pi)\theta^*]^j \sum_{t=0}^b \binom{b}{t} [\pi\eta^*]^{b-t} \cdot \\
 &\quad [(1 - \pi)(1 - \theta^*)]^t \\
 &= \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \pi^{n-j-t} (1 - \pi)^{j+t} (\eta^*)^{(b-t)} (1 - \eta^*)^{a-j} \cdot \\
 &\quad (\theta^*)^j (1 - \theta^*)^t.
 \end{aligned}$$

Assim,

$$L(\pi, \eta^*, \theta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \pi^{n-j-t} (1 - \pi)^{j+t} (\eta^*)^{(b-t)} (1 - \eta^*)^{a-j} (\theta^*)^j (1 - \theta^*)^t. \blacksquare$$

## 2.2 Análise Bayesiana

A razão do uso da análise bayesiana é explicada pelo fato de que a análise clássica apresenta algumas deficiências em relação a estimação da prevalência ( $\pi$ ), por exemplo, esta última análise às vezes fornece estimativas absurdas de prevalência tais como, estimativas que excedam 100% ou que estejam em território negativo. Além disso, na análise clássica é necessário conhecer as taxas de erros  $\eta^* = (1 - \eta)$  e  $\theta^* = (1 - \theta)$ , para que a estimativa da prevalência possa ser calculada, enquanto que na análise bayesiana não há esta necessidade.

A análise bayesiana é uma área da estatística que visa a estimação de quantidade desconhecida utilizando outras informações além da amostra. Estas informações são avaliadas antes dos dados amostrais e podem ser, por exemplo, a opinião de um especialista ou resultados de pesquisas anteriores. Estas informações consideradas subjetivas, são representadas, formalmente pelas distribuições *a priori*.

Seja  $T_1, T_2, \dots, T_n$  uma amostra aleatória de uma distribuição caracterizada por um parâmetro desconhecido  $\omega$ . O objetivo de qualquer procedimento de inferência é refletir o valor verdadeiro desconhecido de  $\omega$ . Logo, com este propósito determina-se  $f(\omega)$  como sendo a informação *a priori* para este parâmetro e expressa o conhecimento sobre o valor "verdadeiro" de  $\omega$  sobre o espaço paramétrico antes da observação da amostra. Os dados amostrais são representados pela função de verossimilhança  $L(\omega|T)$ . A informação combinada de  $\omega$  dessas duas origens é definida pela distribuição *a posteriori*. Formalmente, a função de probabilidade *a posteriori* é, segundo o Teorema de Bayes, definida como

$$f(\omega|T) = \frac{f(\omega)L(\omega|T)}{\int_0^{\infty} f(\omega)L(\omega|T)d\omega} \propto f(\omega)L(\omega|T)$$

onde o denominador é conhecido como constante de normalização.

A distribuição *a posteriori* é considerada uma descrição completa de  $\omega$  vindo da informação *a priori* e dos dados amostrais, descreve qual a probabilidade do verdadeiro valor do parâmetro  $\omega$  estar em diferentes partes do espaço paramétrico. Logo, o procedimento de inferência bayesiana baseia-se na distribuição *a posteriori* de  $\omega$ .

Então, para obter a distribuição *a posteriori* conjunta de  $\pi, \eta^*$  e  $\theta^*$ , foi sugerido uma densidade *a priori* proporcional a um produto de densidades Beta para os parâmetros  $\pi, \eta^*$  e  $\theta^*$ , baseando-se na função de verossimilhança dada em (2.5). A distribuição Beta é freqüentemente utilizada em modelos Bayesianos para interpretar as informações que envolvem proporções. Além disso, como a sensibilidade ( $\eta$ ) e a especificidade ( $\theta$ ) não são calculadas sobre os mesmos indivíduos, ou seja, no cálculo da sensibilidade, utiliza-se apenas os doentes, e no cálculo da especificidade, utiliza-se somente os não doentes, do ponto de vista estatístico, diz-se que estas medidas são entre si independentes, conseqüentemente as taxas de erros  $\eta^*$  e  $\theta^*$  também serão independentes. E ainda, a proporção de doentes (prevalência ( $\pi$ )) observada no estudo do desempenho do teste diagnóstico não interfere no cálculo destas medidas, o que permite afirmar que  $\eta$  e  $\theta$  não sofrem o efeito da prevalência ( $\pi$ ) da doença. Assim, tem-se que os três parâmetros ( $\pi, \eta^*, \theta^*$ ) são independentes entre si. Portanto, a densidade *a priori* para ( $\pi, \eta^*, \theta^*$ ) pode ser vista como um produto de densidades Beta:

$$f(\pi, \eta^*, \theta^*) = f(\pi|\alpha, \beta)f(\eta^*|\beta_1, \alpha_1)f(\theta^*|\beta_2, \alpha_2) \quad (2.6)$$

com  $\alpha, \beta, \beta_1, \alpha_1, \beta_2$  e  $\alpha_2 > 0$ , onde

$$f(\pi|\alpha, \beta) = \frac{\pi^{\alpha-1}(1-\pi)^{\beta-1}}{B(\alpha, \beta)}, \quad (2.7)$$

$$f(\eta^*|\beta_1, \alpha_1) = \frac{(\eta^*)^{\beta_1-1}(1-\eta^*)^{\alpha_1-1}}{B(\beta_1, \alpha_1)},$$

$$f(\theta^*|\beta_2, \alpha_2) = \frac{(\theta^*)^{\beta_2-1}(1-\theta^*)^{\alpha_2-1}}{B(\beta_2, \alpha_2)},$$

são as funções densidades Beta com parâmetros  $\alpha$  e  $\beta$  para  $\pi$ ,  $\beta_1$  e  $\alpha_1$  para  $\eta^*$ ,  $\beta_2$  e  $\alpha_2$  para  $\theta^*$ , e

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \quad B(\beta_1, \alpha_1) = \frac{\Gamma(\beta_1)\Gamma(\alpha_1)}{\Gamma(\beta_1 + \alpha_1)} \quad \text{e} \quad B(\beta_2, \alpha_2) = \frac{\Gamma(\beta_2)\Gamma(\alpha_2)}{\Gamma(\beta_2 + \alpha_2)},$$

são as constantes normalizadoras.

A explicação da escolha de  $\eta^* \sim \text{Beta}(\beta_1, \alpha_1)$  e  $\theta^* \sim \text{Beta}(\beta_2, \alpha_2)$  encontra-se com detalhes no Apêndice A.

Multiplicando a função de verossimilhança, dada em (2.5), pela densidade *a priori* dada em (2.6), e normalizando, como requer o teorema de Bayes, obtém-se a densidade conjunta *a posteriori* exata para  $(\pi, \eta^*, \theta^*)$ , dada pelo Resultado 2.3. Essa densidade conjunta *a posteriori* pode ser expressa como uma mistura de densidades, da mesma forma como em (2.6); mas com novos parâmetros para as densidades Beta.

**Resultado 2.3** - A densidade conjunta *a posteriori* para  $(\pi, \eta^*, \theta^*)$  é dada por:

$$f(\pi, \eta^*, \theta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi, \eta^*, \theta^* | n, a, b, j, t) I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)}, \quad (2.8)$$

onde

$$w_{jt} = \frac{a_{jt}}{\sum_{j=0}^a \sum_{t=0}^b a_{jt}},$$

$$a_{jt} = \binom{a}{j} \binom{b}{t} B(\alpha_{jt}^*, \beta_{jt}^*) B(\beta_{1t}^*, \alpha_{1j}^*) B(\beta_{2j}^*, \alpha_{2t}^*),$$

$$f(\pi, \eta^*, \theta^* | n, a, b, j, t) = f(\pi | \alpha_{jt}^*, \beta_{jt}^*) f(\eta^* | \beta_{1t}^*, \alpha_{1j}^*) f(\theta^* | \beta_{2j}^*, \alpha_{2t}^*), \quad (2.9)$$

$$\alpha_{jt}^* = \alpha + n - j - t,$$

$$\beta_{jt}^* = \beta + j + t,$$

$$\beta_{1t}^* = \beta_1 + b - t,$$

$$\alpha_{1j}^* = \alpha_1 + a - j,$$

$$\beta_{2j}^* = \beta_2 + j, \quad e$$

$$\alpha_{2t}^* = \alpha_2 + t.$$

**Demonstração:** Utilizando a Fórmula de Bayes, obtém-se:

$$f(\pi, \eta^*, \theta^* | n, a, b) = \frac{L(\pi, \eta^*, \theta^* | n, a, b) f(\pi, \eta^*, \theta^*)}{\int_0^1 \int_0^1 \int_0^1 L(\pi, \eta^*, \theta^* | n, a, b) f(\pi, \eta^*, \theta^*) d\pi d\eta^* d\theta^*}$$

onde

$$\begin{aligned} L(\pi, \eta^*, \theta^* | n, a, b) f(\pi, \eta^*, \theta^*) &= \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \pi^{n-j-t} (1-\pi)^{j+t} (\eta^*)^{(b-t)} \cdot \\ &\quad (1-\eta^*)^{a-j} (\theta^*)^j (1-\theta^*)^t f(\pi, \eta^*, \theta^*) \\ &= \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \pi^{n-j-t} (1-\pi)^{j+t} (\eta^*)^{(b-t)} \cdot \\ &\quad (1-\eta^*)^{a-j} (\theta^*)^j (1-\theta^*)^t \frac{\pi^{\alpha-1} (1-\pi)^{\beta-1}}{B(\alpha, \beta)} \cdot \\ &\quad \frac{\eta^{*(\beta_1-1)} (1-\eta^*)^{\alpha_1-1}}{B(\beta_1, \alpha_1)} \frac{\theta^{*(\beta_2-1)} (1-\theta^*)^{\alpha_2-1}}{B(\beta_2, \alpha_2)} \\ &= \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{\pi^{(\alpha+n-j-t)-1} (1-\pi)^{(\beta+j+t)-1}}{B(\alpha, \beta)}. \end{aligned}$$

$$\begin{aligned}
& \frac{(\eta^*)^{(\beta_1+b-t)-1}(1-\eta^*)^{(\alpha_1+a-j)-1}}{B(\beta_1, \alpha_1)} \frac{(\theta^*)^{(\beta_2+j)-1}(1-\theta^*)^{(\alpha_2+t)-1}}{B(\beta_2, \alpha_2)} \\
= & \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{\pi^{\alpha_{jt}^*-1}(1-\pi)^{\beta_{jt}^*-1}}{B(\alpha, \beta)} \frac{(\eta^*)^{\beta_{1t}^*-1}(1-\eta^*)^{\alpha_{1j}^*-1}}{B(\beta_1, \alpha_1)} \frac{(\theta^*)^{\beta_{2j}^*-1}(1-\theta^*)^{\alpha_{2t}^*-1}}{B(\beta_2, \alpha_2)} \\
= & \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{\pi^{\alpha_{jt}^*-1}(1-\pi)^{\beta_{jt}^*-1}}{B(\alpha_{jt}^*, \beta_{jt}^*)} \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} \frac{(\eta^*)^{\beta_{1t}^*-1}(1-\eta^*)^{\alpha_{1j}^*-1}}{B(\beta_{1t}^*, \alpha_{1j}^*)} \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} \cdot \\
& \frac{(\theta^*)^{\beta_{2j}^*-1}(1-\theta^*)^{\alpha_{2t}^*-1}}{B(\beta_{2j}^*, \alpha_{2t}^*)} \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)} \\
= & \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} f(\pi|\alpha_{jt}^*, \beta_{jt}^*) \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} f(\theta^*|\beta_{2j}^*, \alpha_{2t}^*) \cdot \\
& \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)} \\
= & \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)} f(\pi|\alpha_{jt}^*, \beta_{jt}^*) f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) \cdot \\
& f(\theta^*|\beta_{2j}^*, \alpha_{2t}^*).
\end{aligned}$$

Baseando-se em

$$f(\pi, \eta^*, \theta^*) = f(\pi|\alpha, \beta) f(\eta^*|\beta_1, \alpha_1) f(\theta^*|\beta_2, \alpha_2)$$

tem-se que

$$f(\pi, \eta^*, \theta^*|n, a, b, j, t) = f(\pi|\alpha_{jt}^*, \beta_{jt}^*) f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) f(\theta^*|\beta_{2j}^*, \alpha_{2t}^*)$$

Então,

$$L(\pi, \eta^*, \theta^* | n, a, b) f(\pi, \eta^*, \theta^*) = \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)}.$$

$$f(\pi, \eta^*, \theta^* | n, a, b, j, t)$$

e

$$\begin{aligned} \int_0^1 \int_0^1 \int_0^1 L(\pi, \eta^*, \theta^* | n, a, b) f(\pi, \eta^*, \theta^*) d\pi d\eta^* d\theta^* &= \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} \cdot \\ &\quad \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)} \cdot \\ &\quad \underbrace{\int_0^1 \int_0^1 \int_0^1 f(\pi, \eta^*, \theta^* | n, a, b, j, t) d\pi d\eta^* d\theta^*}_{= 1} \\ &= \sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} \cdot \\ &\quad \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)} \end{aligned}$$

Assim,

$$f(\pi, \eta^*, \theta^* | n, a, b) = \frac{\sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)} f(\pi, \eta^*, \theta^* | n, a, b, j, t)}{\sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} \frac{B(\alpha_{jt}^*, \beta_{jt}^*)}{B(\alpha, \beta)} \frac{B(\beta_{1t}^*, \alpha_{1j}^*)}{B(\beta_1, \alpha_1)} \frac{B(\beta_{2j}^*, \alpha_{2t}^*)}{B(\beta_2, \alpha_2)}}$$

$$\begin{aligned}
&= \sum_{j=0}^a \sum_{t=0}^b \frac{\binom{a}{j} \binom{b}{t} B(\alpha_{jt}^*, \beta_{jt}^*) B(\beta_{1t}^*, \alpha_{1j}^*) B(\beta_{2j}^*, \alpha_{2t}^*) f(\pi, \eta^*, \theta^* | n, a, b, j, t)}{\sum_{j=0}^a \sum_{t=0}^b \binom{a}{j} \binom{b}{t} B(\alpha_{jt}^*, \beta_{jt}^*) B(\beta_{1t}^*, \alpha_{1j}^*) B(\beta_{2j}^*, \alpha_{2t}^*)} \\
&= \sum_{j=0}^a \sum_{t=0}^b \frac{a_{jt}}{\sum_{j=0}^a \sum_{t=0}^b a_{jt}} f(\pi, \eta^*, \theta^* | n, a, b, j, t) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi, \eta^*, \theta^* | n, a, b, j, t).
\end{aligned}$$

Logo, baseando-se em (2.9), obtém-se:

$$f(\pi, \eta^*, \theta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi, \eta^*, \theta^* | n, a, b, j, t) I_{\substack{(\pi, \eta^*, \theta^*) \\ [(1-\eta^*) > \theta^*]}}. \blacksquare$$

Considerando que o objetivo principal é fazer inferência sobre  $\pi$ , pode-se obter a distribuição marginal *a posteriori* exata de  $\pi$  da seguinte forma. Como a distribuição *a priori* de  $\pi$  é independente das taxas de erros, tem-se o seguinte resultado:

**Resultado 2.4** - A distribuição marginal *a posteriori* exata para  $\pi$  é dada por:

$$f(\pi | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi | \alpha_{jt}^*, \beta_{jt}^*). \quad (2.10)$$

**Demonstração:** Integrando a distribuição *a posteriori* conjunta  $f(\pi, \eta^*, \theta^* | n, a, b)$  em relação a  $\eta^*$  e  $\theta^*$ , tem-se:

$$\begin{aligned}
f(\pi | n, a, b) &= \int_0^1 \int_0^1 f(\pi, \eta^*, \theta^* | n, a, b) d\eta^* d\theta^* \\
&= \int_0^1 \int_0^1 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi, \eta^*, \theta^* | n, a, b, j, t) d\eta^* d\theta^* \\
&= \int_0^1 \int_0^1 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi | \alpha_{jt}^*, \beta_{jt}^*) f(\eta^* | \beta_{1t}^*, \alpha_{1j}^*) f(\theta^* | \beta_{2j}^*, \alpha_{2t}^*) d\eta^* d\theta^*
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi | \alpha_{jt}^*, \beta_{jt}^*) \int_0^1 \int_0^1 f(\eta^* | \beta_{1t}^*, \alpha_{1j}^*) f(\theta^* | \beta_{2j}^*, \alpha_{2t}^*) d\eta^* d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi | \alpha_{jt}^*, \beta_{jt}^*) \int_0^1 \int_0^1 f(\eta^* | \beta_{1t}^*, \alpha_{1j}^*) d\eta^* \int_0^1 f(\theta^* | \beta_{2j}^*, \alpha_{2t}^*) d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi | \alpha_{jt}^*, \beta_{jt}^*) \underbrace{\int_0^1 \frac{(\eta^*)^{\beta_{1t}^* - 1} (1 - \eta^*)^{\alpha_{1j}^* - 1}}{B(\beta_{1t}^*, \alpha_{1j}^*)} d\eta^*}_{= 1} \\
&\quad \underbrace{\int_0^1 \frac{(\theta^*)^{\beta_{2j}^* - 1} (1 - \theta^*)^{\alpha_{2t}^* - 1}}{B(\beta_{2j}^*, \alpha_{2t}^*)} d\theta^*}_{= 1} = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi | \alpha_{jt}^*, \beta_{jt}^*).
\end{aligned}$$

Como  $\pi$  é independente de  $\eta^*$  e  $\theta^*$ , encontra-se que:

$$f(\pi | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi | \alpha_{jt}^*, \beta_{jt}^*). \blacksquare$$

Considerando as propriedades da distribuição Beta, a média e a variância *a posteriori* de  $\pi$  são expressas pelo seguinte resultado:

**Resultado 2.5** - A média *a posteriori* e a variância *a posteriori* para  $\pi$  são dadas por:

$$E(\pi | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t)}{(\alpha + \beta + n)} \quad \text{e} \quad (2.11)$$

$$V(\pi | n, a, b) = E(\pi^2 | n, a, b) - E^2(\pi | n, a, b), \text{ onde}$$

$$E(\pi^2 | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t + 1)(\alpha + n - j - t)}{(\alpha + \beta + n + 1)(\alpha + \beta + n)}. \quad (2.12)$$

**Demonstração:** Considerando as propriedades de esperança matemática, distribuição Beta e distribuição Gama, obtém-se:

$$\begin{aligned}
E(\pi|n, a, b) &= \int_0^1 \pi f(\pi|n, a, b) d\pi = \int_0^1 \pi \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi|\alpha_{jt}^*, \beta_{jt}^*) d\pi \\
&= \int_0^1 \pi \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi|\alpha + n - j - t, \beta + j + t) d\pi \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \int_0^1 \pi \frac{\pi^{(\alpha+n-j-t)-1} (1-\pi)^{(\beta+j+t)-1}}{B(\alpha + n - j - t, \beta + j + t)} d\pi \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\alpha + n - j - t, \beta + j + t)} \int_0^1 \pi^{(\alpha+n-j-t)} (1-\pi)^{(\beta+j+t)-1} d\pi \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{B(\alpha + n - j - t + 1, \beta + j + t)}{B(\alpha + n - j - t, \beta + j + t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{\Gamma(\alpha + n - j - t + 1) \Gamma(\alpha + \beta + n)}{\Gamma(\alpha + \beta + n + 1) \Gamma(\alpha + n - j - t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t)! (\alpha + \beta + n - 1)!}{(\alpha + \beta + n)! (\alpha + n - j - t - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t) (\alpha + n - j - t - 1)! (\alpha + \beta + n - 1)!}{(\alpha + \beta + n) (\alpha + \beta + n - 1)! (\alpha + n - j - t - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t)}{(\alpha + \beta + n)}.
\end{aligned}$$

Logo, tem-se que:

$$E(\pi|n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t)}{(\alpha + \beta + n)}. \blacksquare$$

Utilizando o mesmo procedimento,  $E(\pi^2|n, a, b)$  pode ser calculado como segue:

$$\begin{aligned}
E(\pi^2|n, a, b) &= \int_0^1 \pi^2 f(\pi|n, a, b) d\pi = \int_0^1 \pi^2 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi|\alpha_{jt}^*, \beta_{jt}^*) d\pi \\
&= \int_0^1 \pi^2 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi|\alpha + n - j - t, \beta + j + t) d\pi \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \int_0^1 \pi^2 \frac{\pi^{(\alpha+n-j-t)-1} (1-\pi)^{(\beta+j+t)-1}}{B(\alpha + n - j - t, \beta + j + t)} d\pi \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\alpha+n-j-t, \beta+j+t)} \int_0^1 \pi^{(\alpha+n-j-t+1)} (1-\pi)^{(\beta+j+t)-1} d\pi \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\alpha+n-j-t, \beta+j+t)} \int_0^1 \pi^{(\alpha+n-j-t+1+1)-1} (1-\pi)^{(\beta+j+t)-1} d\pi \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{B(\alpha + n - j - t + 2, \beta + j + t)}{B(\alpha + n - j - t, \beta + j + t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{\Gamma(\alpha + n - j - t + 2) \Gamma(\alpha + \beta + n)}{\Gamma(\alpha + \beta + n + 2) \Gamma(\alpha + n - j - t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t + 1)! (\alpha + \beta + n - 1)!}{(\alpha + \beta + n + 1)! (\alpha + n - j - t - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha+n-j-t+1)(\alpha+n-j-t)(\alpha+n-j-t-1)(\alpha+\beta+n-1)!}{(\alpha+\beta+n+1)(\alpha+\beta+n)(\alpha+\beta+n-1)(\alpha+n-j-t-1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha + n - j - t + 1)(\alpha + n - j - t)}{(\alpha + \beta + n + 1)(\alpha + \beta + n)}.
\end{aligned}$$

Assim, tem-se que:

$$\begin{aligned} V(\pi|n, a, b) &= E(\pi^2|n, a, b) - E^2(\pi|n, a, b) \\ &= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha+n-j-t+1)(\alpha+n-j-t)}{(\alpha+\beta+n+1)(\alpha+\beta+n)} - \left[ \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\alpha+n-j-t)}{(\alpha+\beta+n)} \right]^2. \blacksquare \end{aligned}$$

Pode-se também fazer inferência sobre  $\eta^*$  e  $\theta^*$ . Suas densidades marginais *a posteriori* exatas podem ser obtidas integrando a densidade conjunta *a posteriori* adequadamente. Como a distribuição *a priori* de  $\eta^*$  é independente de  $\pi$  e  $\theta^*$ , tem-se o seguinte resultado:

**Resultado 2.6** - A distribuição marginal *a posteriori* exata para  $\eta^*$  é dada por:

$$f(\eta^*|n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) I_{[\eta^* < (1-\theta^*)]} \quad (2.13)$$

**Demonstração:** Integrando a distribuição *a posteriori* conjunta  $f(\pi, \eta^*, \theta^*|n, a, b)$  em relação a  $\pi$  e  $\theta^*$ , tem-se:

$$\begin{aligned} f(\eta^*|n, a, b) &= \int_0^1 \int_0^1 f(\pi, \eta^*, \theta^*|n, a, b) d\pi d\theta^* \\ &= \int_0^1 \int_0^1 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi, \eta^*, \theta^*|n, a, b, j, t) d\pi d\theta^* \\ &= \int_0^1 \int_0^1 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi|\alpha_{jt}^*, \beta_{jt}^*) f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) f(\theta^*|\beta_{2j}^*, \alpha_{2t}^*) d\pi d\theta^* \\ &= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) \int_0^1 f(\pi|\alpha_{jt}^*, \beta_{jt}^*) d\pi \int_0^1 f(\theta^*|\beta_{2j}^*, \alpha_{2t}^*) d\theta^* \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^* | \beta_{1t}^*, \alpha_{1j}^*) \underbrace{\int_0^1 \frac{(\pi)^{\alpha_{jt}^* - 1} (1 - \pi)^{\beta_{jt}^* - 1}}{B(\alpha_{jt}^*, \beta_{jt}^*)} d\pi}_{= 1} \\
&\underbrace{\int_0^1 \frac{(\theta^*)^{\beta_{2j}^* - 1} (1 - \theta^*)^{\alpha_{2t}^* - 1}}{B(\beta_{2j}^*, \alpha_{2t}^*)} d\theta^*}_{= 1} = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^* | \beta_{1t}^*, \alpha_{1j}^*).
\end{aligned}$$

Como  $\eta^*$  é independente de  $\pi$  e  $\theta^*$ , encontra-se que:

$$f(\eta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^* | \beta_{1t}^*, \alpha_{1j}^*) I_{[\eta^* < (1 - \theta^*)]}^{(\eta^*)}. \blacksquare$$

Considerando as propriedades da distribuição Beta, a média e a variância *a posteriori* de  $\eta^*$  são expressas pelo seguinte resultado:

**Resultado 2.7** - A média *a posteriori* e a variância *a posteriori* para  $\eta^*$  são dadas por:

$$E(\eta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t)}{(\beta_1 + \alpha_1 + n - t - j)} \quad \text{e} \quad (2.14)$$

$$V(\eta^* | n, a, b) = E((\eta^*)^2 | n, a, b) - E^2(\eta^* | n, a, b), \text{ onde}$$

$$E((\eta^*)^2 | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t + 1)(\beta_1 + b - t)}{(\beta_1 + \alpha_1 + n - t - j + 1)(\beta_1 + \alpha_1 + n - t - j)}. \quad (2.15)$$

**Demonstração:** Considerando as propriedades de esperança matemática, distribuição Beta e distribuição Gama, obtém-se:

$$\begin{aligned}
E(\eta^*|n, a, b) &= \int_0^1 \eta^* f(\eta^*|n, a, b) d\eta^* = \int_0^1 \eta^* \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) d\eta^* \\
&= \int_0^1 \eta^* \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^*|\beta_1 + b - t, \alpha_1 + a - j) d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \int_0^1 \eta^{*(\beta_1 + b - t) - 1} (1 - \eta^*)^{(\alpha_1 + a - j) - 1} \frac{1}{B(\beta_1 + b - t, \alpha_1 + a - j)} d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\beta_1 + b - t, \alpha_1 + a - j)} \int_0^1 \eta^{*(\beta_1 + b - t)} (1 - \eta^*)^{(\alpha_1 + a - j) - 1} d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{B(\beta_1 + b - t + 1, \alpha_1 + a - j)}{B(\beta_1 + b - t, \alpha_1 + a - j)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{\Gamma(\beta_1 + b - t + 1) \Gamma(\beta_1 + \alpha_1 + n - t - j)}{\Gamma(\beta_1 + \alpha_1 + n - t - j + 1) \Gamma(\beta_1 + b - t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t)! (\beta_1 + \alpha_1 + n - t - j - 1)!}{(\beta_1 + \alpha_1 + n - t - j)! (\beta_1 + b - t - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t) (\beta_1 + b - t - 1)! (\beta_1 + \alpha_1 + n - t - j - 1)!}{(\beta_1 + \alpha_1 + n - t - j) (\beta_1 + \alpha_1 + n - t - j - 1)! (\beta_1 + b - t - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t)}{(\beta_1 + \alpha_1 + n - t - j)}.
\end{aligned}$$

Logo, tem-se que:

$$E(\eta^*|n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t)}{(\beta_1 + \alpha_1 + n - t - j)}. \blacksquare$$

Utilizando o mesmo procedimento,  $E((\eta^*)^2|n, a, b)$  pode ser calculado como segue:

$$\begin{aligned}
E((\eta^*)^2|n, a, b) &= \int_0^1 (\eta^*)^2 f(\eta^*|n, a, b) d\eta^* = \int_0^1 (\eta^*)^2 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) d\eta^* \\
&= \int_0^1 (\eta^*)^2 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\eta^*|\beta_1 + b - t, \alpha_1 + a - j) d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \int_0^1 (\eta^*)^2 \frac{\eta^{*(\beta_1+b-t)-1} (1-\eta^*)^{(\alpha_1+a-j)-1}}{B(\beta_1 + b - t, \alpha_1 + a - j)} d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\beta_1+b-t, \alpha_1+a-j)} \eta^{*(\beta_1+b-t+1)} (1-\eta^*)^{(\alpha_1+a-j)-1} d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\beta_1+b-t, \alpha_1+a-j)} \int_0^1 \eta^{*(\beta_1+b-t+1+1)-1} (1-\eta^*)^{(\alpha_1+a-j)-1} d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{B(\beta_1 + b - t + 2, \alpha_1 + a - j)}{B(\beta_1 + b - t, \alpha_1 + a - j)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{\Gamma(\beta_1 + b - t + 2) \Gamma(\beta_1 + \alpha_1 + n - t - j)}{\Gamma(\beta_1 + \alpha_1 + n - t - j + 2) \Gamma(\beta_1 + b - t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t + 1)! (\beta_1 + \alpha_1 + n - t - j - 1)!}{(\beta_1 + \alpha_1 + n - t - j + 1)! (\beta_1 + b - t - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1+b-t+1)(\beta_1+b-t)(\beta_1+b-t-1)!(\beta_1+\alpha_1+n-t-j-1)!}{(\beta_1+\alpha_1+n-t-j+1)(\beta_1+\alpha_1+n-t-j)(\beta_1+\alpha_1+n-t-j-1)!(\beta_1+b-t-1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1 + b - t + 1)(\beta_1 + b - t)}{(\beta_1 + \alpha_1 + n - t - j + 1)(\beta_1 + \alpha_1 + n - t - j)}.
\end{aligned}$$

Assim, tem-se que:

$$\begin{aligned}
V(\eta^*|n, a, b) &= E((\eta^*)^2|n, a, b) - E^2(\eta^*|n, a, b) \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1+b-t+1)(\beta_1+b-t)}{(\beta_1+\alpha_1+n-t-j+1)(\beta_1+\alpha_1+n-t-j)} - \left[ \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_1+b-t)}{(\beta_1+\alpha_1+n-t-j)} \right]^2. \blacksquare
\end{aligned}$$

E como a distribuição *a priori* de  $\theta^*$  é independente de  $\pi$  e  $\eta^*$ , obtém-se o seguinte resultado:

**Resultado 2.8** - A distribuição marginal *a posteriori* exata para  $\theta^*$  é dada por:

$$f(\theta^*|n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^*|\beta_{2t}^*, \alpha_{2j}^*) I_{[\theta^* < (1-\eta^*)]}^{(\theta^*)}. \quad (2.16)$$

**Demonstração:** Integrando a distribuição *a posteriori* conjunta  $f(\pi, \eta^*, \theta^*|n, a, b)$  em relação a  $\pi$  e  $\eta^*$ , tem-se:

$$\begin{aligned}
f(\theta^*|n, a, b) &= \int_0^1 \int_0^1 f(\pi, \eta^*, \theta^*|n, a, b) d\pi d\eta^* \\
&= \int_0^1 \int_0^1 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi, \eta^*, \theta^*|n, a, b, j, t) d\pi d\eta^* \\
&= \int_0^1 \int_0^1 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\pi|\alpha_{jt}^*, \beta_{jt}^*) f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) f(\theta^*|\beta_{2t}^*, \alpha_{2j}^*) d\pi d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^*|\beta_{2t}^*, \alpha_{2j}^*) \int_0^1 \int_0^1 f(\pi|\alpha_{jt}^*, \beta_{jt}^*) f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) d\pi d\eta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^*|\beta_{2t}^*, \alpha_{2j}^*) \int_0^1 f(\pi|\alpha_{jt}^*, \beta_{jt}^*) d\pi \int_0^1 f(\eta^*|\beta_{1t}^*, \alpha_{1j}^*) d\eta^*
\end{aligned}$$

$$\begin{aligned}
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^* | \beta_{2t}^*, \alpha_{2j}^*) \underbrace{\int_0^1 \frac{(\pi)^{\alpha_{jt}^* - 1} (1 - \pi)^{\beta_{jt}^* - 1}}{B(\alpha_{jt}^*, \beta_{jt}^*)} d\pi}_{= 1} \\
&\underbrace{\int_0^1 \frac{(\eta^*)^{\beta_{1j}^* - 1} (1 - \eta^*)^{\alpha_{1t}^* - 1}}{B(\beta_{1j}^*, \alpha_{1t}^*)} d\eta^*}_{= 1} = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^* | \beta_{2t}^*, \alpha_{2j}^*).
\end{aligned}$$

Considerando que  $\theta^*$  é independente de  $\pi$  e  $\eta^*$ , encontra-se que:

$$f(\theta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^* | \beta_{2t}^*, \alpha_{2j}^*) I_{[\theta^* < (1 - \eta^*)]}^{(\theta^*)}. \blacksquare$$

Baseando-se nas propriedades da distribuição Beta, a média e a variância *a posteriori* de  $\theta^*$  são expressas pelo seguinte resultado:

**Resultado 2.9** - A média *a posteriori* e a variância *a posteriori* para  $\theta^*$  são dadas por:

$$E(\theta^* | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j)}{(\beta_2 + \alpha_2 + j + t)} \quad \text{e} \quad (2.17)$$

$$V(\theta^* | n, a, b) = E((\theta^*)^2 | n, a, b) - E^2(\theta^* | n, a, b), \text{ onde}$$

$$E((\theta^*)^2 | n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j + 1)(\beta_2 + j)}{(\beta_2 + \alpha_2 + j + t + 1)(\beta_2 + \alpha_2 + j + t)}. \quad (2.18)$$

**Demonstração:** Considerando as propriedades de esperança matemática, distribuição Beta e distribuição Gama, obtém-se:

$$\begin{aligned}
E(\theta^*|n, a, b) &= \int_0^1 \theta^* f(\theta^*|n, a, b) d\theta^* = \int_0^1 \theta^* \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^*|\beta_{2t}^*, \alpha_{2j}^*) d\theta^* \\
&= \int_0^1 \theta^* \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^*|\beta_2 + j, \alpha_2 + t) d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \int_0^1 \theta^{*(\beta_2+j)-1} (1 - \theta^*)^{(\alpha_2+t)-1} \frac{1}{B(\beta_2 + j, \alpha_2 + t)} d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\beta_2 + j, \alpha_2 + t)} \int_0^1 \theta^{*(\beta_2+j)} (1 - \theta^*)^{(\alpha_2+t)-1} d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{B(\beta_2 + j + 1, \alpha_2 + t)}{B(\beta_2 + j, \alpha_2 + t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{\Gamma(\beta_2 + j + 1)\Gamma(\beta_2 + \alpha_2 + j + t)}{\Gamma(\beta_2 + \alpha_2 + j + t + 1)\Gamma(\beta_2 + j)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j)!(\beta_2 + \alpha_2 + j + t - 1)!}{(\beta_2 + \alpha_2 + j + t)!(\beta_2 + j - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j)(\beta_2 + j - 1)!(\beta_2 + \alpha_2 + j + t - 1)!}{(\beta_2 + \alpha_2 + j + t)(\beta_2 + \alpha_2 + j + t - 1)!(\beta_2 + j - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j)}{(\beta_2 + \alpha_2 + j + t)}.
\end{aligned}$$

Logo, tem-se que:

$$E(\theta^*|n, a, b) = \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j)}{(\beta_2 + \alpha_2 + j + t)}. \blacksquare$$

Utilizando o mesmo procedimento,  $E((\theta^*)^2|n, a, b)$  pode ser calculado como segue:

$$\begin{aligned}
E((\theta^*)^2|n, a, b) &= \int_0^1 (\theta^*)^2 f(\theta^*|n, a, b) d\theta^* = \int_0^1 (\theta^*)^2 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^*|\beta_{2t}^*, \alpha_{2j}^*) d\theta^* \\
&= \int_0^1 (\theta^*)^2 \sum_{j=0}^a \sum_{t=0}^b w_{jt} f(\theta^*|\beta_2 + j, \alpha_2 + t) d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \int_0^1 (\theta^*)^2 \frac{\theta^{*(\beta_2+j)-1} (1 - \theta^*)^{(\alpha_2+t)-1}}{B(\beta_2 + j, \alpha_2 + t)} d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\beta_2+j, \alpha_2+t)} \theta^{*(\beta_2+j+1)} (1 - \theta^*)^{(\alpha_2+t)-1} d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{1}{B(\beta_2+j, \alpha_2+t)} \int_0^1 \theta^{*(\beta_2+j+1+1)-1} (1 - \theta^*)^{(\alpha_2+t)-1} d\theta^* \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{B(\beta_2 + j + 2, \alpha_2 + t)}{B(\beta_2 + j, \alpha_2 + t)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{\Gamma(\beta_2 + j + 2) \Gamma(\beta_2 + \alpha_2 + j + t)}{\Gamma(\beta_2 + \alpha_2 + j + t + 2) \Gamma(\beta_2 + j)} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j + 1)! (\beta_2 + \alpha_2 + j + t - 1)!}{(\beta_2 + \alpha_2 + j + t + 1)! (\beta_2 + j - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j + 1) (\beta_2 + j) (\beta_2 + j - 1)! (\beta_2 + \alpha_2 + j + t - 1)!}{(\beta_2 + \alpha_2 + j + t + 1) (\beta_2 + \alpha_2 + j + t) (\beta_2 + \alpha_2 + j + t - 1)! (\beta_2 + j - 1)!} \\
&= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2 + j + 1) (\beta_2 + j)}{(\beta_2 + \alpha_2 + j + t + 1) (\beta_2 + \alpha_2 + j + t)}.
\end{aligned}$$

Assim, tem-se que:

$$\begin{aligned} V(\theta^*|n, a, b) &= E((\theta^*)^2|n, a, b) - E^2(\theta^*|n, a, b) \\ &= \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2+j+1)(\beta_2+j)}{(\beta_2+\alpha_2+j+t+1)(\beta_2+\alpha_2+j+t)} - \left[ \sum_{j=0}^a \sum_{t=0}^b w_{jt} \frac{(\beta_2+j)}{(\beta_2+\alpha_2+j+t)} \right]^2. \blacksquare \end{aligned}$$

Embora as marginais e as médias *a posteriori* de  $\pi, \eta^*$  e  $\theta^*$  possam ser obtidas de forma analítica, são expressões complexas, justificando assim, o uso do *Gibbs Sampling*.

## 2.3 Distribuições *a priori*

A distribuição *a priori* é o único elemento novo na análise bayesiana em relação à análise clássica, e é também o ponto mais crítico e o mais criticado pelos frequentistas. Esta distribuição tem por intuito representar o conhecimento sobre as quantidades  $\pi, \eta^*$  e  $\theta^*$  desconhecidas, antes de realizar o experimento.

Quando não se tem nenhuma informação adicional além dos dados, uma escolha comum é considerar uma *priori* difusa ou “não informativa” para  $(\pi, \eta^*, \theta^*)$ , como por exemplo, uma densidade uniforme, isto é,  $f(\pi, \eta^*, \theta^*) = 1$ . O problema é que as densidades *a priori* não informativas são geralmente impróprias, não garantindo a existência da *posteriori*.

A distribuição *a priori* informativa, em geral será: (1) em princípio subjetiva; (2) obtida através de especialistas; (3) obtida a partir de dados anteriores ao experimento principal; (4) obtidas por métodos objetivos; (5) obtidas por conjugação.

Neste trabalho, nas aplicações para os dados reais sobre a infecção do parasita *strongyloides* e obesidade em escolares, serão consideradas *prioris* não informativa e informativa, onde esta última será obtida através de especialistas. O procedimento para a obtenção destas *prioris* informativas se encontra na Subseção 4.1.3 do Capítulo 4.

# Capítulo 3

## Inferência Bayesiana para Dados Dicotômicos com Erros de Classificação

Este capítulo é organizado da seguinte maneira. Na Seção 3.1 será apresentado um breve histórico sobre Misturas de Modelos, na Seção 3.2 definições sobre o Modelo Misto são consideradas baseando-se em Diebolt & Robert (1994). Na Subseção 3.2.1 da Seção 3.2, será apresentado o algoritmo *Gibbs Sampling* para amostra ampliada.

Na Seção 3.3, segue o modelo para dados dicotômicos, considerando os dois tipos de erros de classificação  $\eta^* = (1 - \eta)$  e  $\theta^* = (1 - \theta)$  definidos no Capítulo 2. A construção deste modelo foi baseada na Tese de Mestrado de Vanda Donizetti Redondo Silveira, citada no capítulo anterior, e no artigo de Diebolt & Robert (1994). Como o trabalho dela envolve dados dicotômicos, o objetivo foi utilizar a teoria desenvolvida, aplicando-a ao teste diagnóstico. A Subseção 3.3.1 da Seção 3.3, abordará o algoritmo *Gibbs Sampling* com variável latente para o modelo com erros de classificação.

A idéia inicial era desenvolver um modelo misto para teste diagnóstico com erros de classificação, e um algoritmo *Gibbs Sampling* com amostra ampliada para obter os

resumos *a posteriori* dos parâmetros de interesse, a partir das duas referências citadas anteriormente, mas no decorrer deste trabalho, foi encontrado o artigo de Joseph *et al.* (1995) que modificou o andamento do estudo.

No procedimento para o modelo misto que será apresentado na Seção 3.3, é necessário conhecer o resultado de cada teste realizado, enquanto que no procedimento de Joseph *et al.* (1995) que será desenvolvido no Capítulo 4, tem-se apenas a informação sobre o total de testes positivos e negativos, pois destes totais é possível obter as estatísticas *a posteriori* para os parâmetros de interesse.

Como para as duas aplicações de dados reais neste trabalho, são fornecidos apenas os totais de testes positivos e negativos, o algoritmo *Gibbs Sampling* para o modelo com erros de classificação desenvolvido na Subseção 3.3.1 se modificará. Esta modificação será mostrada na Seção 3.4 ao considerar os totais de testes positivos e negativos no algoritmo *Gibbs Sampling* para o modelo misto, como feito no algoritmo *Gibbs Sampling* obtido pelo procedimento de Joseph *et al.* (1995).

É importante ressaltar que no algoritmo para o modelo misto descrito no presente capítulo, utiliza-se uma única variável latente para classificar um indivíduo como doente ou não, e no algoritmo obtido baseando-se no procedimento de Joseph *et al.* (1995), considerando apenas um teste diagnóstico, duas variáveis latentes foram usadas para esta mesma classificação. Mas, devido a consideração dos totais de testes positivos e negativos os algoritmo *Gibbs Sampling* para o modelo com erros de classificação e o obtido pelo procedimento de Joseph *et al.* (1995) mostraram-se equivalentes. Com isso, a restrição  $(1 - \eta^*) > \theta^*$  ou  $\eta > (1 - \theta)$  imposta no modelo misto, se aplica a metodologia de Joseph. Porém, na abordagem de Joseph *et al.* (1995) nenhuma restrição foi imposta, supõe-se que o motivo seja a naturalidade da ocorrência desta restrição, visto que um teste diagnóstico que não seja padrão ouro é aplicado somente quando a taxa de acerto for maior do que a taxa de erro. Então, as estatísticas *a posteriori* dos parâmetros serão obtidas através da implementação do algoritmo *Gibbs Sampling* delineado pelo procedimento de Joseph *et al.* (1995) sem nenhuma restrição. Estas estatísticas se encontram nos Capítulos 4 e 5,

devido as aplicações sobre a infecção do parasita *strongyloides* e sobre dados de obesidade, respectivamente.

### 3.1 Mistura de Modelos

Misturas de modelos podem ser utilizadas em diferentes situações. Por exemplo, quando numa população existem diferentes grupos de observações, justificando a suposição de diferentes distribuições para os indivíduos dentro de cada grupo. Neste caso, o objetivo seria discriminar os diferentes grupos. Outra envolve a identificação de observações discrepantes, possibilitando a suposição de distribuições com mesma média, mas diferentes variabilidades. Além destas, uma outra abordagem é considerar médias ponderadas de modelos. O objetivo, neste caso, seria reduzir a incerteza na escolha de um particular modelo, considerando diferentes formas funcionais em potencial, com suas respectivas chances *a posteriori*. O propósito neste capítulo envolve o primeiro aspecto.

### 3.2 Modelo Misto

Considerando uma amostra de tamanho  $n$  de uma variável aleatória  $T = (T_1, T_2, \dots, T_n)$ , e supondo que esta amostra é obtida de uma população formada por  $k$  componentes, com densidades  $f(t_i|\omega_j)$ ,  $j = 1, \dots, k$ , segundo Diebolt & Robert (1994), o modelo misto é dado por:

$$f(t_i|\omega, p) = \sum_{j=1}^k p_j f_j(t_i|\omega_j), \quad i = 1, \dots, n. \quad (3.1)$$

onde  $f_j(t_i|\omega_j)$ , ( $1 \leq j \leq k$ ) são as densidades conhecidas das componentes que envolvem a mistura, os parâmetros  $p_1, \dots, p_k$  são chamados de pesos, onde  $0 < p_j < 1$ , satisfazendo a restrição  $\sum_{j=1}^k p_j = 1$ , e o vetor  $\omega = (\omega_1, \omega_2, \dots, \omega_k)$  denota os parâmetros desconhecidos presentes em  $f_j(t_i|\omega_j)$ .

A verossimilhança dos parâmetros  $(\boldsymbol{\omega}, \mathbf{p})$ , pode ser construída através do modelo (3.1).

Dada por:

$$L(\boldsymbol{\omega}, \mathbf{p}|\mathbf{t}) = \prod_{i=1}^n f(t_i|\boldsymbol{\omega}, \mathbf{p}) = \prod_{i=1}^n \sum_{j=1}^k p_j f_j(t_i|\omega_j) \quad (3.2)$$

Existem três pontos importantes definindo os parâmetros de interesse. O primeiro ponto é na modelagem dos dados, pois há a necessidade da especificação do número de componentes  $k$ . Estes parâmetros podem ser identificados por uma densidade *a priori* de acordo com a ocasião. Quando o número de componentes  $k$  é desconhecido, o espaço paramétrico é mal definido e com dimensionalidade infinita, comprometendo o uso da densidade *a priori*.

O segundo ponto está na estimação dos parâmetros das componentes  $p_1, p_2, p_3, \dots, p_k$ , e dos parâmetros  $\omega_1, \omega_2, \omega_3, \dots, \omega_k$ .

O terceiro ponto é a questão da classificação, isto é, a determinação do componente para o qual cada  $t_i$ ,  $i = 1, \dots, n$  pertence. Esta classificação, é explicada por Diebolt & Robert (1994) como uma estrutura oculta do modelo, que pode ser vista como "dados incompletos", ou seja, cada observação é associada como um indicador não-observado do componente pelo qual é originada.

Os objetivos são obter informações sobre  $\boldsymbol{\omega}$  baseando-se na amostra  $T_1, T_2, \dots, T_n$ , e tentar classificar cada observação como pertencendo a cada uma das  $k$  componentes.

Segundo Dempster *et al.* (1977), o modelo misto (3.1), pode ser escrito em termos de dados incompletos. Por isso, será introduzida uma variável  $Z_i$  com dimensionalidade  $k$ , indicando a componente do modelo misto da qual a observação  $t_i$  é proveniente, tal que  $Z_{ij} \in \{0, 1\}$  e  $\sum_{j=1}^k Z_{ij} = 1$ , e

$$p_{ij} = \Pr(Z_{ij} = 1|\boldsymbol{\omega}, \mathbf{p}, \mathbf{T}) = \frac{p_j f_j(t_i|\omega_j)}{\sum_{j=1}^k p_j f_j(t_i|\omega_j)}, \quad (3.3)$$

onde os  $p_{ij}$  são os pesos das distribuições multinomiais de  $Z_i$ . Isto é,  $Z_i \sim Multinomial(1 : p_{i1}, \dots, p_{ik})$ , onde  $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{ik})$ . A introdução desta variável latente simplifica a

obtenção da distribuição *a posteriori*.

A distribuição conjunta dos dados  $(T_i, Z_i)$  pode ser escrita como:

$$f(\mathbf{t}, \mathbf{z} | \boldsymbol{\omega}, \mathbf{p}) = \prod_{i=1}^n \prod_{j=1}^k (p_j f(t_i | \omega_j))^{Z_{ij}}. \quad (3.4)$$

**Justificativa:** Seja  $L(\boldsymbol{\omega}, \mathbf{p} | \mathbf{t})$  a função de verossimilhança dada em (3.2):

$$L(\boldsymbol{\omega}, \mathbf{p} | \mathbf{t}) = \prod_{i=1}^n f(t_i | \omega, p) = \prod_{i=1}^n \sum_{j=1}^k p_j f_j(t_i | \omega_j).$$

Seja a matriz  $\mathbf{Z}(n \times k)$ , de variáveis latentes  $Z_{ij}$ ,  $i = 1, \dots, n$  e  $j = 1, \dots, k$  com a seguinte função densidade conjunta:

$$f(\mathbf{Z}) \propto \prod_{i=1}^n \prod_{j=1}^k p_{ij}^{Z_{ij}}.$$

Considerando que as variáveis aleatórias dicotômicas  $T$  e  $Z$  são independentes, a função de verossimilhança é dada por:

$$L(\boldsymbol{\omega}, \mathbf{p} | \mathbf{t}, \mathbf{z}) = \prod_{i=1}^n f(t_i | \boldsymbol{\omega}, \mathbf{p}) \prod_{i=1}^n \prod_{j=1}^k p_{ij}^{Z_{ij}}. \quad (3.5)$$

Dado o modelo em (3.1), e substituindo  $p_{ij}$  pela expressão dada em (3.3), tem-se:

$$\begin{aligned} L(\boldsymbol{\omega}, \mathbf{p} | \mathbf{t}, \mathbf{z}) &= \prod_{i=1}^n \sum_{j=1}^k p_j f(t_i | \omega_j) \prod_{i=1}^n \prod_{j=1}^k \left( \frac{p_j f_j(t_i | \omega_j)}{\sum_{j=1}^k p_j f_j(t_i | \omega_j)} \right)^{Z_{ij}} \\ &= \prod_{i=1}^n \sum_{j=1}^k p_j f_j(t_i | \omega_j) \frac{\prod_{i=1}^n \prod_{j=1}^k (p_j f_j(t_i | \omega_j))^{Z_{ij}}}{\prod_{i=1}^n \prod_{j=1}^k (\sum_{j=1}^k p_j f_j(t_i | \omega_j))^{Z_{ij}}} \end{aligned}$$

$$= \prod_{i=1}^n \sum_{j=1}^k p_j f_j(t_i | \omega_j) \frac{\prod_{i=1}^n \prod_{j=1}^k (p_j f_j(t_i | \omega_j))^{Z_{ij}}}{\prod_{i=1}^n \left( \sum_{j=1}^k p_j f_j(t_i | \omega_j) \right)^{\sum_{j=1}^k Z_{ij}}}.$$

Baseando-se em  $\sum_{j=1}^k Z_{ij} = 1$ , obtém-se que:

$$\begin{aligned} L(\boldsymbol{\omega}, \mathbf{p} | \mathbf{t}, \mathbf{z}) &= \prod_{i=1}^n \sum_{j=1}^k p_j f_j(t_i | \omega_j) \frac{\prod_{i=1}^n \prod_{j=1}^k (p_j f_j(t_i | \omega_j))^{Z_{ij}}}{\prod_{i=1}^n \sum_{j=1}^k p_j f_j(t_i | \omega_j)} \\ &= \prod_{i=1}^n \prod_{j=1}^k (p_j f_j(t_i | \omega_j))^{Z_{ij}}. \end{aligned}$$

Assim,

$$L(\boldsymbol{\omega}, \mathbf{p} | \mathbf{t}, \mathbf{z}) = f(\mathbf{t}, \mathbf{z} | \boldsymbol{\omega}, \mathbf{p}) = \prod_{i=1}^n \prod_{j=1}^k (p_j f_j(t_i | \omega_j))^{Z_{ij}}.$$

O interesse agora é verificar a que distribuição de probabilidade a variável aleatória  $T_i$  é proveniente, através de  $\mathbf{Z}_i$ . Assim, a variável indicadora  $\mathbf{Z}_i$  será introduzida, para completar a observação  $\mathbf{T}_i$ , e indicar de qual distribuição de probabilidade esta variável aleatória é proveniente.

Neste trabalho, o procedimento utilizado para simular os parâmetros de interesse do modelo misto, será o *Gibbs Sampling* baseado em amostra ampliada, descrito na subseção 3.2.1.

### 3.2.1 Algoritmo *Gibbs Sampling* com Amostra Ampliada

Nesta subseção, será introduzido um algoritmo *Gibbs Sampling* com amostra ampliada construído baseando-se em Diebolt & Robert (1994), com o objetivo de obter as estimativas para os parâmetros desconhecidos. Generalizando o *Gibbs Sampling*, descrito no

Apêndice B, é possível descrever uma aproximação para as distribuições marginais *a posteriori*, a partir de um valor inicial para os subvetores  $(\omega_1^{(m)}, \dots, \omega_s^{(m)})$ , onde  $s$  representa o número de parâmetros desconhecidos a serem estimados, com  $s \leq k$ . Deste modo, inicializando com valores  $(\omega_1^{(0)}, \dots, \omega_s^{(0)})$ , o algoritmo na etapa  $m$  é dado por:

$$\begin{aligned} \text{i) Gerar } Z_i^{(m)} &\sim \text{Multinomial}(1 : p_i^{(m)}) & i = 1, \dots, n \\ p_i^{(m)} &= (p_{i1}^{(m)}, \dots, p_{ik}^{(m)}) \end{aligned}$$

onde

$$p_{ij}^{(m)} = \frac{p_j^{(m)} f_j^{(m)}(t_i | \omega_j^{(m)})}{\sum_{j=1}^k p_j^{(m)} f_j^{(m)}(t_i | \omega_j^{(m)})}$$

ii) Na etapa  $m + 1$  gere

$$\begin{aligned} \omega_1^{(m+1)} &\sim p(\omega_1 | t_i, z_i^{(m)}, \omega_2^{(m)}, \dots, \omega_s^{(m)}) \\ &\vdots \\ \omega_s^{(m+1)} &\sim p(\omega_s | t_i, z_i^{(m)}, \omega_1^{(m+1)}, \dots, \omega_{s-1}^{(m+1)}) \end{aligned}$$

Observe que o vetor  $(\omega_1^{(m+1)}, \dots, \omega_s^{(m+1)})$  é gerado pelas suas respectivas distribuições condicionais  $p$ . A questão principal a ser analisada no algoritmo *Gibbs Sampling* é a convergência para as distribuições marginais de cada parâmetro desconhecido a ser estimado. Os Apêndices C, D e E, apresentam respectivamente, os Diagnósticos de Convergência, o Diagnóstico de Gelman e Rubin (1992) que é o de interesse neste trabalho, e o *Software WinBUGS* usado para a implementação dos algoritmos nas aplicações em estudo.

### 3.3 Modelo Misto aplicado a Teste Diagnóstico

Nesta Seção, a teoria desenvolvida na Seção 3.2 sobre modelo misto, baseando-se em Diebolt & Robert (1994), será utilizada para a construção do modelo misto aplicado a

teste diagnóstico.

Suponha que  $n$  pacientes com indício de uma determinada doença sejam selecionados um de cada vez e com reposição para serem testados, com o objetivo de verificar se possuem ou não a doença. Para cada paciente será definida a seguinte variável,  $T_i$ , sujeita a erro de classificação:

$$\begin{cases} T_i = 1 \Rightarrow \text{teste positivo para doença} \\ T_i = 0 \Rightarrow \text{teste negativo para doença} \end{cases}$$

Como a variável  $T_i$  está sujeita a erro de classificação,  $D_i$  será definido, como sendo o valor real (sem erro) de  $T_i$ , e os seguintes parâmetros:

$$\pi = P(D_i = 1) :$$

é a probabilidade do paciente realmente possuir a doença.

$$\eta^* = 1 - \eta = P(T_i = 0 | D_i = 1) :$$

é a probabilidade de um paciente doente ser resgistrado erroneamente como saudável.

$$\theta^* = 1 - \theta = P(T_i = 1 | D_i = 0) :$$

é a probabilidade de um paciente saudável ser registrado erroneamente como doente.

Com isso, é possível construir a distribuição conjunta dos pares  $(T_i, D_i)$  em relação aos parâmetros  $(\pi, \eta^*, \theta^*)$ , dada pela seguinte tabela:

**Tabela 3.1.** Distribuição conjunta de  $(T_i, D_i)$ .

	$D_i = 1$	$D_i = 0$
$T_i = 1$	$\pi(1 - \eta^*)$	$(1 - \pi)\theta^*$
$T_i = 0$	$\pi\eta^*$	$(1 - \pi)(1 - \theta^*)$

A Tabela 3.1 acima sugere uma mistura de distribuições de Bernoulli para a variável  $T_i$ , definida da seguinte forma:

$$f(\mathbf{t}|\pi, \eta^*, \theta^*) = \pi \text{Bern}(\mathbf{t}|1 - \eta^*) + (1 - \pi) \text{Bern}(\mathbf{t}|\theta^*). \quad (3.6)$$

Para evitar a não identificação dos parâmetros, é necessário restringir o modelo para  $(1 - \eta^*) > \theta^*$ ; dessa forma a estimativa de máxima verossimilhança de  $(\pi, \eta^*, \theta^*)$  será única. Esta restrição é realística, pois espera-se que a probabilidade de um paciente doente ser registrado corretamente como doente, seja maior, do que a probabilidade de um paciente não doente ser registrado erroneamente como doente.

O modelo (3.6) é considerado uma mistura de duas distribuições de Bernoulli que envolve dois tipos de erro, e duas componentes, isto é,  $\pi$  e  $(1 - \pi)$  que corresponde a probabilidade de um indivíduo possuir a doença e a probabilidade de um indivíduo não possuir a doença, respectivamente.

Para a implementação da teoria sobre modelo misto da Seção 3.2 no modelo de mistura para o teste diagnóstico, considere  $n$  variáveis aleatórias binárias independentes,  $T_i$ ,  $i = 1, \dots, n$  com o seguinte modelo misto:

$$f(\mathbf{t}|\pi, \eta^*, \theta^*) = \pi \text{Bern}(\mathbf{t}|1 - \eta^*) + (1 - \pi) \text{Bern}(\mathbf{t}|\theta^*),$$

onde  $(1 - \eta^*) > \theta^*$ .

Como o modelo misto definido em (3.6) não informa qual componente do modelo gerou a observação  $t$ , será introduzida uma variável  $Z$  que indicará a proveniência da observação  $t$ , isto é:

$$\mathbf{Z} = \begin{cases} 1 & \text{se a observação } t_i \text{ provem da distribuição Bernoulli com parâmetro } 1 - \eta^* \\ 0 & \text{se a observação } t_i \text{ provem da distribuição Bernoulli com parâmetro } \theta^* \end{cases}$$

Desta maneira, pode-se construir a distribuição conjunta dos dados ampliados  $(T_i, Z_i)$ ,  $i = 1, \dots, n$  que será dada por:

$$f(\mathbf{t}, \mathbf{z} | \pi, \eta^*, \theta^*, p) = \prod_{i=1}^n f(t_i | \pi, \eta^*, \theta^*) p_i^{Z_i} (1 - p_i)^{(1-Z_i)} I_{\substack{(\pi, \eta^*, \theta^*) \\ [(1-\eta^*) > \theta^*]}}. \quad (3.7)$$

Substituindo  $p_i$  pela expressão dada em (3.3), obtém-se o seguinte resultado:

**Resultado 3.1** - A distribuição conjunta para  $(T_i, Z_i)$  é dada por:

$$f(\mathbf{t}, \mathbf{z} | \pi, \eta^*, \theta^*, p) = \prod_{i=1}^n [\pi f(t_i | 1 - \eta^*)]^{Z_i} [(1 - \pi) f(t_i | \theta^*)]^{(1-Z_i)} I_{\substack{(\pi, \eta^*, \theta^*) \\ [(1-\eta^*) > \theta^*]}}. \quad (3.8)$$

**Demonstração:** Seja  $L(\pi, \eta^*, \theta^* | \mathbf{t})$  a função de verossimilhança dada por:

$$\begin{aligned} L(\pi, \eta^*, \theta^* | \mathbf{t}) &= \prod_{i=1}^n f(t_i | \pi, \eta^*, \theta^*) I_{\substack{(\pi, \eta^*, \theta^*) \\ [(1-\eta^*) > \theta^*]}} \\ &= \prod_{i=1}^n [\pi f(t_i | 1 - \eta^*) + (1 - \pi) f(t_i | \theta^*)] I_{\substack{(\pi, \eta^*, \theta^*) \\ [(1-\eta^*) > \theta^*]}}. \end{aligned}$$

E seja o vetor latente  $\mathbf{Z}$ , com a seguinte função densidade conjunta:

$$f(\mathbf{Z}) = \prod_{i=1}^n p_i^{Z_i} (1 - p_i)^{(1-Z_i)}.$$

Considerando que as variáveis aleatórias dicotômicas  $T$  e  $Z$  são independentes, segue que:

$$\begin{aligned} L(\pi, \eta^*, \theta^*, p | \mathbf{t}, \mathbf{z}) &= \prod_{i=1}^n f(t_i | \pi, \eta^*, \theta^*) I_{\substack{(\pi, \eta^*, \theta^*) \\ [(1-\eta^*) > \theta^*]}} \prod_{i=1}^n p_i^{Z_i} (1 - p_i)^{(1-Z_i)} \\ &= \prod_{i=1}^n f(t_i | \pi, \eta^*, \theta^*) p_i^{Z_i} (1 - p_i)^{(1-Z_i)} I_{\substack{(\pi, \eta^*, \theta^*) \\ [(1-\eta^*) > \theta^*]}}. \end{aligned}$$

Dado o modelo em (3.6), e substituindo  $p_i$  pela expressão dada em (3.3), tem-se:

$$\begin{aligned}
L(\pi, \eta^*, \theta^*, p|\mathbf{t}, \mathbf{z}) &= \prod_{i=1}^n \pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*) \\
&\quad \left( \frac{\pi f(t_i|1 - \eta^*)}{\pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*)} \right)^{Z_i} \cdot \\
&\quad \left( \frac{(1 - \pi)f(t_i|\theta^*)}{\pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*)} \right)^{(1-Z_i)} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\
&= \prod_{i=1}^n \pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*) \\
&\quad \frac{(\pi f(t_i|1 - \eta^*))^{Z_i}}{(\pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*))^{Z_i}} \cdot \\
&\quad \frac{((1 - \pi)f(t_i|\theta^*))^{(1-Z_i)}}{(\pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*))^{(1-Z_i)}} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\
&= \prod_{i=1}^n \frac{\pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*)}{\pi f(t_i|1 - \eta^*) + (1 - \pi)f(t_i|\theta^*)} \cdot \\
&\quad (\pi f(t_i|1 - \eta^*))^{Z_i} ((1 - \pi)f(t_i|\theta^*))^{(1-Z_i)} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\
&= \prod_{i=1}^n (\pi f(t_i|1 - \eta^*))^{Z_i} ((1 - \pi)f(t_i|\theta^*))^{(1-Z_i)} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \cdot
\end{aligned}$$

Logo, obtém-se que:

$$L(\pi, \eta^*, \theta^*, p|\mathbf{t}, \mathbf{z}) = \prod_{i=1}^n [\pi f(t_i|\omega = 1 - \eta^*)]^{Z_i} [(1 - \pi)f(t_i|\omega = \theta^*)]^{(1-Z_i)} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \cdot \blacksquare$$

Desenvolvendo a expressão dada em (3.8), obtém-se a função de verossimilhança expressa pelo seguinte resultado:

**Resultado 3.2** - A função de verossimilhança baseada nos dados ampliados é dada por:

$$L(\pi, \eta^*, \theta^*, p | \mathbf{t}, \mathbf{z}) = \pi^{\sum_{i=1}^n Z_i} (1 - \pi)^{n - \sum_{i=1}^n Z_i} (\eta^*)^{\sum_{i=1}^n (1 - T_i) Z_i} (1 - \eta^*)^{\sum_{i=1}^n T_i Z_i} \cdot \quad (3.9)$$

$$(\theta^*)^{\sum_{i=1}^n T_i (1 - Z_i)} (1 - \theta^*)^{\sum_{i=1}^n (1 - T_i) (1 - Z_i)} I_{[(1 - \eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)}.$$

**Demonstração:** Desenvolvendo a expressão em (3.8), segue que:

$$\begin{aligned} L(\pi, \eta^*, \theta^* | \mathbf{t}, \mathbf{z}) &= \prod_{i=1}^n [\pi f(t_i | 1 - \eta^*)]^{Z_i} [(1 - \pi) f(t_i | \theta^*)]^{(1 - Z_i)} I_{[(1 - \eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\ &= \prod_{i=1}^n [\pi (1 - \eta^*)^{T_i} (\eta^*)^{(1 - T_i)}]^{Z_i} [(1 - \pi) (\theta^*)^{T_i} (1 - \theta^*)^{(1 - T_i)}]^{(1 - Z_i)} I_{[(1 - \eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\ &= \prod_{i=1}^n [\pi^{Z_i} (1 - \eta^*)^{T_i Z_i} (\eta^*)^{(1 - T_i) Z_i}] [(1 - \pi)^{(1 - Z_i)} (\theta^*)^{T_i (1 - Z_i)}] \cdot \\ &\quad (1 - \theta^*)^{(1 - T_i) (1 - Z_i)} I_{[(1 - \eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\ &= \prod_{i=1}^n [\pi^{Z_i} (1 - \pi)^{(1 - Z_i)} (1 - \eta^*)^{T_i Z_i} (\eta^*)^{(1 - T_i) Z_i} (\theta^*)^{T_i (1 - Z_i)}] \cdot \\ &\quad (1 - \theta^*)^{(1 - T_i) (1 - Z_i)} I_{[(1 - \eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\ &= \pi^{\sum_{i=1}^n Z_i} (1 - \pi)^{n - \sum_{i=1}^n Z_i} (\eta^*)^{\sum_{i=1}^n (1 - T_i) Z_i} (1 - \eta^*)^{\sum_{i=1}^n T_i Z_i} \cdot \\ &\quad (\theta^*)^{\sum_{i=1}^n T_i (1 - Z_i)} (1 - \theta^*)^{\sum_{i=1}^n (1 - T_i) (1 - Z_i)} I_{[(1 - \eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)}. \end{aligned}$$

Logo,

$$L(\pi, \eta^*, \theta^* | \mathbf{t}, \mathbf{z}) = \pi^{\sum_{i=1}^n Z_i} (1 - \pi)^{n - \sum_{i=1}^n Z_i} (\eta^*)^{\sum_{i=1}^n (1 - T_i) Z_i} (1 - \eta^*)^{\sum_{i=1}^n T_i Z_i} \cdot (\theta^*)^{\sum_{i=1}^n T_i (1 - Z_i)} (1 - \theta^*)^{\sum_{i=1}^n (1 - T_i) (1 - Z_i)} I_{[(1 - \eta^*) > \theta^*]}(\pi, \eta^*, \theta^*) \cdot \blacksquare$$

Baseando-se no núcleo da função de verossimilhança dada em (3.9), sugere-se as seguintes densidades *a priori* para os parâmetros  $\pi, \eta^*$  e  $\theta^*$ :

$$\pi \sim \text{Beta}[\alpha, \beta] \tag{3.10}$$

$$\eta^* \sim \text{Beta}[\beta_1, \alpha_1]$$

$$\theta^* \sim \text{Beta}[\beta_2, \alpha_2], \quad \text{onde } \alpha, \beta, \beta_1, \alpha_1, \beta_2 \text{ e } \alpha_2 \text{ são conhecidos.}$$

A razão de  $\eta^* \sim \text{Beta}[\beta_1, \alpha_1]$  e  $\theta^* \sim \text{Beta}[\beta_2, \alpha_2]$  encontra-se com detalhes no Apêndice A.

Multiplicando a função de verossimilhança dada em (3.9) pelas densidades *a priori* dadas em (3.10) e normalizando como requer o teorema de Bayes, obtém-se a densidade conjunta *a posteriori* para  $(\pi, \eta^*, \theta^*)$ , dada pelo seguinte resultado:

**Resultado 3.3** - A densidade conjunta *a posteriori* para  $(\pi, \eta^*, \theta^*)$  baseada nos dados ampliados é dada por:

Dada as *prioris* para  $(\pi, \eta^*, \theta^*)$  :

$$\pi \sim \text{Beta}[\alpha, \beta] \Rightarrow f(\pi) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \pi^{\alpha-1} (1 - \pi)^{\beta-1},$$

$$\eta^* \sim \text{Beta}[\beta_1, \alpha_1] \Rightarrow f(\eta^*) = \frac{\Gamma(\beta_1 + \alpha_1)}{\Gamma(\beta_1)\Gamma(\alpha_1)} (\eta^*)^{\beta_1-1} (1 - \eta^*)^{\alpha_1-1},$$

$$\theta^* \sim \text{Beta}[\beta_2, \alpha_2] \Rightarrow f(\theta^*) = \frac{\Gamma(\beta_2 + \alpha_2)}{\Gamma(\beta_2)\Gamma(\alpha_2)} (\theta^*)^{\beta_2-1} (1 - \theta^*)^{\alpha_2-1}.$$

segue,

$$\begin{aligned}
f(\eta^*, \theta^*, \pi | \mathbf{t}, \mathbf{z}) &\propto f(\pi) f(\eta^*) f(\theta^*) L(\pi, \eta^*, \theta^* | \mathbf{t}, \mathbf{z}) \\
&\propto \pi^{\alpha-1} (1-\pi)^{\beta-1} (\eta^*)^{\beta_1-1} (1-\eta^*)^{\alpha_1-1} (\theta^*)^{\beta_2-1} (1-\theta^*)^{\alpha_2-1} \cdot \\
&\quad \pi^{\sum_{i=1}^n Z_i} (1-\pi)^{n-\sum_{i=1}^n Z_i} (\eta^*)^{\sum_{i=1}^n (1-T_i)Z_i} (1-\eta^*)^{\sum_{i=1}^n T_i Z_i} \cdot \\
&\quad (\theta^*)^{\sum_{i=1}^n T_i(1-Z_i)} (1-\theta^*)^{\sum_{i=1}^n (1-T_i)(1-Z_i)} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)} \\
&\propto \pi^{\alpha + \sum_{i=1}^n Z_i - 1} (1-\pi)^{\beta + n - \sum_{i=1}^n Z_i - 1} (\eta^*)^{\beta_1 + \sum_{i=1}^n (1-T_i)Z_i - 1} (1-\eta^*)^{\alpha_1 + \sum_{i=1}^n T_i Z_i - 1} \cdot \\
&\quad (\theta^*)^{\beta_2 + \sum_{i=1}^n T_i(1-Z_i) - 1} (1-\theta^*)^{\alpha_2 + \sum_{i=1}^n (1-T_i)(1-Z_i) - 1} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)}.
\end{aligned}$$

Assim,

$$f(\eta^*, \theta^*, \pi | \mathbf{t}, \mathbf{z}) \propto \pi^{\alpha + \sum_{i=1}^n Z_i - 1} (1-\pi)^{\beta + n - \sum_{i=1}^n Z_i - 1}. \quad (3.11)$$

$$(\eta^*)^{\beta_1 + \sum_{i=1}^n (1-T_i)Z_i - 1} (1-\eta^*)^{\alpha_1 + \sum_{i=1}^n T_i Z_i - 1}.$$

$$(\theta^*)^{\beta_2 + \sum_{i=1}^n T_i(1-Z_i) - 1} (1-\theta^*)^{\alpha_2 + \sum_{i=1}^n (1-T_i)(1-Z_i) - 1} I_{[(1-\eta^*) > \theta^*]}^{(\pi, \eta^*, \theta^*)}.$$

Portanto baseando-se em (3.11), é possível obter as seguintes distribuições condicionais:

$$\pi | \{t_i\} \{z_i\} \sim \text{Beta}[\alpha + \sum_{i=1}^n Z_i; \beta + n - \sum_{i=1}^n Z_i]$$

$$\eta^* | \{t_i\} \{z_i\} \sim \text{Beta}[\beta_1 + \sum_{i=1}^n (1-T_i)Z_i; \alpha_1 + \sum_{i=1}^n T_i Z_i] I_{[(1-\theta^*) > \eta^*]}^{(\eta^*)}$$

$$\theta^* | \{t_i\} \{z_i\} \sim \text{Beta}[\beta_2 + \sum_{i=1}^n T_i(1-Z_i); \alpha_2 + \sum_{i=1}^n (1-T_i)(1-Z_i)] I_{[(1-\eta^*) > \theta^*]}^{(\theta^*)}$$

$$Z_i | \omega_i, p \sim \text{Bernoulli}(p_i) \quad i = 1, \dots, n$$

onde

$$p_i = \frac{\pi f(t_i|\omega = 1 - \eta^*)}{\pi f(t_i|\omega = 1 - \eta^*) + (1 - \pi)f(t_i|\omega = \theta^*)}.$$

O modelo considerado para teste diagnóstico é uma mistura de duas distribuições de bernoulli envolvendo dois tipos de erros. Devido a facilidade de se trabalhar com essa distribuição, pois envolve variáveis dicotômicas, a Inferência Bayesiana aplicada ao problema produziu distribuições condicionais conhecidas. Por isso, para inferir sobre os parâmetros do modelo misto, será introduzido o algoritmo *Gibbs Sampling* com amostra ampliada descrito na Subseção 3.3.1.

### 3.3.1 Algoritmo *Gibbs Sampling* para o Modelo com Erros de Classificação

Nesta seção será abordado o algoritmo para estimar os parâmetros do modelo misto dado em (3.6) com duas componentes,  $\pi$  e  $(1 - \pi)$ . Este algoritmo é baseado no algoritmo computacional proposto por Diebolt & Robert (1994), descrito na Subseção 3.2.1.

Então, baseando-se na Subseção 3.2.1, pode-se formular o seguinte algoritmo para avaliar os parâmetros  $\pi, \eta^*$  e  $\theta^*$ .

Iniciando com valores  $(\pi^{(0)}, \eta^{*(0)}, \theta^{*(0)})$ , o algoritmo na etapa  $m$ , é dado por:

i) Gerar  $Z_i^{(m)} \sim \text{Bernoulli}(p_i^{(m)})$ ,  $i = 1, \dots, n$

$$p_i^{(m)} = \frac{\pi^{(m)} f(t_i|\omega = 1 - \eta^{*(m)})}{\pi^{(m)} f(t_i|\omega = 1 - \eta^{*(m)}) + (1 - \pi^{(m)})f(t_i|\omega = \theta^{*(m)})}.$$

ii) Na etapa  $m + 1$  gere

$$\begin{aligned} \pi^{(m+1)}|\{t_i\}\{z_i^{(m)}\} &\sim \text{Beta}\left[\alpha + \sum_{i=1}^n Z_i^{(m)}; \beta + n - \sum_{i=1}^n Z_i^{(m)}\right], \\ \eta^{*(m+1)}|\{t_i\}\{z_i^{(m)}\} &\sim \text{Beta}\left[\beta_1 + \sum_{i=1}^n (1 - T_i)Z_i^{(m)}; \alpha_1 + \sum_{i=1}^n T_i Z_i^{(m)}\right] I_{\left[\frac{(\eta^*)}{(1-\theta^*)} > \eta^*\right]}, \\ \theta^{*(m+1)}|\{t_i\}\{z_i^{(m)}\} &\sim \text{Beta}\left[\beta_2 + \sum_{i=1}^n T_i(1 - Z_i^{(m)}); \alpha_2 + \sum_{i=1}^n (1 - T_i)(1 - Z_i^{(m)})\right] I_{\left[\frac{(\theta^*)}{(1-\eta^*)} > \theta^*\right]}. \end{aligned}$$

### 3.4 Algoritmo

Esta Seção tem por objetivo, apresentar o algoritmo *Gibbs Sampling* para o Modelo com erros de classificação, dado na Subseção 3.3.1, após a sua modificação devido a consideração dos resultados dos testes positivos e negativos.

Como dito anteriormente, a idéia inicial era desenvolver um modelo misto para teste diagnóstico com erros de classificação, e um algoritmo *Gibbs Sampling* com amostra ampliada para obter os resumos *a posteriori* dos parâmetros de interesse, baseando-se na Tese de Mestrado de Vanda Donizetti Redondo e Silveira e no artigo de Diebolt & Robert (1994), mas no decorrer deste trabalho, foi encontrado o artigo de Joseph *et al.* (1995) que modificou o andamento do estudo.

No procedimento para o modelo misto apresentado na Seção 3.3, é necessário conhecer o resultado de cada teste realizado, enquanto que no procedimento de Joseph *et al.* (1995) desenvolvido no Capítulo 4, tem-se apenas a informação sobre o total de testes positivos e negativos, pois destes totais é possível obter as estatísticas *a posteriori* para os parâmetros de interesse.

Como para as aplicações dos dados reais deste trabalho, são fornecidos apenas os totais de testes positivos e negativos, o algoritmo *Gibbs Sampling* para o modelo com erros de classificação desenvolvido na Subseção 3.3.1 vai ser modificado. Esta modificação será mostrada ao se considerar os totais de testes positivos e negativos no algoritmo *Gibbs*

*Sampling* para o modelo misto, como feito no algoritmo *Gibbs Sampling* obtido pelo procedimento de Joseph *et al.* (1995).

É importante ressaltar que no algoritmo para o modelo misto descrito no presente capítulo, utiliza-se uma única variável latente para classificar um indivíduo como doente ou não, e no algoritmo obtido baseando-se no procedimento de Joseph *et al.* (1995), considerando apenas um teste diagnóstico, duas variáveis latentes foram usadas, para esta mesma classificação. Mas, devido a consideração dos totais de testes positivos e negativos os algoritmo *Gibbs Sampling* para o modelo com erros de classificação e o obtido pelo procedimento de Joseph *et al.* (1995) mostraram-se equivalentes. Com isso, a restrição  $(1 - \eta^*) > \theta^*$  ou  $\eta > (1 - \theta)$  imposta no modelo misto, se aplica a metodologia de Joseph. Porém, na abordagem de Joseph *et al.* (1995) nenhuma restrição foi imposta, supõe-se que o motivo seja a naturalidade da ocorrência desta restrição, visto que um teste diagnóstico que não seja padrão ouro é aplicado somente quando a taxa de acerto for maior do que a taxa de erro.

Então, para esta modificação considere o algoritmo *Gibbs Sampling* para o modelo com erros de classificação dado na Subseção 3.3.1, e o algoritmo *Gibbs Sampling* obtido pelo procedimento de Joseph *et al.* (1995), descritos como:

- **Algoritmo *Gibbs Sampling* para o Modelo com Erros de Classificação**

Iniciando com valores  $(\pi^{(0)}, \eta^{*(0)}, \theta^{*(0)})$ , o algoritmo na etapa  $m$ , é dado por:

i) Gerar  $Z_i^{(m)} \sim \text{Bernoulli}(p_i^{(m)})$ ,  $i = 1, \dots, n$

$$p_i^{(m)} = \frac{\pi^{(m)} f(t_i | \omega = 1 - \eta^{*(m)})}{\pi^{(m)} f(t_i | \omega = 1 - \eta^{*(m)}) + (1 - \pi^{(m)}) f(t_i | \omega = \theta^{*(m)})}.$$

ii) Na etapa  $m + 1$  gere

$$\pi^{(m+1)}|\{t_i\}\{z_i^{(m)}\} \sim \text{Beta}[\alpha + \sum_{i=1}^n Z_i^{(m)}; \beta + n - \sum_{i=1}^n Z_i^{(m)}],$$

$$\eta^{*(m+1)}|\{t_i\}\{z_i^{(m)}\} \sim \text{Beta}[\beta_1 + \sum_{i=1}^n (1 - T_i)Z_i^{(m)}; \alpha_1 + \sum_{i=1}^n T_i Z_i^{(m)}] I_{[(1-\theta^*) > \eta^*]}^{(\eta^*)},$$

$$\theta^{*(m+1)}|\{t_i\}\{z_i^{(m)}\} \sim \text{Beta}[\beta_2 + \sum_{i=1}^n T_i(1 - Z_i^{(m)}); \alpha_2 + \sum_{i=1}^n (1 - T_i)(1 - Z_i^{(m)})] I_{[(1-\eta^*) > \theta^*]}^{(\eta^*)}.$$

• **Algoritmo Gibbs Sampling obtido pelo procedimento de Joseph *et al.* (1995)**

Iniciando com valores  $(\pi^{(0)}, \eta^{(0)}, \theta^{(0)})$ , o algoritmo na etapa  $m$ , é dado por:

i) Gerar  $Y_{1i}^{(m)} \sim \text{Binomial}(a, p_1^{(m)})$ ,  $i = 1, \dots, a$

$$p_1^{(m)} = \frac{\pi^{(m)}\eta^{(m)}}{\pi^{(m)}\eta^{(m)} + (1 - \pi^{(m)})(1 - \theta^{(m)})}.$$

$Y_{2i}^{(m)} \sim \text{Binomial}(b, p_2^{(m)})$ ,  $i = 1, \dots, b$

$$p_2^{(m)} = \frac{\pi^{(m)}(1 - \eta^{(m)})}{\pi^{(m)}(1 - \eta^{(m)}) + (1 - \pi^{(m)})\theta^{(m)}}.$$

ii) Na etapa  $m + 1$  gere

$$\pi^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha + Y_{1i}^{(m)} + Y_{2i}^{(m)}; \beta + n - Y_{1i}^{(m)} - Y_{2i}^{(m)}],$$

$$\eta^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha_1 + Y_{1i}^{(m)}; \beta_1 + Y_{2i}^{(m)}],$$

$$\theta^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha_2 + b - Y_{2i}^{(m)}; \beta_2 + a - Y_{1i}^{(m)}].$$

O desenvolvimento deste algoritmo *Gibbs Sampling* obtido pelo procedimento de Joseph *et al.* (1995), está descrito com detalhes no Capítulo 4.

É relevante notar que no algoritmo *Gibbs Sampling* para o modelo com erros de classificação, dado na Subseção 3.3.1, estão sendo consideradas as taxas de erro  $\eta^* = (1 - \eta)$  e  $\theta^* = (1 - \theta)$ , enquanto que no algoritmo *Gibbs Sampling* obtido pelo procedimento de Joseph *et al.* (1995), estão sendo considerados os parâmetros  $\eta$  (sensibilidade) e  $\theta$  (especificidade).

Então, a partir do Teorema para o Método de Transformação de Variáveis, descrito no Apêndice A, com o intuito de mostrar que  $\eta = (1 - \eta^*) \sim \text{Beta}(\alpha_1, \beta_1)$  é equivalente a  $\eta^* = (1 - \eta) \sim \text{Beta}(\beta_1, \alpha_1)$ , tem-se que o algoritmo *Gibbs Sampling* para o modelo com erros de classificação fica como segue:

Iniciando com valores  $(\pi^{(0)}, \eta^{(0)}, \theta^{(0)})$ , o algoritmo na etapa  $m$ , é dado por:

i) Gerar  $Z_i^{(m)} \sim \text{Bernoulli}(p_i^{(m)})$ ,  $i = 1, \dots, n$

$$p_i^{(m)} = \frac{\pi^{(m)} f(t_i | \omega = \eta^{(m)})}{\pi^{(m)} f(t_i | \omega = \eta^{(m)}) + (1 - \pi^{(m)}) f(t_i | \omega = 1 - \theta^{(m)})}.$$

ii) Na etapa  $m + 1$  gere

$$\begin{aligned} \pi^{(m+1)} | \{t_i\} \{z_i^{(m)}\} &\sim \text{Beta}[\alpha + \sum_{i=1}^n Z_i^{(m)}; \beta + n - \sum_{i=1}^n Z_i^{(m)}], \\ \eta^{(m+1)} | \{t_i\} \{z_i^{(m)}\} &\sim \text{Beta}[\alpha_1 + \sum_{i=1}^n T_i Z_i^{(m)}; \beta_1 + \sum_{i=1}^n (1 - T_i) Z_i^{(m)}] I_{[\eta > (1-\theta)]}^{(\eta)}, \\ \theta^{(m+1)} | \{t_i\} \{z_i^{(m)}\} &\sim \text{Beta}[\alpha_2 + \sum_{i=1}^n (1 - T_i)(1 - Z_i^{(m)}); \beta_2 + \sum_{i=1}^n T_i(1 - Z_i^{(m)})] I_{[(1-\eta) < \theta]}^{(\theta)}. \end{aligned}$$

Considerando agora, os totais de testes positivos e negativos,  $a$  e  $b$ , respectivamente, o algoritmo *Gibbs Sampling* para o modelo com erros de classificação será modificado por:

- **Modificação do algoritmo Gibbs Sampling para o Modelo com Erros de Classificação**

Iniciando com valores  $(\pi^{(0)}, \eta^{(0)}, \theta^{(0)})$ , o algoritmo na etapa  $m$ , é dado por:

i) Gerar  $Z_{1i}^{(m)} \sim \text{Bernoulli}(p_{1i}^{(m)})$ ,  $i = 1, \dots, a$

$$\begin{aligned}
 p_{1i}^{(m)} &= \frac{\pi^{(m)} f(t_i = 1 | \omega = \eta^{(m)})}{\pi^{(m)} f(t_i = 1 | \omega = \eta^{(m)}) + (1 - \pi^{(m)}) f(t_i = 1 | \omega = 1 - \theta^{(m)})} \\
 &= \frac{\pi^{(m)} [(\eta^{(m)})^{t_i=1} (1 - \eta^{(m)})^{1-t_i=1}]}{\pi^{(m)} [(\eta^{(m)})^{t_i=1} (1 - \eta^{(m)})^{1-t_i=1}] + (1 - \pi^{(m)}) [(1 - \theta^{(m)})^{t_i=1} (\theta^{(m)})^{1-t_i=1}]} \\
 &= \frac{\pi^{(m)} \eta^{(m)}}{\pi^{(m)} \eta^{(m)} + (1 - \pi^{(m)}) (1 - \theta^{(m)})}.
 \end{aligned}$$

$Z_{2i}^{(m)} \sim \text{Bernoulli}(p_{2i}^{(m)})$ ,  $i = 1, \dots, b$

$$\begin{aligned}
 p_{2i}^{(m)} &= \frac{\pi^{(m)} f(t_i = 0 | \omega = 1 - \eta^{(m)})}{\pi^{(m)} f(t_i = 0 | \omega = 1 - \eta^{(m)}) + (1 - \pi^{(m)}) f(t_i = 0 | \omega = \theta^{(m)})} \\
 &= \frac{\pi^{(m)} [(1 - \eta^{(m)})^{t_i=0} (\eta^{(m)})^{1-t_i=0}]}{\pi^{(m)} [(1 - \eta^{(m)})^{t_i=0} (\eta^{(m)})^{1-t_i=0}] + (1 - \pi^{(m)}) [(1 - \theta^{(m)})^{t_i=0} (\theta^{(m)})^{1-t_i=0}]} \\
 &= \frac{\pi^{(m)} (1 - \eta^{(m)})}{\pi^{(m)} (1 - \eta^{(m)}) + (1 - \pi^{(m)}) \theta^{(m)}}.
 \end{aligned}$$

Considerando que o  $Z_{1i}$  e  $Z_{2i}$  representam a soma dos totais de testes positivos e negativos,  $a$  e  $b$ , respectivamente, é possível descrevê-los como  $Y_{1i} = Z_{1i} \sim \text{Binomial}(a, p_1)$ ,  $i = 1, \dots, a$  e  $Y_{2i} = Z_{2i} \sim \text{Binomial}(b, p_2)$ ,  $i = 1, \dots, b$ , com

$$p_1^{(m)} = \frac{\pi^{(m)}\eta^{(m)}}{\pi^{(m)}\eta^{(m)} + (1 - \pi^{(m)})(1 - \theta^{(m)})},$$

$$p_2^{(m)} = \frac{\pi^{(m)}(1 - \eta^{(m)})}{\pi^{(m)}(1 - \eta^{(m)}) + (1 - \pi^{(m)})\theta^{(m)}}.$$

ii) Na etapa  $m + 1$  gere

$$\begin{aligned} \pi^{(m+1)}|\{z_{1i}^{(m)}\}\{z_{2i}^{(m)}\} &\sim \text{Beta}[\alpha + \sum_{i=1}^n Z_i^{(m)}; \beta + n - \sum_{i=1}^n Z_i^{(m)}] \\ &\sim \text{Beta}[\alpha + \sum_{i=1}^a Z_{1i}^{(m)} + \sum_{i=1}^b Z_{2i}^{(m)}; \beta + n - \sum_{i=1}^a Z_{1i}^{(m)} - \sum_{i=1}^b Z_{2i}^{(m)}] \\ &\sim \text{Beta}[\alpha + Y_{1i}^{(m)} + Y_{2i}^{(m)}; \beta + n - Y_{1i}^{(m)} - Y_{2i}^{(m)}], \end{aligned}$$

$$\begin{aligned} \eta^{(m+1)}|\{z_{1i}^{(m)}\}\{z_{2i}^{(m)}\} &\sim \text{Beta}[\alpha_1 + \sum_{i=1}^n T_i Z_i^{(m)}; \beta_1 + \sum_{i=1}^n (1 - T_i) Z_i^{(m)}] I_{[\eta > (1-\theta)]}^{(\eta)} \\ &\sim \text{Beta}[\alpha_1 + \sum_{i=1}^a Z_{1i}^{(m)}; \beta_1 + \sum_{i=1}^n (Z_i^{(m)} - T_i Z_i^{(m)})] I_{[\eta > (1-\theta)]}^{(\eta)} \\ &\sim \text{Beta}[\alpha_1 + \sum_{i=1}^a Z_{1i}^{(m)}; \beta_1 + \sum_{i=1}^n Z_i^{(m)} - \sum_{i=1}^a T_i Z_{1i}^{(m)}] I_{[\eta > (1-\theta)]}^{(\eta)} \\ &\sim \text{Beta}[\alpha_1 + \sum_{i=1}^a Z_{1i}^{(m)}; \beta_1 + \sum_{i=1}^n Z_i^{(m)} - \sum_{i=1}^a Z_{1i}^{(m)}] I_{[\eta > (1-\theta)]}^{(\eta)} \\ &\sim \text{Beta}[\alpha_1 + \sum_{i=1}^a Z_{1i}^{(m)}; \beta_1 + \sum_{i=1}^b Z_{2i}^{(m)}] I_{[\eta > (1-\theta)]}^{(\eta)} \\ &\sim \text{Beta}[\alpha_1 + Y_{1i}^{(m)}; \beta_1 + Y_{2i}^{(m)}] I_{[\eta > (1-\theta)]}^{(\eta)}, \end{aligned}$$

$$\begin{aligned}
\theta^{(m+1)} | \{z_{1i}^{(m)}\} \{z_{2i}^{(m)}\} &\sim \text{Beta}[\alpha_2 + \sum_{i=1}^n (1 - T_i)(1 - Z_i^{(m)}); \beta_2 + \sum_{i=1}^n T_i(1 - Z_i^{(m)})] I_{[(1-\eta) < \theta]}^{(\theta)} \\
&\sim \text{Beta}[\alpha_2 + \sum_{i=1}^n (1 - Z_i^{(m)} - T_i + T_i Z_i^{(m)}); \beta_2 + \sum_{i=1}^n (T_i - T_i Z_{1i}^{(m)})] I_{[(1-\eta) < \theta]}^{(\theta)} \\
&\sim \text{Beta}[\alpha_2 + n - \sum_{i=1}^n Z_i^{(m)} - \sum_{i=1}^a T_i + \sum_{i=1}^a T_i Z_{1i}^{(m)}; \beta_2 + \sum_{i=1}^a T_i - \sum_{i=1}^a T_i Z_{1i}^{(m)}] I_{[(1-\eta) < \theta]}^{(\theta)} \\
&\sim \text{Beta}[\alpha_2 + n - \sum_{i=1}^n Z_i^{(m)} - a + \sum_{i=1}^a Z_{1i}^{(m)}; \beta_2 + a - \sum_{i=1}^a Z_{1i}^{(m)}] I_{[(1-\eta) < \theta]}^{(\theta)} \\
&\sim \text{Beta}[\alpha_2 + b - \sum_{i=1}^b Z_{2i}^{(m)}; \beta_2 + a - \sum_{i=1}^a Z_{1i}^{(m)}] I_{[(1-\eta) < \theta]}^{(\theta)} \\
&\sim \text{Beta}[\alpha_2 + b - Y_{2i}^{(m)}; \beta_2 + a - Y_{1i}^{(m)}] I_{[(1-\eta) < \theta]}^{(\theta)}.
\end{aligned}$$

Logo, o algoritmo *Gibbs Sampling* modificado, para o modelo com erros de classificação pode ser escrito como segue:

Iniciando com valores  $(\pi^{(0)}, \eta^{(0)}, \theta^{(0)})$ , o algoritmo na etapa  $m$ , é dado por:

i) Gerar  $Y_{1i}^{(m)} \sim \text{Binomial}(a, p_1^{(m)})$ ,  $i = 1, \dots, a$

$$p_1^{(m)} = \frac{\pi^{(m)} \eta^{(m)}}{\pi^{(m)} \eta^{(m)} + (1 - \pi^{(m)})(1 - \theta^{(m)})},$$

$Y_{2i}^{(m)} \sim \text{Binomial}(b, p_2^{(m)})$ ,  $i = 1, \dots, b$

$$p_2^{(m)} = \frac{\pi^{(m)}(1 - \eta^{(m)})}{\pi^{(m)}(1 - \eta^{(m)}) + (1 - \pi^{(m)})\theta^{(m)}}.$$

ii) Na etapa  $m + 1$  gere

$$\pi^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha + Y_{1i}^{(m)} + Y_{2i}^{(m)}; \beta + n - Y_{1i}^{(m)} - Y_{2i}^{(m)}],$$

$$\eta^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha_1 + Y_{1i}^{(m)}; \beta_1 + Y_{2i}^{(m)}]I_{[\eta > (1-\theta)]}^{(\eta)},$$

$$\theta^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha_2 + b - Y_{2i}^{(m)}; \beta_2 + a - Y_{1i}^{(m)}]I_{[(1-\eta) < \theta]}^{(\theta)}.$$

Dessa forma, nota-se que estes algoritmos são equivalentes. Assim, as estatísticas *a posteriori* dos parâmetros serão obtidas a partir do algoritmo *Gibbs Sampling* obtido pelo procedimento de Joseph *et al.* (1995). Estas estatísticas encontram-se nos Capítulos 4 e 5, devido o desenvolvimento das aplicações sobre a infecção do parasita *strongyloides* e sobre a obesidade em escolares, respectivamente.

# Capítulo 4

## Aplicação: Diagnóstico do Parasita *Strongyloides*

Neste capítulo, será apresentada uma aplicação com dados reais sobre a infecção causada pelo parasita *strongyloides*, com o objetivo de estimar os parâmetros de um único teste diagnóstico, assim como de dois testes combinados na ausência de um padrão ouro. Na Seção 4.1 será apresentada uma introdução sobre a Estimativa Bayesiana dos Parâmetros de Teste Diagnóstico na ausência de um padrão ouro, nas Subseções 4.1.1 e 4.1.2, definições, propriedades e a obtenção dos algoritmos *Gibbs Sampling* no caso de um teste e de dois testes diagnósticos são consideradas baseando-se em Joseph *et al.* (1995). E na Subseção 4.1.3 da Seção 4.1, será descrito o procedimento para obtenção dos coeficientes das distribuições *a priori* informativas para as sensibilidades e especificidades do teste sorológico e exame de fezes, e os hiperparâmetros da distribuição *a priori* não informativa para a prevalência da infecção.

Na Seção 4.2, segue os resultados obtidos pelo teste sorológico e exame de fezes para a infecção causada pelo parasita *strongyloides* em 162 refugiados do Camboja que chegaram em Montreal, Canadá entre julho de 1982 a fevereiro de 1983. A Seção 4.3 apresenta os resultados obtidos a partir da implementação do algoritmo *Gibbs Sampling* obtido

baseando-se no procedimento descrito em Joseph *et al.* (1995) para cada teste diagnóstico separadamente, assim como para a combinação de ambos os testes.

Uma discussão sobre a indicação de convergência dos parâmetros envolvidos na aplicação será apresentado na Seção 4.4. Na Seção 4.5, segue os comentários sobre os valores estimados nesta aplicação.

## 4.1 Estimação Bayesiana dos Parâmetros de Teste Diagnóstico na ausência de um Padrão Ouro

É comum na determinação da prevalência de uma doença através de um exame médico na população ou na investigação de novos testes diagnósticos obter resultados de um ou mais testes que investigam esta mesma doença, nenhum do qual pode ser considerado um padrão ouro. Por exemplo, dois métodos usados com frequência para estimar a prevalência de uma infecção parasitária na população observada são, o exame de fezes e o teste sorológico. Contudo, sabe-se que os resultados dos exames de fezes geralmente subestimam a prevalência, enquanto o teste sorológico geralmente resulta em superestimação. Por este motivo, nenhum destes dois métodos é considerado um padrão ouro. Entretanto, esta é praticamente sempre a situação, visto que poucos testes são considerados 100% precisos.

Apesar destas limitações, é de importância clínica e de saúde pública ter as melhores estimativas possíveis de prevalência da doença ( $\pi$ ) e parâmetros do teste, tal como a sensibilidade ( $\eta$ ), especificidade ( $\theta$ ), e os valores preditivos positivo ( $\gamma$ ) e negativo ( $\delta$ ).

A partir do uso de uma abordagem bayesiana, foi possível obter inferências simultâneas sobre a prevalência da população, sensibilidade, especificidade, e valores preditivos positivo e negativo de cada teste diagnóstico (teste sorológico e exame de fezes). Os métodos que serão apresentados neste capítulo, foram aplicados a cada teste separadamente e para os dois testes combinados (teste sorológico e exame de fezes). As densidades marginais *a posteriori* de todos os parâmetros foram estimadas usando o *Gibbs Sampling*.

As técnicas foram aplicadas para estimar a prevalência da infecção causada pelo parasita *strongyloides* e para investigar as propriedades dos testes diagnósticos, teste sorológico e exame de fezes, respectivamente, usando dados de uma pesquisa sobre 162 refugiados do Camboja que chegaram em Montreal, Canadá durante um período de 8 meses.

O motivo do uso da abordagem bayesiana se dá pelo fato de que na inferência clássica, no caso de um teste diagnóstico, é necessário assumir que a sensibilidade ( $\eta$ ) e especificidade ( $\theta$ ) sejam exatamente conhecidas, e no caso de dois testes diagnósticos, é necessário supor que a sensibilidade ( $\eta$ ) e especificidade ( $\theta$ ) de um dos dois testes sejam completamente conhecidas, enquanto que na inferência bayesiana estas necessidades não precisam ser impostas.

Dessa forma, a idéia básica por trás da abordagem bayesiana apresentada neste capítulo é eliminar a necessidade para estas suposições, construindo uma distribuição *a priori* com quantidades desconhecidas. Os dados, devido a função de verossimilhança, são então combinados com a distribuição *a priori* para obter distribuições *a posteriori* usando o Teorema de Bayes. Isto permite fazer inferências simultâneas em todos os parâmetros. As densidades *a posteriori* contêm opiniões atualizadas sobre os valores dos parâmetros do modelo, após levar em conta as informações providas pelos dados.

Como o cálculo direto das distribuições *a posteriori* pode ser difícil, o *Gibbs Sampling* será utilizado. Então este capítulo tem por meta demonstrar como aproximar densidades marginais *a posteriori* de todos parâmetros de interesse no caso de um e dois testes diagnósticos na ausência de um padrão ouro usando o *Gibbs Sampling*, e obter as estimativas destes parâmetros.

#### 4.1.1 Um Teste Diagnóstico

O problema considerado nesta subseção pode ser descrito como segue: Os resultados de um único teste diagnóstico para uma certa doença estão disponíveis em uma amostra

aleatória de indivíduos. O teste padrão ouro não está disponível, porque não existe, por causa do tamanho do erro medido, ou porque este tipo de teste praticamente não pode ser realizado. Esta última situação freqüentemente ocorre devido custos proibitivos. O objetivo é descrever inferências sobre a prevalência,  $\pi$ , da doença da qual a população foi delineada, assim como a sensibilidade,  $\eta$ , e a especificidade,  $\theta$ , do teste, e os valores preditivos positivo,  $\gamma$ , e negativo,  $\delta$ , para a população.

Seja  $a$  e  $b$  o número observado de resultados positivo e negativo do teste, respectivamente, na amostra de  $a + b = n$  indivíduos. Seja  $Y_1$  e  $Y_2$  a informação que está ausente quando não há padrão ouro, isto é, o número de resultados verdadeiros positivos do teste fora de  $a$  e  $b$ , respectivamente. Então,  $Y_1$  é o número desconhecido de verdadeiros positivos entre os  $a$  indivíduos com resultados positivos, e  $Y_2$  é o número desconhecido de verdadeiros positivos entre os  $b$  indivíduos com resultados negativos. Veja Tabela 4.1. Tal ausência de informação tem sido denominada "dado latente" por Tanner & Wong (1987), e análises usando tais dados latentes é nomeada "classe de análise latente" por Kaldor & Clayton (1985) e Walter & Irwig (1988).

**Tabela 4.1.** Resultados observados e latentes no caso de um teste diagnóstico na ausência de padrão ouro.

		DOENÇA		Total
		Presente	Ausente	
TESTE	Positivo	$Y_1$	$a - Y_1$	$a$
	Negativo	$Y_2$	$b - Y_2$	$b$
Total		$Y_1 + Y_2$	$n - (Y_1 + Y_2)$	$n$

Como qualquer indivíduo que possua ou não a doença em questão, pode ter resultado positivo ou negativo em um teste diagnóstico, existem quatro combinações possíveis para tal ocorrência. Estas combinações são representadas por:

$$P(T = 1, D = 1) = P(D = 1)P(T = 1|D = 1) = \pi\eta,$$

$$P(T = 0, D = 1) = P(D = 1)P(T = 0|D = 1) = \pi(1 - \eta),$$

$$P(T = 1, D = 0) = P(D = 0)P(T = 1|D = 0) = (1 - \pi)(1 - \theta),$$

$$P(T = 0, D = 0) = P(D = 0)P(T = 0|D = 0) = (1 - \pi)\theta.$$

A situação é resumida na Tabela 4.2.

**Tabela 4.2.** Contribuições da verossimilhança de todas as combinações possíveis de dados observados e latentes para o caso de um teste diagnóstico.

Número de indivíduos	Doença	Resultado do teste	Contribuição da verossimilhança
$Y_1$	presente	positivo	$\pi\eta$
$Y_2$	presente	negativo	$\pi(1 - \eta)$
$a - Y_1$	ausente	positivo	$(1 - \pi)(1 - \theta)$
$b - Y_2$	ausente	negativo	$(1 - \pi)\theta$

Baseando-se nas Tabelas 4.1 e 4.2, a função de verossimilhança correspondente será proporcional a distribuição binomial dada por:

$$L(\pi, \eta, \theta|n, a, b) \propto [\pi\eta + (1 - \pi)(1 - \theta)]^a [\pi(1 - \eta) + (1 - \pi)\theta]^b. \quad (4.1)$$

A função de verossimilhança (4.1) foi obtida da seguinte maneira: A probabilidade de um indivíduo pertencer a cela  $a$  (número de indivíduos com resultados positivos) da Tabela 4.1 é  $\pi\eta + (1 - \pi)(1 - \theta)$ , isto é, a doença está presente e o resultado do teste foi

positivo ou a doença está ausente e o resultado do teste foi positivo; e a probabilidade de um indivíduo pertencer a cela  $b$  (número de indivíduos com resultados negativos) da Tabela 4.1 é  $\pi(1 - \eta) + (1 - \pi)\theta$ , ou seja, a doença está presente e o resultado do teste foi negativo ou a doença está ausente e o resultado do teste foi negativo.

Como existem apenas duas celas para classificar o indivíduo como doente ou não, a função de verossimilhança será proporcional a distribuição binomial.

Uma alternativa para obter a densidade conjunta *a posteriori* de  $(\pi, \eta, \theta)$  é o algoritmo *Gibbs Sampling* com variável latente:

Sejam,

$$Y_{1i} = \begin{cases} 1 & \text{se o indivíduo é doente } (D = 1), p_1 = \pi\eta/\pi\eta + (1 - \pi)(1 - \theta) \\ 0 & \text{se o indivíduo é saudável } (D = 0), p_1 = (1 - \pi)(1 - \theta)/\pi\eta + (1 - \pi)(1 - \theta) \end{cases}$$

e

$$Y_{2i} = \begin{cases} 1 & \text{se o indivíduo é doente } (D = 1), p_2 = \pi(1 - \eta)/\pi(1 - \eta) + (1 - \pi)\theta \\ 0 & \text{se o indivíduo é saudável } (D = 0), p_2 = (1 - \pi)\theta/\pi(1 - \eta) + (1 - \pi)\theta \end{cases}$$

Onde,

$$L(Y_{1i}) = \left( \frac{\pi\eta}{\pi\eta + (1 - \pi)(1 - \theta)} \right)^{Y_{1i}} \left( \frac{(1 - \pi)(1 - \theta)}{\pi\eta + (1 - \pi)(1 - \theta)} \right)^{1 - Y_{1i}}, i = 1, \dots, a$$

$$L(Y_{2i}) = \left( \frac{\pi(1 - \eta)}{\pi(1 - \eta) + (1 - \pi)\theta} \right)^{Y_{2i}} \left( \frac{(1 - \pi)\theta}{\pi(1 - \eta) + (1 - \pi)\theta} \right)^{1 - Y_{2i}}, i = 1, \dots, b.$$

Então,

$$\begin{aligned}
L(Y_1) &= \prod_{i=1}^a \left( \frac{\pi\eta}{\pi\eta + (1-\pi)(1-\theta)} \right)^{Y_{1i}} \left( \frac{(1-\pi)(1-\theta)}{\pi\eta + (1-\pi)(1-\theta)} \right)^{1-Y_{1i}} \\
&= \left( \frac{\pi\eta}{\pi\eta + (1-\pi)(1-\theta)} \right)^{\sum_{i=1}^a Y_{1i}} \left( \frac{(1-\pi)(1-\theta)}{\pi\eta + (1-\pi)(1-\theta)} \right)^{a - \sum_{i=1}^a Y_{1i}}.
\end{aligned}$$

e

$$\begin{aligned}
L(Y_2) &= \prod_{i=1}^b \left( \frac{\pi(1-\eta)}{\pi(1-\eta) + (1-\pi)\theta} \right)^{Y_{2i}} \left( \frac{(1-\pi)\theta}{\pi(1-\eta) + (1-\pi)\theta} \right)^{1-Y_{2i}} \\
&= \left( \frac{\pi(1-\eta)}{\pi(1-\eta) + (1-\pi)\theta} \right)^{\sum_{i=1}^b Y_{2i}} \left( \frac{(1-\pi)\theta}{\pi(1-\eta) + (1-\pi)\theta} \right)^{b - \sum_{i=1}^b Y_{2i}}.
\end{aligned}$$

Logo,

$$\begin{aligned}
Y_1 &= \sum_{i=1}^a Y_{1i} \sim \text{Binomial} \left( a, \frac{\pi\eta}{\pi\eta + (1-\pi)(1-\theta)} \right), \\
Y_2 &= \sum_{i=1}^b Y_{2i} \sim \text{Binomial} \left( b, \frac{\pi(1-\eta)}{\pi(1-\eta) + (1-\pi)\theta} \right).
\end{aligned}$$

Assim, a função de verossimilhança baseada em dados ampliados é dada por:

$$L(\pi, \eta, \theta, p | y_1, y_2) \propto L(\pi, \eta, \theta | a, b, n) L(Y_1) L(Y_2) \quad (4.2)$$

$$\propto [\pi\eta + (1-\pi)(1-\theta)]^a [\pi(1-\eta) + (1-\pi)\theta]^b.$$

$$\left( \frac{\pi\eta}{\pi\eta + (1-\pi)(1-\theta)} \right)^{Y_1} \left( \frac{(1-\pi)(1-\theta)}{\pi\eta + (1-\pi)(1-\theta)} \right)^{a-Y_1}.$$

$$\begin{aligned}
& \left( \frac{\pi(1-\eta)}{\pi(1-\eta) + (1-\pi)\theta} \right)^{Y_2} \left( \frac{(1-\pi)\theta}{\pi(1-\eta) + (1-\pi)\theta} \right)^{b-Y_2} \\
\propto & [\pi\eta + (1-\pi)(1-\theta)]^a [\pi(1-\eta) + (1-\pi)\theta]^b. \\
& \frac{(\pi\eta)^{Y_1}}{[\pi\eta + (1-\pi)(1-\theta)]^{Y_1}} \frac{((1-\pi)(1-\theta))^{a-Y_1}}{[\pi\eta + (1-\pi)(1-\theta)]^{a-Y_1}}. \\
& \frac{(\pi(1-\eta))^{Y_2}}{[\pi(1-\eta) + (1-\pi)\theta]^{Y_2}} \frac{((1-\pi)\theta)^{b-Y_2}}{[\pi(1-\eta) + (1-\pi)\theta]^{b-Y_2}} \\
\propto & [\pi\eta + (1-\pi)(1-\theta)]^a [\pi(1-\eta) + (1-\pi)\theta]^b. \\
& \frac{(\pi\eta)^{Y_1} ((1-\pi)(1-\theta))^{a-Y_1}}{[\pi\eta + (1-\pi)(1-\theta)]^a} \frac{(\pi(1-\eta))^{Y_2} ((1-\pi)\theta)^{b-Y_2}}{[\pi(1-\eta) + (1-\pi)\theta]^b} \\
\propto & (\pi\eta)^{Y_1} ((1-\pi)(1-\theta))^{a-Y_1} (\pi(1-\eta))^{Y_2} ((1-\pi)\theta)^{b-Y_2} \\
\propto & \pi^{Y_1+Y_2} (1-\pi)^{n-Y_1-Y_2} \eta^{Y_1} (1-\eta)^{Y_2} \theta^{b-Y_2} (1-\theta)^{a-Y_1}.
\end{aligned}$$

Baseando-se na função de verossimilhança (4.2), sugere-se densidades *a priori* Beta para os parâmetros  $(\pi, \eta, \theta)$  :

$$\pi \sim \text{Beta}[\alpha, \beta] \tag{4.3}$$

$$\eta \sim \text{Beta}[\alpha_1, \beta_1]$$

$$\theta \sim \text{Beta}[\alpha_2, \beta_2], \text{ onde } \alpha, \beta, \alpha_1, \beta_1, \alpha_2 \text{ e } \beta_2 \text{ são conhecidos.}$$

onde

$$\begin{aligned}\pi &\sim \text{Beta}[\alpha, \beta] \Rightarrow f(\pi) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \pi^{\alpha-1} (1 - \pi)^{\beta-1}, \\ \eta &\sim \text{Beta}[\alpha_1, \beta_1] \Rightarrow f(\eta) = \frac{\Gamma(\alpha_1 + \beta_1)}{\Gamma(\alpha_1)\Gamma(\beta_1)} \eta^{\alpha_1-1} (1 - \eta)^{\beta_1-1}, \\ \theta &\sim \text{Beta}[\alpha_2, \beta_2] \Rightarrow f(\theta) = \frac{\Gamma(\alpha_2 + \beta_2)}{\Gamma(\alpha_2)\Gamma(\beta_2)} \theta^{\alpha_2-1} (1 - \theta)^{\beta_2-1}.\end{aligned}$$

Esta família de distribuições foi selecionada já que sua região de densidade positiva, de 0 a 1, é igual as regiões de todos os parâmetros de interesse neste estudo, e porque é uma família flexível, na qual uma ampla variedade de formas das densidades podem ser obtidas pela escolha de diferentes  $\alpha, \beta, \alpha_1, \beta_1, \alpha_2$  e  $\beta_2$ . Também tem a vantagem de ser a distribuição *a priori* conjugada para a verossimilhança binomial, uma propriedade que simplifica a obtenção das distribuições *a posteriori*.

Multiplicando a função de verossimilhança dada em (4.2) pelas densidades *a priori* dadas em (4.3), tem-se que a densidade conjunta *a posteriori* para  $(\pi, \eta, \theta)$  é dada por:

$$\begin{aligned}f(\pi, \eta, \theta | y_1, y_2) &\propto f(\pi)f(\eta)f(\theta)L(\pi, \eta, \theta, p | y_1, y_2) & (4.4) \\ &\propto \pi^{\alpha-1}(1 - \pi)^{\beta-1}\eta^{\alpha_1-1}(1 - \eta)^{\beta_1-1}\theta^{\alpha_2-1}(1 - \theta)^{\beta_2-1}. \\ &\quad \pi^{Y_1+Y_2}(1 - \pi)^{n-Y_1-Y_2}\eta^{Y_1}(1 - \eta)^{Y_2}\theta^{b-Y_2}(1 - \theta)^{a-Y_1} \\ &\propto \pi^{Y_1+Y_2+\alpha-1}(1 - \pi)^{n-Y_1-Y_2+\beta-1}\eta^{Y_1+\alpha_1-1}(1 - \eta)^{Y_2+\beta_1-1}. \\ &\quad \theta^{b-Y_2+\alpha_2-1}(1 - \theta)^{a-Y_1+\beta_2-1}.\end{aligned}$$

Como os dados latentes  $Y_1$  e  $Y_2$  não são observados, impedem o uso direto da equação (4.4) para calcular as densidades marginais *a posteriori* de  $\pi, \eta$  e  $\theta$ . No entanto, a inferência torna-se possível usando um algoritmo *Gibbs Sampling*.

A implementação do *Gibbs Sampling* requer a especificação de distribuições completas, isto é, distribuições condicionais de cada parâmetro dado os valores de todos os outros

parâmetros. Geralmente, a distribuição condicional completa de cada parâmetro não depende sempre de todos os outros parâmetros, que leva a algumas simplificações adicionais.

Portanto de (4.4), obtém-se as seguintes distribuições condicionais:

$$\pi|a, b, Y_1, Y_2, \alpha, \beta \sim \text{Beta}[Y_1+Y_2+\alpha; n - Y_1-Y_2+\beta],$$

$$\eta|Y_1, Y_2, \alpha_1, \beta_1 \sim \text{Beta}[Y_1+\alpha_1; Y_2+\beta_1],$$

$$\theta|a, b, Y_1, Y_2, \alpha_2, \beta_2 \sim \text{Beta}[b - Y_2+\alpha_2; a - Y_1+\beta_2],$$

$$Y_1|a, \pi, \eta, \theta \sim \text{Binomial}\left(a, \frac{\pi\eta}{\pi\eta + (1-\pi)(1-\theta)}\right),$$

$$Y_2|b, \pi, \eta, \theta \sim \text{Binomial}\left(b, \frac{\pi(1-\eta)}{\pi(1-\eta) + (1-\pi)\theta}\right).$$

Logo, o algoritmo *Gibbs Sampling* para avaliar os parâmetros  $(\pi, \eta, \theta)$  é dado por:

Iniciando com valores  $(\pi^{(0)}, \eta^{(0)}, \theta^{(0)})$ , o algoritmo na etapa  $m$ , é dado por:

(a) Gerar  $Y_{1i}^{(m)} \sim \text{Binomial}(a, p_1^{(m)})$ ,  $i = 1, \dots, a$

$$p_1^{(m)} = \frac{\pi^{(m)}\eta^{(m)}}{\pi^{(m)}\eta^{(m)} + (1-\pi^{(m)})(1-\theta^{(m)})}, \quad (4.5)$$

e

$Y_{2i}^{(m)} \sim \text{Binomial}(b, p_2^{(m)})$ ,  $i = 1, \dots, b$

$$p_2^{(m)} = \frac{\pi^{(m)}(1-\eta^{(m)})}{\pi^{(m)}(1-\eta^{(m)}) + (1-\pi^{(m)})\theta^{(m)}}. \quad (4.6)$$

(b) Na etapa  $m + 1$  gerar:

$$\pi^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha + Y_{1i}^{(m)} + Y_{2i}^{(m)}; \beta + n - (Y_{1i}^{(m)} + Y_{2i}^{(m)})], \quad (4.7)$$

$$\eta^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha_1 + Y_{1i}^{(m)}; \beta_1 + Y_{2i}^{(m)}], \quad (4.8)$$

$$\theta^{(m+1)}|\{y_{1i}^{(m)}\}\{y_{2i}^{(m)}\} \sim \text{Beta}[\alpha_2 + b - Y_{2i}^{(m)}; \beta_2 + a - Y_{1i}^{(m)}]. \quad (4.9)$$

O *Gibbs Sampling* funciona como segue: Valores iniciais arbitrários são escolhidos para cada parâmetro. Uma amostra de tamanho  $m$  é então delineada de cada distribuição condicional completa, no passeio. Os valores amostrados das iterações anteriores são usados nas distribuições condicionais para iterações posteriores. Um ciclo do algoritmo é completado quando todas distribuições condicionais tiverem sido amostradas pelo menos uma vez. O ciclo inteiro é repetido um número grande de vezes. As amostras aleatórias, geradas para cada parâmetro pode ser considerada como uma amostra aleatória da distribuição marginal *a posteriori* exata.

Para o algoritmo acima,  $Y_1$  e  $Y_2$  são gerados das expressões (4.5) e (4.6), respectivamente, dado os valores iniciais dos outros parâmetros. Então,  $\pi$  é gerado da equação condicional (4.7) usando os valores  $Y_1$  e  $Y_2$  amostrados. Obtém-se  $\eta$  e  $\theta$  das densidades dadas pelas expressões (4.8) e (4.9), respectivamente, usando os mesmos valores de  $Y_1$  e  $Y_2$ , completando o primeiro ciclo.

As amostras aleatórias geradas, repetindo o ciclo descrito pelo número desejado de vezes, são então usadas para reconstruir as densidades marginais *a posteriori* de cada parâmetro, e para encontrar os resumos de cada parâmetro destas densidades marginais *a posteriori*, tal como suas médias, medianas, desvios padrões, assim como os intervalos de credibilidade.

Os resumos *a posteriori* obtidos para os parâmetros de interesse no caso de um único teste diagnóstico, apresentam-se na Seção 4.3.

## 4.1.2 Dois Testes Diagnósticos

Os métodos da Subseção 4.1.1 podem ser estendidos à situação onde os resultados de dois testes diagnósticos para uma mesma doença estão disponíveis em uma amostra aleatoriamente selecionada de indivíduos, onde nenhum teste pode ser considerado um padrão ouro. De interesse são as densidades marginais *a posteriori* da prevalência da doença na população,  $\pi$ , da qual a amostra foi selecionada, assim como as sensibilidades,  $\eta_1$  e  $\eta_2$ , especificidades,  $\theta_1$  e  $\theta_2$ , e valores preditivos positivo,  $\gamma_1$  e  $\gamma_2$ , e negativo,  $\delta_1$  e  $\delta_2$ , dado os dados e qualquer informação *a priori* disponível. Os dados são coletados como mostra a Tabela 4.3.

**Tabela 4.3.** Dados observados de dois testes diagnósticos, na ausência de um padrão ouro.

		TESTE 2		
		Positivo	Negativo	Total
TESTE 1	Positivo	$u$	$v$	$u + v$
	Negativo	$w$	$x$	$w + x$
	Total	$u + w$	$v + x$	$n$

Sejam os dados latentes não observados  $Y_1, Y_2, Y_3$  e  $Y_4$  constituindo o número de indivíduos com resultados verdadeiros positivos fora dos valores observados das celas,  $u, v, w$  e  $x$ , respectivamente, nos dados da Tabela 4.3. Dessa forma,  $Y_1$  é o número desconhecido de verdadeiros positivos entre os  $u$  indivíduos com resultados positivos em ambos os testes,  $Y_2$  é o número desconhecido de verdadeiros positivos entre os  $v$  indivíduos com resultado positivo no primeiro teste e resultado negativo no segundo teste,  $Y_3$  é o número desconhecido de resultados verdadeiros positivos entre os  $w$  indivíduos com resultado negativo no primeiro teste e resultado positivo no segundo teste, e  $Y_4$  é o número desconhecido de resultados verdadeiros positivos entre os  $x$  indivíduos com resultados negativos em ambos os testes. Como qualquer indivíduo que realmente possua ou não a doença em questão,

pode ter resultado positivo ou negativo em cada teste, existem oito combinações possíveis para tal ocorrência. Estas combinações são representadas por:

$$P(T = 1, T = 1, D = 1) = P(D = 1)P(T = 1|D = 1)P(T = 1|T = 1, D = 1).$$

Considerando que o primeiro teste e o segundo teste são independentes, tem-se que:

$$P(T = 1, T = 1, D = 1) = P(D = 1)P(T = 1|D = 1)P(T = 1|D = 1) = \pi\eta_1\eta_2.$$

O cálculo das probabilidades de ocorrência das outras sete combinações, são obtidas de modo análogo:

$$P(T = 1, T = 0, D = 1) = P(D = 1)P(T = 1|D = 1)P(T = 0|D = 1) = \pi\eta_1(1 - \eta_2),$$

$$P(T = 0, T = 1, D = 1) = P(D = 1)P(T = 0|D = 1)P(T = 1|D = 1) = \pi(1 - \eta_1)\eta_2,$$

$$P(T = 0, T = 0, D = 1) = P(D = 1)P(T = 0|D = 1)P(T = 0|D = 1) = \pi(1 - \eta_1)(1 - \eta_2),$$

$$\begin{aligned} P(T = 1, T = 1, D = 0) &= P(D = 0)P(T = 1|D = 0)P(T = 1|D = 0) \\ &= (1 - \pi)(1 - \theta_1)(1 - \theta_2), \end{aligned}$$

$$P(T = 1, T = 0, D = 0) = P(D = 0)P(T = 1|D = 0)P(T = 0|D = 0) = (1 - \pi)(1 - \theta_1)\theta_2,$$

$$P(T = 0, T = 1, D = 0) = P(D = 0)P(T = 0|D = 0)P(T = 1|D = 0) = (1 - \pi)\theta_1(1 - \theta_2),$$

$$P(T = 0, T = 0, D = 0) = P(D = 0)P(T = 0|D = 0)P(T = 0|D = 0) = (1 - \pi)\theta_1\theta_2.$$

A situação é sumarizada na Tabela 4.4.

**Tabela 4.4.** Contribuições da verossimilhança de todas as combinações possíveis de dados observados e latentes para o caso de dois testes diagnósticos.

Número de indivíduos	Doença	Resultado do teste 1	Resultado do teste 2	Contribuição da verossimilhança
$Y_1$	presente	positivo	positivo	$\pi\eta_1\eta_2$
$Y_2$	presente	positivo	negativo	$\pi\eta_1(1 - \eta_2)$
$Y_3$	presente	negativo	positivo	$\pi(1 - \eta_1)\eta_2$
$Y_4$	presente	negativo	negativo	$\pi(1 - \eta_1)(1 - \eta_2)$
$u - Y_1$	ausente	positivo	positivo	$(1 - \pi)(1 - \theta_1)(1 - \theta_2)$
$v - Y_2$	ausente	positivo	negativo	$(1 - \pi)(1 - \theta_1)\theta_2$
$w - Y_3$	ausente	negativo	positivo	$(1 - \pi)\theta_1(1 - \theta_2)$
$x - Y_4$	ausente	negativo	negativo	$(1 - \pi)\theta_1\theta_2$

Baseando-se nas Tabelas 4.3 e 4.4, a função de verossimilhança correspondente será proporcional a distribuição multinomial dada por:

$$\begin{aligned}
 L(\pi, \eta_1, \eta_2, \theta_1, \theta_2 | u, v, w, x) &\propto [\pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2)]^u \cdot & (4.10) \\
 &[\pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2]^v \cdot \\
 &[\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)]^w \cdot \\
 &[\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2]^x.
 \end{aligned}$$

A função de verossimilhança (4.10) foi obtida da seguinte maneira: A probabilidade de um indivíduo pertencer a cela  $u$  (número de indivíduos com resultados positivos em ambos os testes) da Tabela 4.3 é  $\pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2)$ , isto é, a doença está presente e o resultado do primeiro e segundo teste foi positivo ou a doença está ausente e o resultado do primeiro e segundo teste foi positivo; a probabilidade de um indivíduo pertencer a cela  $v$  (número de indivíduos com resultados positivos no primeiro teste e

negativos no segundo) da Tabela 4.3 é  $\pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2$ , ou seja, a doença está presente e o resultado do primeiro teste foi positivo e do segundo foi negativo ou a doença está ausente e o resultado do primeiro teste foi positivo e do segundo foi negativo. Analogamente, a probabilidade de um indivíduo pertencer a cela  $w$  (número de indivíduos com resultados negativos no primeiro teste e positivos no segundo) da Tabela 4.3 é  $\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)$ , desta forma, tem-se que a doença está presente e o resultado do primeiro teste foi negativo e do segundo foi positivo ou a doença está ausente e o resultado do primeiro teste foi negativo e do segundo foi positivo, e a probabilidade de um indivíduo pertencer a cela  $x$  (número de indivíduos com resultados negativos em ambos os testes) da Tabela 4.3 é  $\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2$ , isto é, a doença está presente e o resultado do primeiro e segundo teste foi negativo ou a doença está ausente e o resultado do primeiro e segundo teste foi negativo.

Assim, como há quatro celas para classificar o indivíduo como doente ou não, a função de verossimilhança será proporcional a distribuição multinomial.

Uma alternativa para obter a densidade conjunta *a posteriori* de  $(\pi, \eta_1, \eta_2, \theta_1, \theta_2)$  é o algoritmo *Gibbs Sampling* com variável latente:

Sejam,

$$Y_{1i} = \begin{cases} 1 \text{ se doente, } p_1 = \pi\eta_1\eta_2 / \pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2) \\ 0 \text{ se saudável, } p_1 = (1 - \pi)(1 - \theta_1)(1 - \theta_2) / \pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2) \end{cases}$$

$$Y_{2i} = \begin{cases} 1 \text{ se doente, } p_2 = \pi\eta_1(1 - \eta_2) / \pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2 \\ 0 \text{ se saudável, } p_2 = (1 - \pi)(1 - \theta_1)\theta_2 / \pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2 \end{cases}$$

$$Y_{3i} = \begin{cases} 1 \text{ se doente, } p_3 = \pi(1 - \eta_1)\eta_2 / \pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2) \\ 0 \text{ se saudável, } p_3 = (1 - \pi)\theta_1(1 - \theta_2) / \pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2) \end{cases}$$

$$Y_{4i} = \begin{cases} 1 \text{ se doente, } p_4 = \pi(1 - \eta_1)(1 - \eta_2) / \pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2 \\ 0 \text{ se saudável, } p_4 = (1 - \pi)\theta_1\theta_2 / \pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2 \end{cases}$$

Onde,

$$L(Y_{1i}) = \left( \frac{\pi\eta_1\eta_2}{\pi\eta_1\eta_2 + (1-\pi)(1-\theta_1)(1-\theta_2)} \right)^{Y_{1i}} \left( \frac{(1-\pi)(1-\theta_1)(1-\theta_2)}{\pi\eta_1\eta_2 + (1-\pi)(1-\theta_1)(1-\theta_2)} \right)^{1-Y_{1i}},$$

$$L(Y_{2i}) = \left( \frac{\pi\eta_1(1-\eta_2)}{\pi\eta_1(1-\eta_2) + (1-\pi)(1-\theta_1)\theta_2} \right)^{Y_{2i}} \left( \frac{(1-\pi)(1-\theta_1)\theta_2}{\pi\eta_1(1-\eta_2) + (1-\pi)(1-\theta_1)\theta_2} \right)^{1-Y_{2i}},$$

$$L(Y_{3i}) = \left( \frac{\pi(1-\eta_1)\eta_2}{\pi(1-\eta_1)\eta_2 + (1-\pi)\theta_1(1-\theta_2)} \right)^{Y_{3i}} \left( \frac{(1-\pi)\theta_1(1-\theta_2)}{\pi(1-\eta_1)\eta_2 + (1-\pi)\theta_1(1-\theta_2)} \right)^{1-Y_{3i}},$$

$$L(Y_{4i}) = \left( \frac{\pi(1-\eta_1)(1-\eta_2)}{\pi(1-\eta_1)(1-\eta_2) + (1-\pi)\theta_1\theta_2} \right)^{Y_{4i}} \left( \frac{(1-\pi)\theta_1\theta_2}{\pi(1-\eta_1)(1-\eta_2) + (1-\pi)\theta_1\theta_2} \right)^{1-Y_{4i}}.$$

Então,

$$\begin{aligned} L(Y_1) &= \prod_{i=1}^u \left( \frac{\pi\eta_1\eta_2}{\pi\eta_1\eta_2 + (1-\pi)(1-\theta_1)(1-\theta_2)} \right)^{Y_{1i}} \left( \frac{(1-\pi)(1-\theta_1)(1-\theta_2)}{\pi\eta_1\eta_2 + (1-\pi)(1-\theta_1)(1-\theta_2)} \right)^{1-Y_{1i}} \\ &= \left( \frac{\pi\eta_1\eta_2}{\pi\eta_1\eta_2 + (1-\pi)(1-\theta_1)(1-\theta_2)} \right)^{\sum_{i=1}^u Y_{1i}} \left( \frac{(1-\pi)(1-\theta_1)(1-\theta_2)}{\pi\eta_1\eta_2 + (1-\pi)(1-\theta_1)(1-\theta_2)} \right)^{u - \sum_{i=1}^u Y_{1i}} \end{aligned}$$

$$\begin{aligned} L(Y_2) &= \prod_{i=1}^v \left( \frac{\pi\eta_1(1-\eta_2)}{\pi\eta_1(1-\eta_2) + (1-\pi)(1-\theta_1)\theta_2} \right)^{Y_{2i}} \left( \frac{(1-\pi)(1-\theta_1)\theta_2}{\pi\eta_1(1-\eta_2) + (1-\pi)(1-\theta_1)\theta_2} \right)^{1-Y_{2i}} \\ &= \left( \frac{\pi\eta_1(1-\eta_2)}{\pi\eta_1(1-\eta_2) + (1-\pi)(1-\theta_1)\theta_2} \right)^{\sum_{i=1}^v Y_{2i}} \left( \frac{(1-\pi)(1-\theta_1)\theta_2}{\pi\eta_1(1-\eta_2) + (1-\pi)(1-\theta_1)\theta_2} \right)^{v - \sum_{i=1}^v Y_{2i}} \end{aligned}$$

$$\begin{aligned}
L(Y_3) &= \prod_{i=1}^w \left( \frac{\pi(1-\eta_1)\eta_2}{\pi(1-\eta_1)\eta_2+(1-\pi)\theta_1(1-\theta_2)} \right)^{Y_{3i}} \left( \frac{(1-\pi)\theta_1(1-\theta_2)}{\pi(1-\eta_1)\eta_2+(1-\pi)\theta_1(1-\theta_2)} \right)^{1-Y_{3i}} \\
&= \left( \frac{\pi(1-\eta_1)\eta_2}{\pi(1-\eta_1)\eta_2+(1-\pi)\theta_1(1-\theta_2)} \right)^{\sum_{i=1}^w Y_{3i}} \left( \frac{(1-\pi)\theta_1(1-\theta_2)}{\pi(1-\eta_1)\eta_2+(1-\pi)\theta_1(1-\theta_2)} \right)^{w-\sum_{i=1}^w Y_{3i}}
\end{aligned}$$

$$\begin{aligned}
L(Y_4) &= \prod_{i=1}^x \left( \frac{\pi(1-\eta_1)(1-\eta_2)}{\pi(1-\eta_1)(1-\eta_2)+(1-\pi)\theta_1\theta_2} \right)^{Y_{4i}} \left( \frac{(1-\pi)\theta_1\theta_2}{\pi(1-\eta_1)(1-\eta_2)+(1-\pi)\theta_1\theta_2} \right)^{1-Y_{4i}} \\
&= \left( \frac{\pi(1-\eta_1)(1-\eta_2)}{\pi(1-\eta_1)(1-\eta_2)+(1-\pi)\theta_1\theta_2} \right)^{\sum_{i=1}^x Y_{4i}} \left( \frac{(1-\pi)\theta_1\theta_2}{\pi(1-\eta_1)(1-\eta_2)+(1-\pi)\theta_1\theta_2} \right)^{x-\sum_{i=1}^x Y_{4i}}.
\end{aligned}$$

Logo,

$$\begin{aligned}
Y_1 &= \sum_{i=1}^u Y_{1i} \sim \text{Binomial} \left( u, \frac{\pi\eta_1\eta_2}{\pi\eta_1\eta_2+(1-\pi)(1-\theta_1)(1-\theta_2)} \right), \\
Y_2 &= \sum_{i=1}^v Y_{2i} \sim \text{Binomial} \left( v, \frac{\pi\eta_1(1-\eta_2)}{\pi\eta_1(1-\eta_2)+(1-\pi)(1-\theta_1)\theta_2} \right), \\
Y_3 &= \sum_{i=1}^w Y_{3i} \sim \text{Binomial} \left( w, \frac{\pi(1-\eta_1)\eta_2}{\pi(1-\eta_1)\eta_2+(1-\pi)\theta_1(1-\theta_2)} \right), \\
Y_4 &= \sum_{i=1}^x Y_{4i} \sim \text{Binomial} \left( x, \frac{\pi(1-\eta_1)(1-\eta_2)}{\pi(1-\eta_1)(1-\eta_2)+(1-\pi)\theta_1\theta_2} \right).
\end{aligned}$$

Assim, a função de verossimilhança baseada em dados ampliados é dada por:

$$L(\pi, \eta_1, \eta_2, \theta_1, \theta_2, p | y_1, y_2, y_3, y_4) \propto L(\pi, \eta_1, \eta_2, \theta_1, \theta_2 | u, v, w, x) L(Y_1) L(Y_2) L(Y_3) L(Y_4) \tag{4.11}$$

$$\propto [\pi\eta_1\eta_2+(1-\pi)(1-\theta_1)(1-\theta_2)]^u.$$

$$\propto [\pi\eta_1(1-\eta_2)+(1-\pi)(1-\theta_1)\theta_2]^v$$

$$[\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)]^w [\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2]^x.$$

$$\left( \frac{\pi\eta_1\eta_2}{\pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2)} \right)^{Y_1} \left( \frac{(1 - \pi)(1 - \theta_1)(1 - \theta_2)}{\pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2)} \right)^{u - Y_1}.$$

$$\left( \frac{\pi\eta_1(1 - \eta_2)}{\pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2} \right)^{Y_2} \left( \frac{(1 - \pi)(1 - \theta_1)\theta_2}{\pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2} \right)^{Y_2}.$$

$$\left( \frac{\pi(1 - \eta_1)\eta_2}{\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)} \right)^{Y_3} \left( \frac{(1 - \pi)\theta_1(1 - \theta_2)}{\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)} \right)^{w - Y_3}.$$

$$\left( \frac{\pi(1 - \eta_1)(1 - \eta_2)}{\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2} \right)^{Y_4} \left( \frac{(1 - \pi)\theta_1\theta_2}{\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2} \right)^{x - Y_4}$$

$$\propto [\pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2)]^u [\pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2]^v.$$

$$[\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)]^w [\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2]^x.$$

$$\frac{(\pi\eta_1\eta_2)^{Y_1}}{[\pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2)]^{Y_1}} \frac{(\pi\eta_1\eta_2)^{Y_1}}{[\pi\eta_1\eta_2 + (1 - \pi)(1 - \theta_1)(1 - \theta_2)]^{Y_1}}.$$

$$\frac{(\pi\eta_1(1 - \eta_2))^{Y_2}}{[\pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2]^{Y_2}} \frac{((1 - \pi)(1 - \theta_1)\theta_2)^{v - Y_2}}{[\pi\eta_1(1 - \eta_2) + (1 - \pi)(1 - \theta_1)\theta_2]^{v - Y_2}}.$$

$$\frac{(\pi(1 - \eta_1)\eta_2)^{Y_3}}{[\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)]^{Y_3}} \frac{((1 - \pi)\theta_1(1 - \theta_2))^{w - Y_3}}{[\pi(1 - \eta_1)\eta_2 + (1 - \pi)\theta_1(1 - \theta_2)]^{w - Y_3}}.$$

$$\frac{(\pi(1 - \eta_1)(1 - \eta_2))^{Y_4}}{[\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2]^{Y_4}} \frac{((1 - \pi)\theta_1\theta_2)^{x - Y_4}}{[\pi(1 - \eta_1)(1 - \eta_2) + (1 - \pi)\theta_1\theta_2]^{x - Y_4}}$$

$$\begin{aligned}
&\propto [\pi\eta_1\eta_2+(1-\pi)(1-\theta_1)(1-\theta_2)]^u[\pi\eta_1(1-\eta_2)+(1-\pi)(1-\theta_1)\theta_2]^v. \\
&[\pi(1-\eta_1)\eta_2+(1-\pi)\theta_1(1-\theta_2)]^w[\pi(1-\eta_1)(1-\eta_2)+(1-\pi)\theta_1\theta_2]^x. \\
&\frac{(\pi\eta_1\eta_2)^{Y_1}((1-\pi)(1-\theta_1)(1-\theta_2))^{u-Y_1}}{[\pi\eta_1\eta_2+(1-\pi)(1-\theta_1)(1-\theta_2)]^u} \frac{(\pi\eta_1(1-\eta_2))^{Y_2}((1-\pi)(1-\theta_1)\theta_2)^{v-Y_2}}{[\pi\eta_1(1-\eta_2)+(1-\pi)(1-\theta_1)\theta_2]^v}. \\
&\frac{(\pi(1-\eta_1)\eta_2)^{Y_3}((1-\pi)\theta_1(1-\theta_2))^{w-Y_3}}{[\pi(1-\eta_1)\eta_2+(1-\pi)\theta_1(1-\theta_2)]^w} \frac{(\pi(1-\eta_1)(1-\eta_2))^{Y_4}((1-\pi)\theta_1\theta_2)^{x-Y_4}}{[\pi(1-\eta_1)(1-\eta_2)+(1-\pi)\theta_1\theta_2]^x} \\
&\propto (\pi\eta_1\eta_2)^{Y_1}((1-\pi)(1-\theta_1)(1-\theta_2))^{u-Y_1}(\pi\eta_1(1-\eta_2))^{Y_2}((1-\pi)(1-\theta_1)\theta_2)^{v-Y_2}. \\
&(\pi(1-\eta_1)\eta_2)^{Y_3}((1-\pi)\theta_1(1-\theta_2))^{w-Y_3}(\pi(1-\eta_1)(1-\eta_2))^{Y_4}((1-\pi)\theta_1\theta_2)^{x-Y_4} \\
&\propto \pi^{Y_1+Y_2+Y_3+Y_4}(1-\pi)^{n-(Y_1+Y_2+Y_3+Y_4)}\eta_1^{Y_1+Y_2}(1-\eta_1)^{Y_3+Y_4}\eta_2^{Y_1+Y_3}(1-\eta_2)^{Y_2+Y_4}. \\
&\theta_1^{w+x-(Y_3+Y_4)}(1-\theta_1)^{u+v-(Y_1+Y_2)}\theta_2^{v+x-(Y_2+Y_4)}(1-\theta_2)^{u+w-(Y_1+Y_3)}.
\end{aligned}$$

Baseando-se na função de verossimilhança (4.11), sugere-se densidades *a priori* Beta para os parâmetros  $(\pi, \eta_1, \eta_2, \theta_1, \theta_2)$  :

$$\begin{aligned}
\pi &\sim \text{Beta}[\alpha, \beta] \\
\eta_1 &\sim \text{Beta}[\alpha_1, \beta_1] \\
\eta_2 &\sim \text{Beta}[\alpha_2, \beta_2] \\
\theta_1 &\sim \text{Beta}[\alpha_3, \beta_3] \\
\theta_2 &\sim \text{Beta}[\alpha_4, \beta_4], \quad \text{onde } \alpha, \beta, \alpha_1, \beta_1, \alpha_2, \beta_2, \alpha_3, \beta_3, \alpha_4, \beta_4 \text{ são conhecidos.}
\end{aligned} \tag{4.12}$$

Multiplicando a função de verossimilhança dada em (4.11) pelas densidades *a priori* dadas em (4.12), tem-se que a densidade conjunta *a posteriori* para  $(\pi, \eta_1, \eta_2, \theta_1, \theta_2)$  é dada por:

$$\begin{aligned}
f(\pi, \eta_1, \eta_2, \theta_1, \theta_2 | y_1, y_2, y_3, y_4) &\propto f(\pi) f(\eta_1) f(\theta_1) f(\eta_2) f(\theta_2) L(\pi, \eta_1, \eta_2, \theta_1, \theta_2, p | y_1, y_2, y_3, y_4) \\
&\propto \pi^{\alpha-1} (1-\pi)^{\beta-1} \eta_1^{\alpha_1-1} (1-\eta_1)^{\beta_1-1} \eta_2^{\alpha_2-1} (1-\eta_2)^{\beta_2-1} \\
&\quad \theta_1^{\alpha_3-1} (1-\theta_1)^{\beta_3-1} \theta_2^{\alpha_4-1} (1-\theta_2)^{\beta_4-1}. \\
&\quad \pi^{Y_1+Y_2+Y_3+Y_4} (1-\pi)^{n-(Y_1+Y_2+Y_3+Y_4)} \eta_1^{Y_1+Y_2} (1-\eta_1)^{Y_3+Y_4}. \\
&\quad \eta_2^{Y_1+Y_3} (1-\eta_2)^{Y_2+Y_4} \theta_1^{w+x-(Y_3+Y_4)} (1-\theta_1)^{u+v-(Y_1+Y_2)}. \\
&\quad \theta_2^{v+x-(Y_2+Y_4)} (1-\theta_2)^{u+w-(Y_1+Y_3)} \\
&\propto \pi^{Y_1+Y_2+Y_3+Y_4+\alpha-1} (1-\pi)^{n-(Y_1+Y_2+Y_3+Y_4)+\beta-1} \eta_1^{Y_1+Y_2+\alpha_1-1}. \\
&\quad (1-\eta_1)^{Y_3+Y_4+\beta_1-1} \eta_2^{Y_1+Y_3+\alpha_2-1} (1-\eta_2)^{Y_2+Y_4+\beta_2-1}. \\
&\quad \theta_1^{w+x-(Y_3+Y_4)+\alpha_3-1} (1-\theta_1)^{u+v-(Y_1+Y_2)+\beta_3-1}. \\
&\quad \theta_2^{v+x-(Y_2+Y_4)+\alpha_4-1} (1-\theta_2)^{u+w-(Y_1+Y_3)+\beta_4-1}. \tag{4.13}
\end{aligned}$$

Baseando-se em (4.13), obtém-se as seguintes distribuições condicionais:

$$\begin{aligned}
\pi | u, v, w, v, Y_1, Y_2, Y_3, Y_4, \alpha, \beta &\sim \text{Beta}[Y_1+Y_2+Y_3+Y_4+\alpha; n - (Y_1+Y_2+Y_3+Y_4) + \beta], \\
\eta_1 | Y_1, Y_2, Y_3, Y_4, \alpha_{\eta_1}, \beta_1 &\sim \text{Beta}[Y_1+Y_2+\alpha_1; Y_3+Y_4+\beta_1], \\
\eta_2 | Y_1, Y_2, Y_3, Y_4, \alpha_2, \beta_2 &\sim \text{Beta}[Y_1+Y_3+\alpha_2; Y_2+Y_4+\beta_2], \\
\theta_1 | u, v, w, v, Y_1, Y_2, Y_3, Y_4, \alpha_3, \beta_3 &\sim \text{Beta}[w+x-(Y_3+Y_4)+\alpha_3; u+v-(Y_1+Y_2)+\beta_3], \\
\theta_2 | u, v, w, v, Y_1, Y_2, Y_3, Y_4, \alpha_4, \beta_4 &\sim \text{Beta}[v+x-(Y_2+Y_4)+\alpha_4; u+w-(Y_1+Y_3)+\beta_4], \\
Y_1 | u, \pi, \eta_1, \eta_2, \theta_1, \theta_2 &\sim \text{Binomial} \left( u, \frac{\pi \eta_1 \eta_2}{\pi \eta_1 \eta_2 + (1-\pi)(1-\theta_1)(1-\theta_2)} \right), \\
Y_2 | v, \pi, \eta_1, \eta_2, \theta_1, \theta_2 &\sim \text{Binomial} \left( v, \frac{\pi \eta_1 (1-\eta_2)}{\pi \eta_1 (1-\eta_2) + (1-\pi)(1-\theta_1)\theta_2} \right),
\end{aligned}$$

$$Y_3|w, \pi, \eta_1, \eta_2, \theta_1, \theta_2 \sim \text{Binomial} \left( w, \frac{\pi(1-\eta_1)\eta_2}{\pi(1-\eta_1)\eta_2 + (1-\pi)\theta_1(1-\theta_2)} \right),$$

$$Y_4|x, \pi, \eta_1, \eta_2, \theta_1, \theta_2 \sim \text{Binomial} \left( x, \frac{\pi(1-\eta_1)(1-\eta_2)}{\pi(1-\eta_1)(1-\eta_2) + (1-\pi)\theta_1\theta_2} \right).$$

Logo, o algoritmo *Gibbs Sampling* para avaliar os parâmetros  $(\pi, \eta_1, \eta_2, \theta_1, \theta_2)$  é dado por:

Iniciando com valores  $(\pi^{(0)}, \eta_1^{(0)}, \eta_2^{(0)}, \theta_1^{(0)}, \theta_2^{(0)})$ , o algoritmo na etapa  $m$ , é dado por:

(a) Gerar  $Y_{1i}^{(m)} \sim \text{Binomial}(u, p_1^{(m)})$ ,  $i = 1, \dots, u$

$$p_1^{(m)} = \frac{\pi^{(m)}\eta_1^{(m)}\eta_2^{(m)}}{\pi^{(m)}\eta_1^{(m)}\eta_2^{(m)} + (1-\pi^{(m)})(1-\theta_1^{(m)})(1-\theta_2^{(m)})}, \quad (4.14)$$

$Y_{2i}^{(m)} \sim \text{Binomial}(v, p_2^{(m)})$ ,  $i = 1, \dots, v$

$$p_2^{(m)} = \frac{\pi^{(m)}\eta_1^{(m)}(1-\eta_2^{(m)})}{\pi^{(m)}\eta_1^{(m)}(1-\eta_2^{(m)}) + (1-\pi^{(m)})(1-\theta_1^{(m)})\theta_2^{(m)}}, \quad (4.15)$$

$Y_{3i}^{(m)} \sim \text{Binomial}(w, p_3^{(m)})$ ,  $i = 1, \dots, w$

$$p_3^{(m)} = \frac{\pi^{(m)}(1-\eta_1^{(m)})\eta_2^{(m)}}{\pi^{(m)}(1-\eta_1^{(m)})\eta_2^{(m)} + (1-\pi^{(m)})\theta_1^{(m)}(1-\theta_2^{(m)})}, \quad (4.16)$$

$Y_{4i}^{(m)} \sim \text{Binomial}(x, p_4^{(m)})$ ,  $i = 1, \dots, x$

$$p_4^{(m)} = \frac{\pi^{(m)}(1-\eta_1^{(m)})(1-\eta_2^{(m)})}{\pi^{(m)}(1-\eta_1^{(m)})(1-\eta_2^{(m)}) + (1-\pi^{(m)})\theta_1^{(m)}\theta_2^{(m)}}. \quad (4.17)$$

(b) Na etapa  $m + 1$  gerar:

$$\pi^{(m+1)} | \{y_{1i}^{(m)}\} \{y_{2i}^{(m)}\} \{y_{3i}^{(m)}\} \{y_{4i}^{(m)}\} \quad (4.18)$$

$$\sim \text{Beta}[Y_{1i}^{(m)} + Y_{2i}^{(m)} + Y_{3i}^{(m)} + Y_{4i}^{(m)} + \alpha; n - (Y_{1i}^{(m)} + Y_{2i}^{(m)} + Y_{3i}^{(m)} + Y_{4i}^{(m)}) + \beta],$$

$$\eta_1^{(m+1)} | \{y_{1i}^{(m)}\} \{y_{2i}^{(m)}\} \{y_{3i}^{(m)}\} \{y_{4i}^{(m)}\} \sim \text{Beta}[Y_{1i}^{(m)} + Y_{2i}^{(m)} + \alpha_1; Y_{3i}^{(m)} + Y_{4i}^{(m)} + \beta_1], \quad (4.19)$$

$$\eta_2^{(m+1)} | \{y_{1i}^{(m)}\} \{y_{2i}^{(m)}\} \{y_{3i}^{(m)}\} \{y_{4i}^{(m)}\} \sim \text{Beta}[Y_{1i}^{(m)} + Y_{3i}^{(m)} + \alpha_2; Y_{2i}^{(m)} + Y_{4i}^{(m)} + \beta_2], \quad (4.20)$$

$$\begin{aligned} \theta_1^{(m+1)} | \{y_{1i}^{(m)}\} \{y_{2i}^{(m)}\} \{y_{3i}^{(m)}\} \{y_{4i}^{(m)}\} \\ \sim \text{Beta}[w + x - (Y_{3i}^{(m)} + Y_{4i}^{(m)}) + \alpha_3; u + v - (Y_{1i}^{(m)} + Y_{2i}^{(m)}) + \beta_3], \end{aligned} \quad (4.21)$$

$$\begin{aligned} \theta_2^{(m+1)} | \{y_{1i}^{(m)}\} \{y_{2i}^{(m)}\} \{y_{3i}^{(m)}\} \{y_{4i}^{(m)}\} \\ \sim \text{Beta}[v + x - (Y_{2i}^{(m)} + Y_{4i}^{(m)}) + \alpha_4; u + w - (Y_{1i}^{(m)} + Y_{3i}^{(m)}) + \beta_4]. \end{aligned} \quad (4.22)$$

O *Gibbs Sampling* é usado para amostrar o passeio das distribuições (4.14) - (4.22) de uma maneira semelhante ao procedimento utilizado no caso de um teste diagnóstico, traçado na Subseção 4.1.1.

Assim, o algoritmo *Gibbs Sampling*, descrito acima, fornece amostras aleatórias para cada parâmetro de interesse, e estas amostras aleatórias podem ser usadas para resumir as densidades marginais *a posteriori* de cada parâmetro, tal como suas médias, medianas, desvios padrões e os intervalos de credibilidade.

Os resumos *a posteriori* obtidos para os parâmetros de interesse no caso de dois testes diagnósticos, são mostrados na Seção 4.3.

### 4.1.3 Distribuições *a priori*

Um passo importante em qualquer análise bayesiana é obter uma distribuição *a priori* para todos os parâmetros do modelo.

Quando não se tem nenhuma informação adicional além dos dados, uma escolha comum é considerar uma *priori* difusa ou “não informativa”, como por exemplo, uma densidade uniforme. O problema é que as densidades *a priori* não informativas são geralmente

impróprias, não garantindo a existência da *posteriori*.

A distribuição *a priori* informativa, em geral será: (1) em princípio subjetiva; (2) obtida através de especialistas; (3) obtida a partir de dados anteriores ao experimento principal; (4) obtidas por métodos objetivos; (5) obtidas por conjugação.

Para o problema presente, os parâmetros do modelo são a sensibilidade e a especificidade de cada teste diagnóstico, assim como a prevalência da população. Para a sensibilidade e especificidade de cada teste será considerada *prioris* informativas.

O procedimento para a obtenção destas *prioris* informativas é descrito da seguinte maneira: Intervalos de probabilidade de 95% (isto é, 2,5% em cada limite do intervalo) são determinados para a sensibilidade e especificidade de cada teste. Estes intervalos são obtidos de uma análise relevante da literatura, e opinião clínica.

A densidade *a priori* beta para cada parâmetro do teste é selecionada combinando o centro do intervalo com a média da distribuição beta, dada por  $\frac{\alpha}{\alpha+\beta}$ , e combinando o desvio padrão da distribuição beta, dado por  $\sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}}$ , com um quarto do intervalo total. Estas duas condições definem  $\alpha$  e  $\beta$ . A partir deste procedimento as *prioris* informativas para a sensibilidade e especificidade de cada teste podem ser obtidas. Considere primeiramente o teste sorológico:

- **Teste Sorológico**

- a) Sensibilidade**

Um especialista supõe que o intervalo de probabilidade de 95% de  $\eta$  é de 65% a 95%, isto é,  $P(0,65 \leq \eta \leq 0,95) = 0,95$ . Logo, os hiperparâmetros  $\alpha_1$  e  $\beta_1$  da beta, são calculados através da solução do seguinte sistema:

$$\left\{ \begin{array}{l} \frac{\alpha_1}{\alpha_1+\beta_1} = 0,80 \\ \sqrt{\frac{\alpha_1\beta_1}{(\alpha_1+\beta_1)^2(\alpha_1+\beta_1+1)}} = 0,075 \end{array} \right.$$

onde 0,80  $(0,95 + 0,65/2)$  é o centro do intervalo de 65% a 95%, e 0,075  $(0,95 - 0,65/4)$

é um quarto do intervalo total.

Assim,

$$\frac{\alpha_1}{\alpha_1 + \beta_1} = 0,80 \Rightarrow \alpha_1 = 0,80\alpha_1 + 0,80\beta_1 \Rightarrow 0,20\alpha_1 = 0,80\beta_1 \Rightarrow \alpha_1 = 4\beta_1$$

e

$$\begin{aligned} \sqrt{\frac{\alpha_1\beta_1}{(\alpha_1 + \beta_1)^2(\alpha_1 + \beta_1 + 1)}} &= 0,075 \\ \Rightarrow \frac{4\beta_1^2}{25\beta_1^2(5\beta_1 + 1)} &= 0,005625 \\ \Rightarrow 4\beta_1^2 &= 0,140625\beta_1^2(5\beta_1 + 1) \\ \Rightarrow 4\beta_1^2 &= 0,703125\beta_1^3 + 0,140625\beta_1^2 \\ \Rightarrow 0,703125\beta_1^3 &= 3,859375\beta_1^2 \Rightarrow \beta_1 = 5,49. \end{aligned}$$

Substituindo  $\beta_1 = 5,49$  em  $\alpha_1 = 4\beta_1$ , tem-se que  $\alpha_1 = 21,96$ .

## b) Especificidade

Um especialista supõe que o intervalo de probabilidade de 95% de  $\theta$  é de 35% a 100%, isto é,  $P(0,35 \leq \theta \leq 1,00) = 0,95$ . Logo, os hiperparâmetros  $\alpha_2$  e  $\beta_2$  da beta, são calculados através da solução do seguinte sistema:

$$\begin{cases} \frac{\alpha_2}{\alpha_2 + \beta_2} = 0,675 \\ \sqrt{\frac{\alpha_2\beta_2}{(\alpha_2 + \beta_2)^2(\alpha_2 + \beta_2 + 1)}} = 0,1625 \end{cases}$$

onde 0,675  $(1,00 + 0,35/2)$  é o centro do intervalo de 35% a 100%, e 0,1625  $(1,00 - 0,35/4)$  é um quarto do intervalo total.

Assim,

$$\frac{\alpha_2}{\alpha_2 + \beta_2} = 0,675 \Rightarrow \alpha_2 = 0,675\alpha_2 + 0,675\beta_2 \Rightarrow 0,325\alpha_2 = 0,675\beta_2 \Rightarrow \beta_2 = 0,481481481\alpha_2$$

e

$$\begin{aligned}\sqrt{\frac{\alpha_2\beta_2}{(\alpha_2 + \beta_2)^2(\alpha_2 + \beta_2 + 1)}} &= 0,1625 \\ \Rightarrow \frac{0,481481481\alpha_2^2}{2,19478738\alpha_2^2(1,481481481\alpha_2 + 1)} &= 0,02640625 \\ \Rightarrow 0,481481481\alpha_2^2 &= 0,057956104\alpha_2^2(1,481481481\alpha_2 + 1) \\ \Rightarrow 0,481481481\alpha_2^2 &= 0,085860895\alpha_2^3 + 0,057956104\alpha_2^2 \\ \Rightarrow 0,085860895\alpha_2^3 &= 0,423525377\alpha_2^2 \Rightarrow \alpha_2 = 4,93.\end{aligned}$$

Substituindo  $\alpha_2 = 4,93$  em  $\beta_2 = 0,481481481\alpha_2$ , tem-se que  $\beta_2 = 2,38$ .

Portanto, os hiperparâmetros da distribuição *a priori* beta para a sensibilidade e especificidade do teste sorológico são dados por  $(21,96, 5,49)$  e  $(4,93, 2,38)$ , respectivamente. Da mesma forma os hiperparâmetros da distribuição *a priori* beta para a sensibilidade e especificidade do exame de fezes podem ser obtidos.

- **Exame de Fezes**

- a) **Sensibilidade**

Um especialista supõe que o intervalo de probabilidade de 95% de  $\eta$  é de 5% a 45%, isto é,  $P(0,05 \leq \eta \leq 0,45) = 0,95$ . Logo, os hiperparâmetros  $\alpha_1$  e  $\beta_1$  da beta, são calculados através da solução do seguinte sistema:

$$\begin{cases} \frac{\alpha_1}{\alpha_1 + \beta_1} = 0,25 \\ \sqrt{\frac{\alpha_1\beta_1}{(\alpha_1 + \beta_1)^2(\alpha_1 + \beta_1 + 1)}} = 0,10 \end{cases}$$

onde  $0,25$   $(0,45 + 0,05/2)$  é o centro do intervalo de 5% a 45%, e  $0,10$   $(0,45 - 0,05/4)$  é um quarto do intervalo total.

Assim,

$$\frac{\alpha_1}{\alpha_1 + \beta_1} = 0,25 \Rightarrow \alpha_1 = 0,25\alpha_1 + 0,25\beta_1 \Rightarrow 0,75\alpha_1 = 0,25\beta_1 \Rightarrow \beta_1 = 3\alpha_1$$

e

$$\begin{aligned} \sqrt{\frac{\alpha_1\beta_1}{(\alpha_1 + \beta_1)^2(\alpha_1 + \beta_1 + 1)}} &= 0,10 \\ \Rightarrow \frac{3\alpha_1^2}{16\alpha_1^2(4\alpha_1 + 1)} &= 0,010 \\ \Rightarrow 3\alpha_1^2 &= 0,16\alpha_1^2(4\alpha_1 + 1) \\ \Rightarrow 3\alpha_1^2 &= 0,64\alpha_1^3 + 0,16\alpha_1^2 \\ \Rightarrow 0,64\alpha_1^3 &= 2,84\alpha_1^2 \Rightarrow \alpha_1 = 4,44. \end{aligned}$$

Substituindo  $\alpha_1 = 4,44$  em  $\beta_1 = 3\alpha_1$ , tem-se que  $\beta_1 = 13,31$ .

## b) Especificidade

Um especialista supõe que o intervalo de probabilidade de 95% de  $\theta$  é de 90% a 100%, isto é,  $P(0,90 \leq \theta \leq 1,00) = 0,95$ . Logo, os hiperparâmetros  $\alpha_2$  e  $\beta_2$  da beta, são calculados através da solução do seguinte sistema:

$$\begin{cases} \frac{\alpha_2}{\alpha_2 + \beta_2} = 0,95 \\ \sqrt{\frac{\alpha_2\beta_2}{(\alpha_2 + \beta_2)^2(\alpha_2 + \beta_2 + 1)}} = 0,025 \end{cases}$$

onde 0,95  $(1,00 + 0,90/2)$  é o centro do intervalo de 90% a 100%, e 0,025  $(1,00 - 0,90/4)$  é um quarto do intervalo total.

Assim,

$$\frac{\alpha_2}{\alpha_2 + \beta_2} = 0,95 \Rightarrow \alpha_2 = 0,95\alpha_2 + 0,95\beta_2 \Rightarrow 0,05\alpha_2 = 0,95\beta_2 \Rightarrow \alpha_2 = 19\beta_2$$

e

$$\begin{aligned}
 \sqrt{\frac{\alpha_2 \beta_2}{(\alpha_2 + \beta_2)^2 (\alpha_2 + \beta_2 + 1)}} &= 0,025 \\
 \Rightarrow \frac{19\beta_2^2}{400\beta_2^2(20\beta_2 + 1)} &= 0,000625 \\
 \Rightarrow 19\beta_2^2 &= 0,25\beta_2^2(20\beta_2 + 1) \\
 \Rightarrow 19\beta_2^2 &= 5\beta_2^3 + 0,25\beta_2^2 \\
 \Rightarrow 5\beta_2^3 &= 18,75\beta_2^2 \Rightarrow \beta_2 = 3,75.
 \end{aligned}$$

Substituindo  $\beta_2 = 3,75$  em  $\alpha_2 = 19\beta_2$ , tem-se que  $\alpha_2 = 71,25$ .

Portanto, os hiperparâmetros da distribuição *a priori* beta para a sensibilidade e especificidade do exame de fezes são dados por (4,44, 13,31) e (71,25, 3,75), respectivamente. A Tabela 4.5 apresenta os valores dos hiperparâmetros da distribuição *a priori* beta e o intervalo de probabilidade de 95% para a sensibilidade e especificidade do teste sorológico e exame de fezes.

**Tabela 4.5.** Probabilidade de 95% dos intervalos, e hiperparâmetros das densidades *a priori* Beta para os parâmetros dos testes diagnósticos da infecção causada pelo parasita *strongyloides*.

	Sensibilidade	Especificidade
Teste Sorológico:		
Intervalo	65% a 95%	35% a 100%
Hiperparâmetros da Beta	$\alpha_1 = 21,96, \beta_1 = 5,49$	$\alpha_2 = 4,93, \beta_2 = 2,38$
Exame de Fezes		
Intervalo	5% a 45%	90% a 100%
Hiperparâmetros da Beta	$\alpha_1 = 4,44, \beta_1 = 13,31$	$\alpha_2 = 71,25, \beta_2 = 3,75$

Para a prevalência da doença na população de refugiados de cada teste diagnóstico, será considerada uma *priori* não informativa, pois não se tem nenhuma informação sobre

este parâmetro. Uma densidade uniforme no intervalo  $[0, 1]$  será usada para a distribuição *a priori* da prevalência de *strongyloides* na população de refugiados, que equivale a uma distribuição *a priori* beta com hiperparâmetros  $(\alpha = 1, \beta = 1)$ .

As estimativas dos parâmetros de interesse, obtidas utilizando as *prioris* descritas nesta subseção, serão apresentadas na Seção 4.3.

## 4.2 Infecção de *Strongyloides* em Refugiados do Camboja

Esta aplicação apresentada em Joseph *et al.* (1995), teve por objetivo estimar os parâmetros (sensibilidade, especificidade, valores preditivos positivo e negativo, assim como a prevalência da infecção causada pelo parasita *strongyloides*) de cada teste diagnóstico (teste sorológico e exame de fezes) e da sua combinação em uma situação onde não há nenhum diagnóstico padrão ouro. Conforme Neves *et al.* (2000) este tipo de parasita é encontrado nas fezes de homens ou animais (cães, gatos e macaco) infectados. A transmissão geralmente ocorre através da penetração da pele, devido o contato com alimento contaminado por água de irrigação poluída com fezes ou pelo contato dos pés com o solo contaminado pelas fezes, neste caso a penetração ocorre através da pele dos pés das pessoas que não usam calçados. Este último modo de transmissão parece ser o mais freqüente. A infecção causada pela presença deste parasita no intestino do indivíduo portador geralmente provoca enterite (inflamação nos intestinos); ocorrência de diarreia e vômito.

Os dados consistem de 162 refugiados do Camboja que chegaram em Montreal, Canadá entre julho de 1982 a fevereiro de 1983. Os resultados obtidos pelo teste sorológico e exame de fezes para a infecção causada pelo parasita *strongyloides* apresentam-se na Tabela 4.6.

**Tabela 4.6.** Resultados do teste sorológico e exame de fezes para a infecção causada pelo parasita *strongyloides* em 162 refugiados do Camboja que chegaram em Montreal, Canadá entre julho de 1982 a fevereiro de 1983.

		Exame de Fezes		
		Positivo	Negativo	Total
Sorológico	Positivo	38	87	125
	Negativo	2	35	37
Total		40	122	162

O objetivo nesta seção é estimar a prevalência ( $\pi$ ) na população de refugiados, a sensibilidade ( $\eta$ ), especificidade ( $\theta$ ), e os valores preditivos positivo ( $\gamma$ ) e negativo ( $\delta$ ), assim como as taxas de erros, taxa de um resultado falso negativo ( $\eta^*$ ) e falso positivo ( $\theta^*$ ), para os resultados de cada teste separadamente (teste sorológico e exame de fezes) e para a combinação dos resultados de ambos os testes. Para a estimação destes parâmetros os algoritmos *Gibbs Sampling* apresentados nas Subseções 4.1.1 e 4.1.2 da Seção 4.1, foram implementados utilizando o *Software WinBUGS*. As implementações computacionais destes algoritmos em forma de programa, encontram-se no Apêndice F. Os parâmetros *a priori* para as sensibilidades e especificidades de cada teste diagnóstico, apresentados na Tabela 4.5, e os parâmetros *a priori* para a prevalência da infecção de *strongyloides* na população de refugiados dados no final da Subseção 4.1.3, foram usados.

Os resultados foram obtidos a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas. Os resultados para médias, desvios padrões, medianas e intervalos de credibilidade são apresentados na Seção 4.3.

### 4.3 Resultados

Esta Seção apresenta os resultados obtidos pelo teste sorológico e exame de fezes separadamente, assim como para a combinação destes dois testes. Estes resultados foram

obtidos a partir da implementação dos algoritmos *Gibbs Sampling* apresentados nas Subseções 4.1.1 e 4.1.2, respectivamente.

### 4.3.1 Teste Sorológico

Esta aplicação com o teste sorológico, envolve  $a = 125$  refugiados com resultados positivos e  $b = 37$  refugiados com resultados negativos. Os resultados obtidos a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas, são apresentados na Tabela 4.7. A Tabela 4.8 sumariza os resultados obtidos para os dados latentes  $Y_1$  e  $Y_2$  do teste sorológico.

A Figura 4.1 mostra as distribuições estimadas via *Gibbs Sampling* para os dados latentes usando a informação do teste sorológico. As Figuras 4.2 e 4.3 apresentam, respectivamente, as densidades estimadas via *Gibbs Sampling* e o comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta$  e  $\theta$  usando dados do teste sorológico.

**Tabela 4.7.** Valores estimados via *Gibbs Sampling* para os parâmetros do teste sorológico, com  $a = 125$  e  $b = 37$ .

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,797	0,183	0,266	0,848	0,991
$\eta$	0,828	0,049	0,734	0,827	0,923
$\theta$	0,588	0,188	0,240	0,598	0,914
$\gamma$	0,862	0,182	0,294	0,932	0,997
$\delta$	0,408	0,236	0,025	0,396	0,868
$\eta^*$	0,172	0,049	0,077	0,173	0,266
$\theta^*$	0,412	0,187	0,086	0,402	0,760

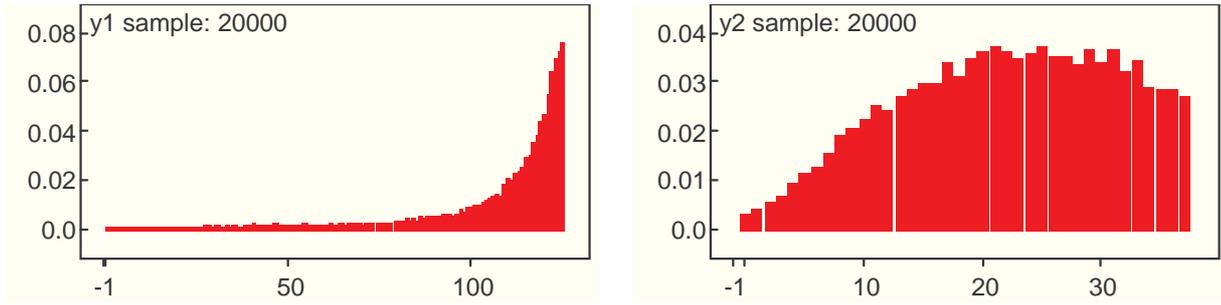
Baseando-se na Tabela 4.7 tem-se que a prevalência ( $\pi$ ), sensibilidade ( $\eta$ ), especificidade ( $\theta$ ), valores preditivos positivo ( $\gamma$ ) e negativo ( $\delta$ ) da infecção causada pelo parasita

*strongyloides* foram em média 79,7%, 82,8%, 58,8%, 86,2% e 40,8%, respectivamente. As taxas de erros para o teste sorológico foram 17,2% para um resultado falso negativo ( $\eta^*$ ) e 41,2% para um resultado falso positivo ( $\theta^*$ ), com isto, verifica-se a importância de considerar erros nos dados. Os intervalos de credibilidade foram (26,6%, 99,1%) para a prevalência da infecção, (73,4%, 92,3%) para sensibilidade, (24,0%, 91,4%) para especificidade, (29,4%, 99,7%) para o valor preditivo positivo, (2,5%, 86,8%) para o valor preditivo negativo, (7,7%, 26,6%) para a taxa do resultado falso negativo e (8,6%, 76,0%) para a taxa do resultado falso positivo. Com respeito aos intervalos de credibilidade, a amplitude do intervalo para a prevalência foi de 72,5%, 18,9%, 67,4%, 70,3%, 84,3%, 18,9% e 67,4% para a sensibilidade, especificidade, valor preditivo positivo, valor preditivo negativo, taxa do resultado falso negativo e taxa do resultado falso positivo, respectivamente.

Em relação aos refugiados do Camboja, o teste sorológico mostrou-se mais sensível (82,8%) para diagnosticar a infecção, e menos específico (58,8%) para não diagnosticar a infecção. Conseqüentemente, produziu uma taxa do resultado falso negativo (17,2%) menor do que a taxa do resultado falso positivo (41,2%), e valor preditivo negativo (40,8%) melhor do que o valor preditivo positivo (86,2%), pois quanto mais sensível o teste, melhor seu valor preditivo negativo, isto é, maior a segurança de que um refugiado com resultado negativo não tenha a infecção.

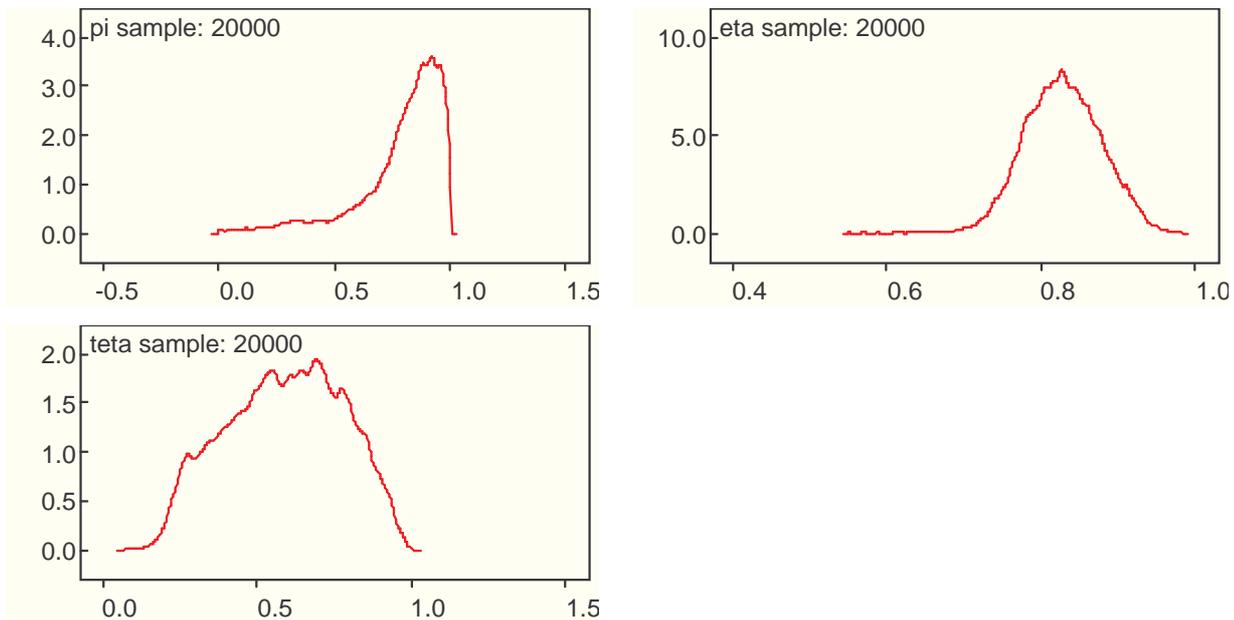
**Tabela 4.8.** Valores estimados via *Gibbs Sampling* para os dados latentes do teste sorológico, com  $a = 125$  e  $b = 37$ .

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	107,7	22,93	36,0	117,0	125,0
$Y_2$	21,92	9,117	4,0	22,0	37,0



**Figura 4.1.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1$  e  $Y_2$  do teste sorológico.

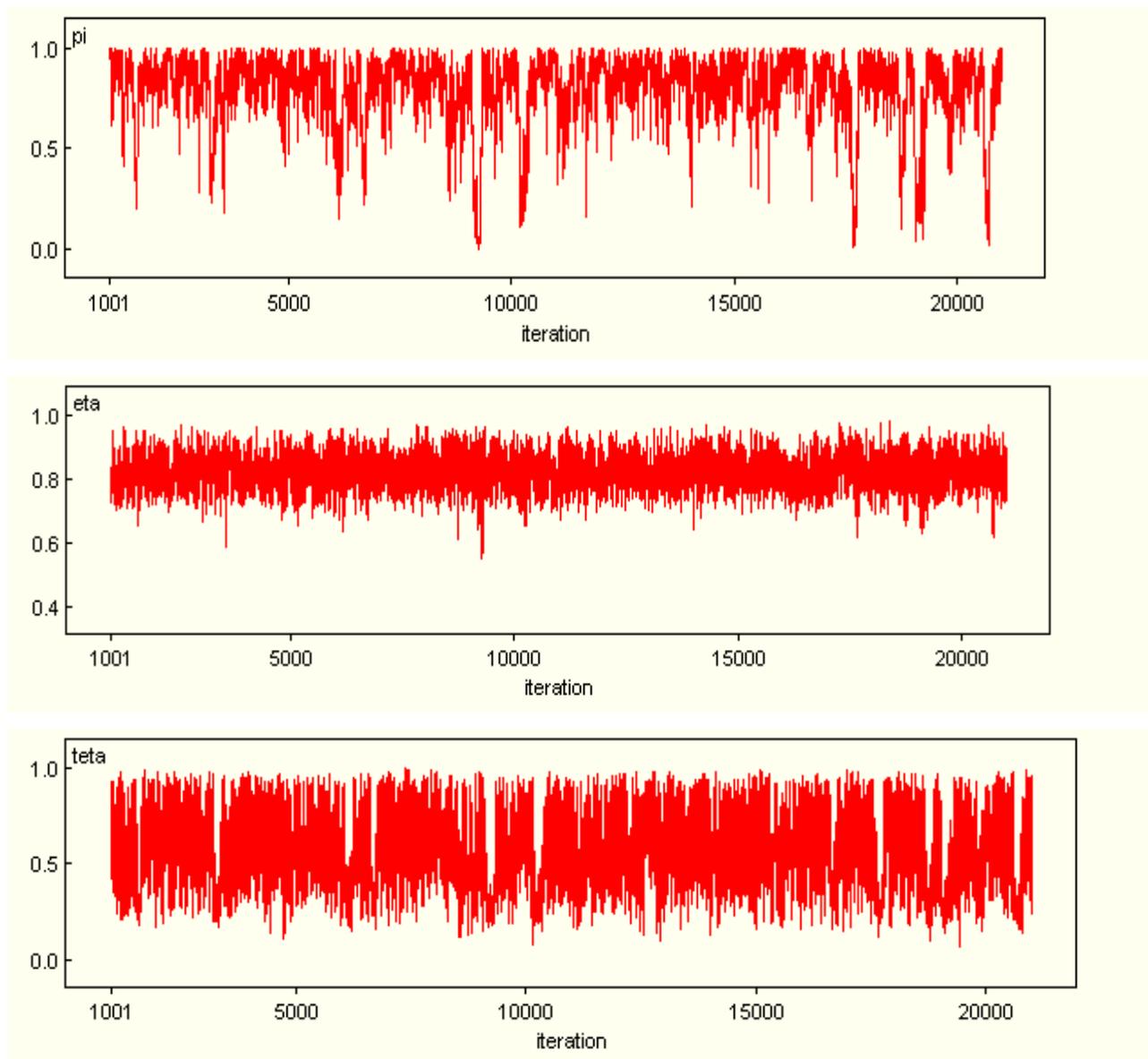
A partir da Tabela 4.8 e da Figura 4.1 tem-se que aproximadamente 108 refugiados dos 125 com resultados positivos para o teste sorológico tenham resultados verdadeiros positivos, e que aproximadamente 22 refugiados dos 37 com resultados negativos para o teste sorológico tenham resultados verdadeiros positivos.



**Figura 4.2.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do teste sorológico.

Como é evidente da Tabela 4.7 e da Figura 4.2, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da prevalência usando dados do

teste sorológico foi 79,7%, embora o intervalo de credibilidade de 95% foi de (26,6%, 9961%), esta mesma situação também ocorre com os outros parâmetros.



**Figura 4.3.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do teste sorológico.

Através da Figura 4.3, tem-se uma indicação de convergência para todos os parâmetros do teste sorológico.

### 4.3.2 Exame de Fezes

Esta aplicação com o exame de fezes, envolve  $a = 40$  refugiados com resultados positivos e  $b = 122$  refugiados com resultados negativos. Os resultados obtidos a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas, são apresentados na Tabela 4.9. A Tabela 4.10 sumariza os resultados obtidos para os dados latentes  $Y_1$  e  $Y_2$  do exame de fezes.

A Figura 4.4 mostra as distribuições estimadas via *Gibbs Sampling* para os dados latentes usando a informação do exame de fezes. As Figuras 4.5 e 4.6 apresentam, respectivamente, as densidades estimadas via *Gibbs Sampling* e o comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta$  e  $\theta$  usando dados do exame de fezes.

**Tabela 4.9.** Valores estimados via *Gibbs Sampling* para os parâmetros do exame de fezes, com  $a = 40$  e  $b = 122$ .

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,743	0,162	0,407	0,758	0,989
$\eta$	0,313	0,067	0,209	0,303	0,473
$\theta$	0,948	0,026	0,885	0,952	0,986
$\gamma$	0,938	0,063	0,777	0,957	0,998
$\delta$	0,318	0,195	0,014	0,303	0,713
$\eta^*$	0,687	0,067	0,527	0,697	0,791
$\theta^*$	0,052	0,026	0,014	0,048	0,115

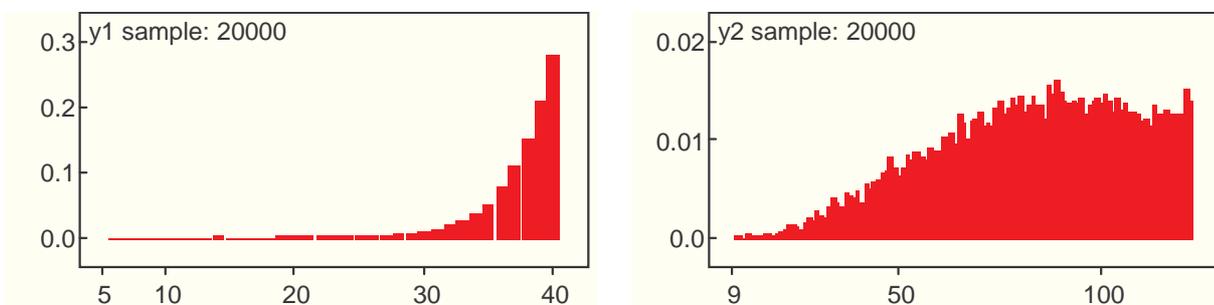
Baseando-se na Tabela 4.9 tem-se que a prevalência ( $\pi$ ), sensibilidade ( $\eta$ ), especificidade ( $\theta$ ), valores preditivos positivo ( $\gamma$ ) e negativo ( $\delta$ ) da infecção causada pelo parasita *strongyloides* foram em média 74,3%, 31,3%, 94,8%, 93,8% e 31,8%, respectivamente. As taxas de erros para o exame de fezes foram 68,7% para um resultado falso negativo ( $\eta^*$ ) e 5,2% para um resultado falso positivo ( $\theta^*$ ), desta forma, verifica-se a importância de

considerar erros nos dados. Os intervalos de credibilidade foram (40,7%, 98,9%) para a prevalência da infecção, (20,9%, 47,3%) para sensibilidade, (88,5%, 98,6%) para especificidade, (77,7%, 99,8%) para o valor preditivo positivo, (1,4%, 71,3%) para o valor preditivo negativo, (52,7%, 79,1%) para a taxa do resultado falso negativo e (1,4%, 11,5%) para a taxa do resultado falso positivo. Com respeito aos intervalos de credibilidade, a amplitude do intervalo para a prevalência foi de 58,2%, 26,4%, 10,1%, 22,1%, 69,9%, 26,4% e 10,1% para a sensibilidade, especificidade, valor preditivo positivo, valor preditivo negativo, taxa do resultado falso negativo e taxa do resultado falso positivo, respectivamente.

Em relação aos refugiados do Camboja, o exame de fezes mostrou-se mais específico (94,8%) para não diagnosticar a infecção, e menos sensível (31,3%) para diagnosticar a infecção. Conseqüentemente, produziu uma taxa do resultado falso negativo (68,7%) maior do que a taxa do resultado falso positivo (5,2%), e valor preditivo positivo (93,8%) melhor do que o valor preditivo negativo (31,8%), pois quanto mais específico o teste, melhor seu valor preditivo positivo, isto é, maior a segurança de que um refugiado com resultado positivo tenha a infecção.

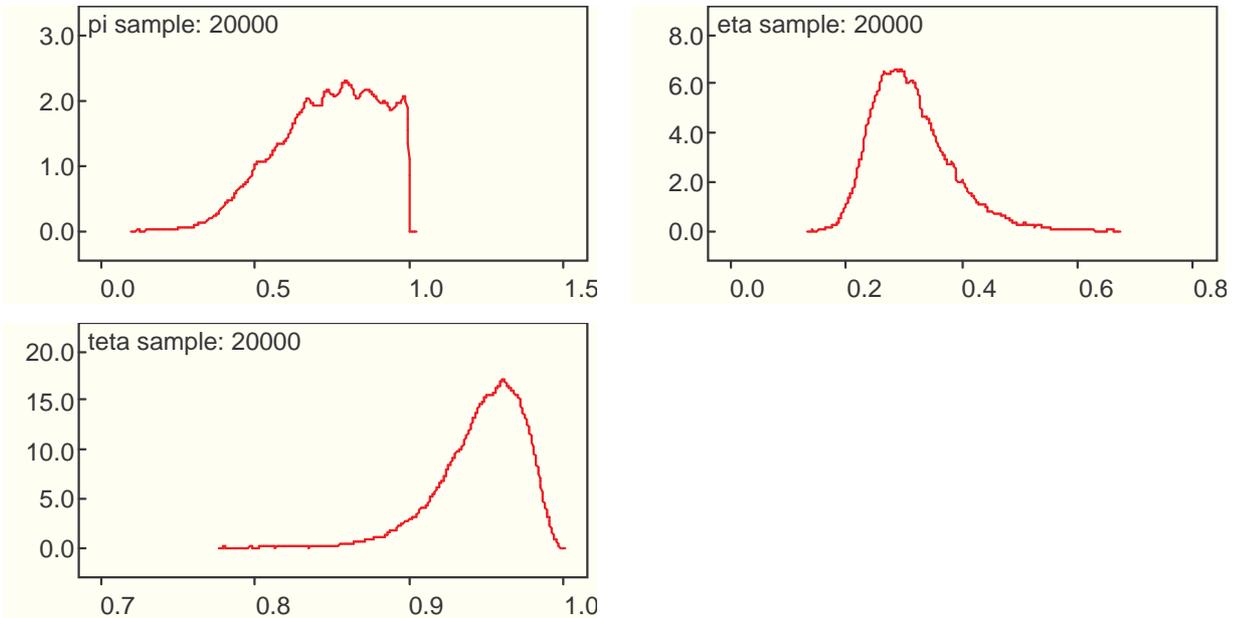
**Tabela 4.10.** Valores estimados via *Gibbs Sampling* para os dados latentes do exame de fezes, com  $a = 40$  e  $b = 122$ .

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	37,56	2,909	30,0	38,0	40,0
$Y_2$	83,25	2,25	34,0	85,0	121,0



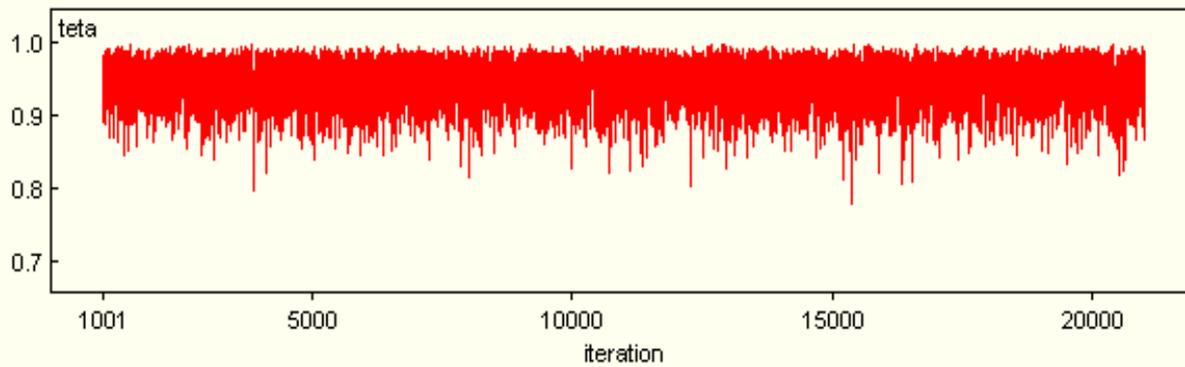
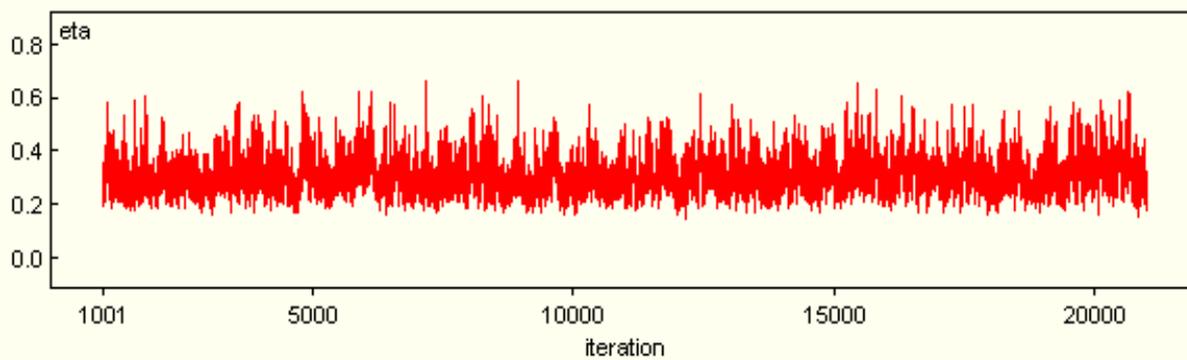
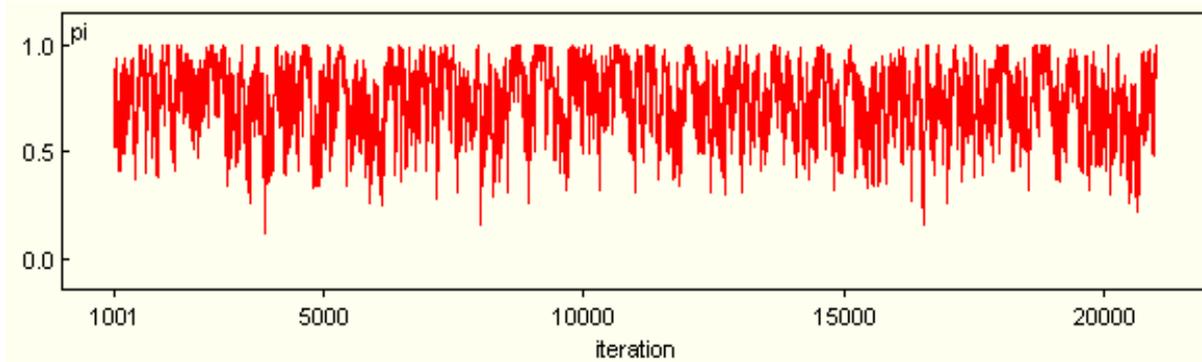
**Figura 4.4.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1$  e  $Y_2$  do exame de fezes.

A partir da Tabela 4.10 e da Figura 4.4 tem-se que aproximadamente 38 refugiados dos 40 com resultados positivos para o exame de fezes tenham resultados verdadeiros positivos, e que aproximadamente 83 refugiados dos 122 com resultados negativos para o exame de fezes tenham resultados verdadeiros positivos.



**Figura 4.5.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do exame de fezes.

Como é evidente da Tabela 4.9 e da Figura 4.5, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da prevalência usando dados do exame de fezes foi 74,3%, embora o intervalo de credibilidade de 95% foi de (40,7%, 98,9%), esta mesma situação também ocorre com os outros parâmetros.



**Figura 4.6.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do exame de fezes.

Através da Figura 4.6, tem-se uma indicação de convergência para todos os parâmetros do exame de fezes.

### 4.3.3 Combinação do Teste Sorológico e Exame de Fezes

Esta aplicação com a combinação do teste sorológico e exame de fezes, envolve  $u = 38$  refugiados com resultados positivos em ambos os testes,  $v = 87$  refugiados com resultados positivos no sorológico mas negativos no exame de fezes,  $w = 2$  refugiados com resultados negativos no sorológico mas positivos no exame de fezes, e  $x = 35$  refugiados com resultados negativos em ambos os testes.

Os resultados obtidos a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas, são apresentados na Tabela 4.11. A Tabela 4.12 sumariza os resultados obtidos para os dados latentes  $Y_1, Y_2, Y_3$  e  $Y_4$  da combinação destes dois testes.

A Figura 4.7 mostra as distribuições estimadas via *Gibbs Sampling* para os dados latentes usando a informação da combinação dos testes. As Figuras 4.8 e 4.9 apresentam, respectivamente, as densidades estimadas via *Gibbs Sampling* e o comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  usando dados da combinação dos dois testes.

É importante salientar que quando os testes são combinados, os parâmetros (sensibilidade, especificidade, valor preditivo positivo e negativo), por exemplo, do teste sorológico são estimados considerando os dados do exame de fezes, e vice-versa. Para a presente aplicação da combinação de ambos os testes os parâmetros  $\eta_1, \theta_1, \gamma_1$  e  $\delta_1$  são do teste sorológico e  $\eta_2, \theta_2, \gamma_2$  e  $\delta_2$  do exame de fezes.

**Tabela 4.11.** Valores estimados via *Gibbs Sampling* para os parâmetros da combinação do teste sorológico e exame de fezes, com  $u = 38$ ,  $v = 87$ ,  $w = 2$  e  $x = 35$ .

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,750	0,101	0,516	0,764	0,919
$\eta_1$	0,883	0,042	0,789	0,888	0,953
$\theta_1$	0,662	0,152	0,371	0,667	0,928
$\gamma_1$	0,875	0,097	0,625	0,902	0,988
$\delta_1$	0,632	0,160	0,245	0,657	0,874
$\eta_2$	0,313	0,053	0,224	0,308	0,433
$\theta_2$	0,957	0,021	0,906	0,960	0,989
$\gamma_2$	0,955	0,033	0,868	0,963	0,993
$\delta_2$	0,315	0,125	0,103	0,299	0,600
$\eta_1^*$	0,117	0,042	0,047	0,112	0,211
$\theta_1^*$	0,338	0,152	0,072	0,333	0,629
$\eta_2^*$	0,687	0,053	0,567	0,692	0,776
$\theta_2^*$	0,043	0,021	0,011	0,039	0,094

Baseando-se na Tabela 4.11 tem-se que a prevalência ( $\pi$ ) da infecção causada pelo parasita *strongyloides* foi em média 75,0%, e o intervalo de credibilidade (51,6%, 91,9%) com amplitude de 40,3%; a sensibilidade ( $\eta_1$ ), especificidade ( $\theta_1$ ), valores preditivos positivo ( $\gamma_1$ ) e negativo ( $\delta_1$ ) estimados utilizando os dados do teste sorológico na presença dos dados do exame de fezes foram em média 88,3%, 66,2%, 87,5% e 63,2%, com intervalos de credibilidade de (78,9%, 95,3%), (37,1%, 92,8%), (62,5%, 98,8%) e (24,5%, 87,4%), de amplitudes 16,4%, 55,7%, 36,3% e 62,9%, respectivamente. As taxas de erros foram 11,7% para um resultado falso negativo ( $\eta_1^*$ ) e 33,8% para um resultado falso positivo ( $\theta_1^*$ ), com intervalos de (4,7%, 21,1%) e (7,2%, 62,9%), e amplitudes de 16,4% e 55,7%.

Em relação aos refugiados do Camboja, o teste sorológico mostrou-se mais sensível (88,3%) para diagnosticar a infecção, e menos específico (66,2%) para não diagnosticar a infecção. Conseqüentemente, produziu uma taxa do resultado falso negativo (11,7%)

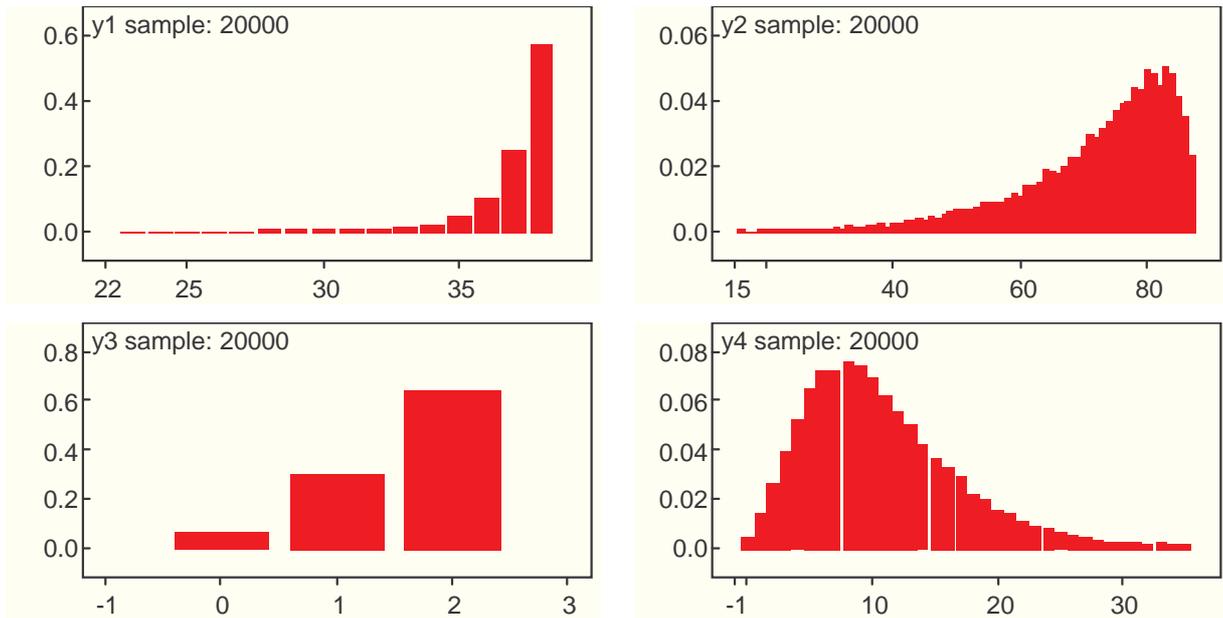
menor do que a taxa do resultado falso positivo (33,8%), e valor preditivo negativo (63,2%) melhor do que o valor preditivo positivo (87,5%), pois quanto mais sensível o teste, melhor seu valor preditivo negativo, isto é, maior a segurança de que um refugiado com resultado negativo não tenha a infecção.

Já o exame de fezes apresentou sensibilidade ( $\eta_2$ ), especificidade ( $\theta_2$ ), valores preditivos positivo ( $\gamma_2$ ) e negativo ( $\delta_2$ ) de 31,3%, 95,7%, 95,5% e 31,5%, com intervalos de credibilidade de (22,4%, 43,3%), (90,6%, 98,9%), (86,8%, 99,3%) e (10,3%, 60,0%), de amplitudes 20,9%, 8,3%, 12,5% e 49,7%, respectivamente. As taxas de erros foram 68,7% para um resultado falso negativo ( $\eta_2^*$ ) e 4,3% para um resultado falso positivo ( $\theta_2^*$ ), com intervalos de (56,7%, 77,6%) e (1,10%, 9,40%), e amplitudes de 20,9% e 8,3%.

Neste caso, o exame de fezes mostrou-se mais específico (95,7%) para não diagnosticar a infecção, e menos sensível (31,3%) para diagnosticar a infecção. Conseqüentemente, produziu uma taxa do resultado falso negativo (68,7%) maior do que a taxa do resultado falso positivo (4,3%), e valor preditivo positivo (95,5%) melhor do que o valor preditivo negativo (31,5%), pois quanto mais específico o teste, melhor seu valor preditivo positivo, isto é, maior a segurança de que um refugiado com resultado positivo tenha a infecção.

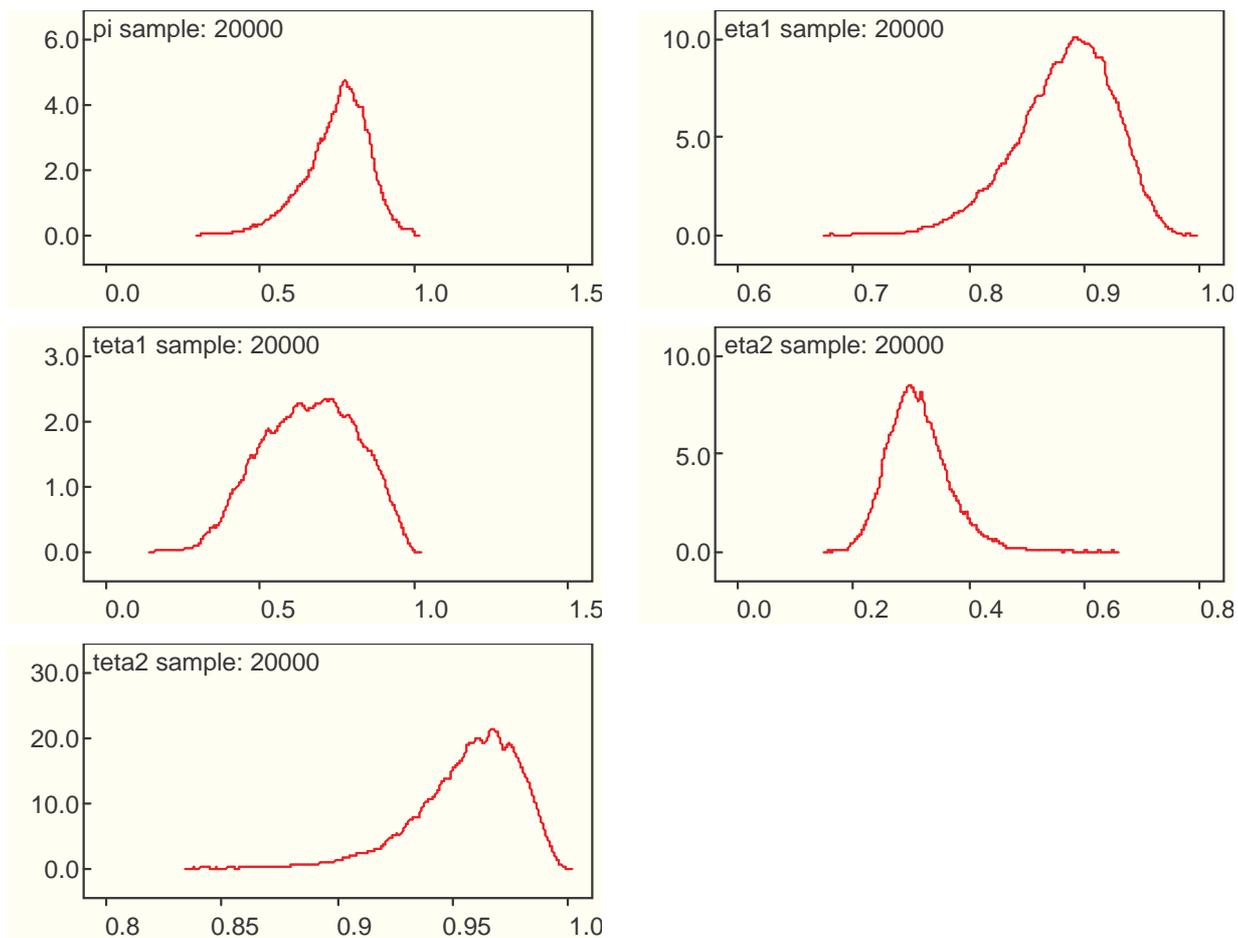
**Tabela 4.12.** Valores estimados via *Gibbs Sampling* para os dados latentes da combinação do teste sorológico e exame de fezes, com  $u = 38$ ,  $v = 87$ ,  $w = 2$  e  $x = 35$ .

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	37,23	1,217	34,0	38,0	38,0
$Y_2$	72,62	11,53	43,0	76,0	86,0
$Y_3$	1,583	0,603	0,0	2,0	2,0
$Y_4$	10,63	6,07	2,0	10,0	25,0



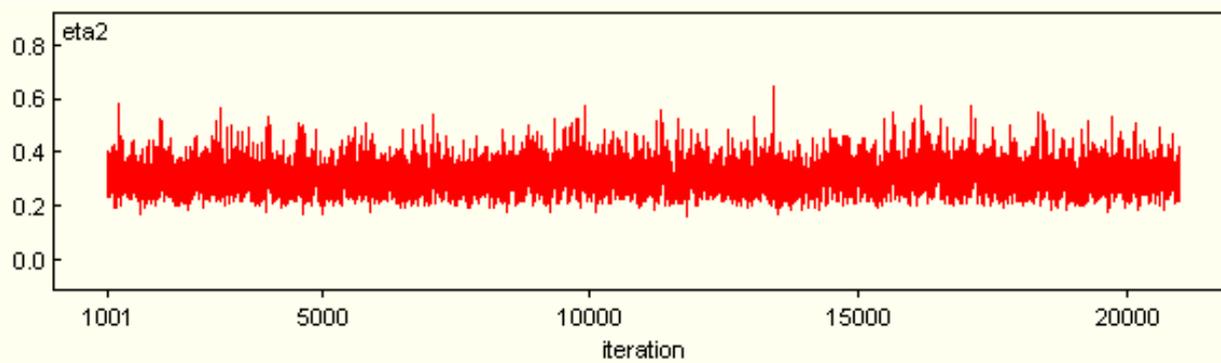
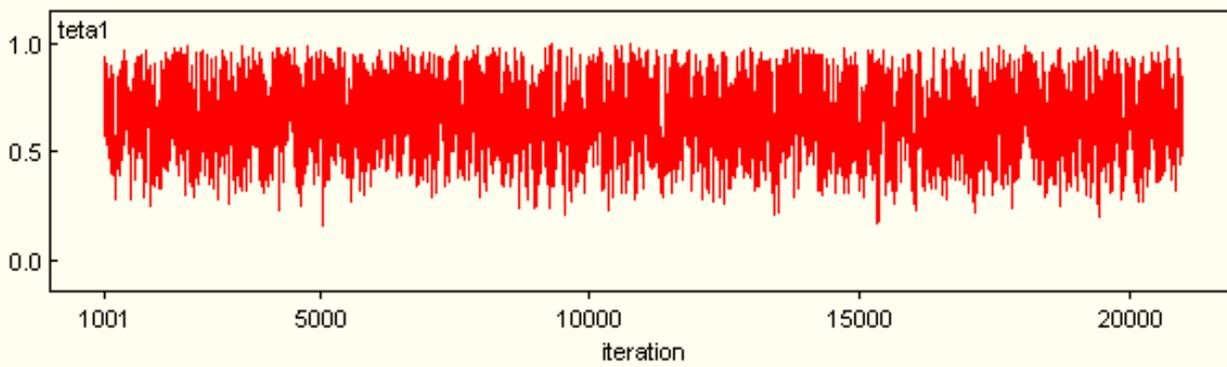
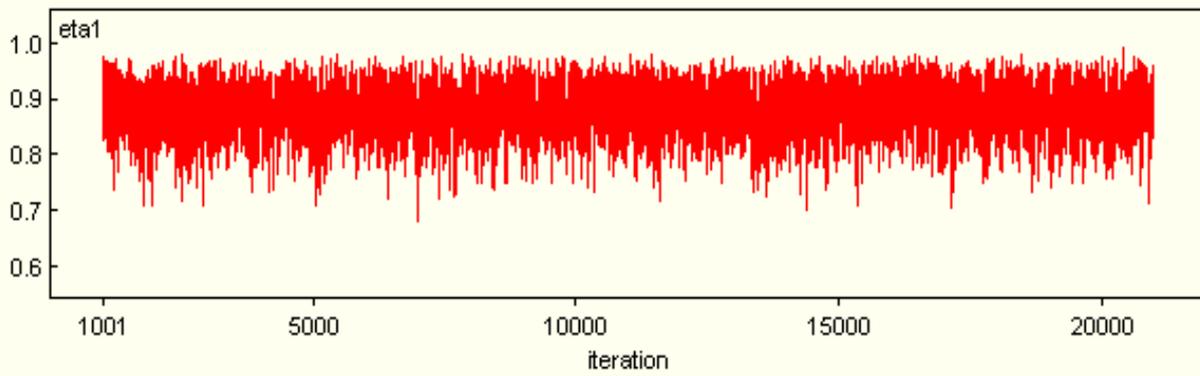
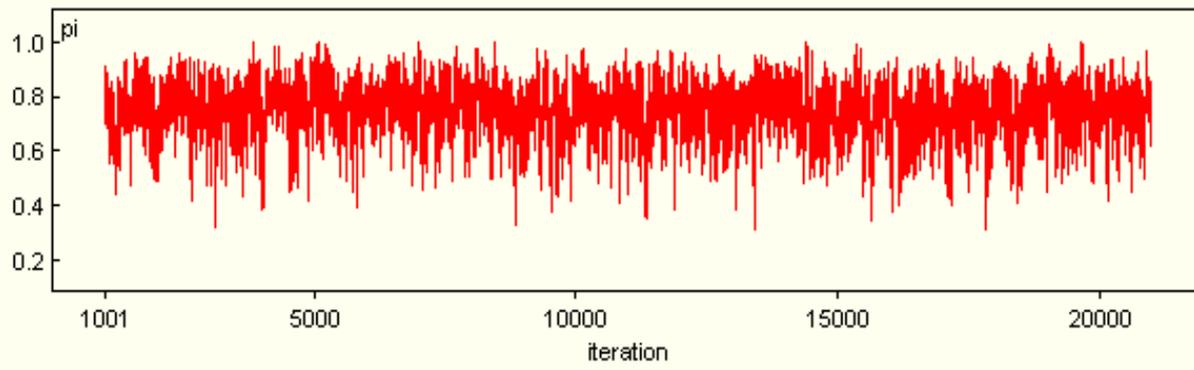
**Figura 4.7.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1, Y_2, Y_3$  e  $Y_4$  da combinação dos dois testes.

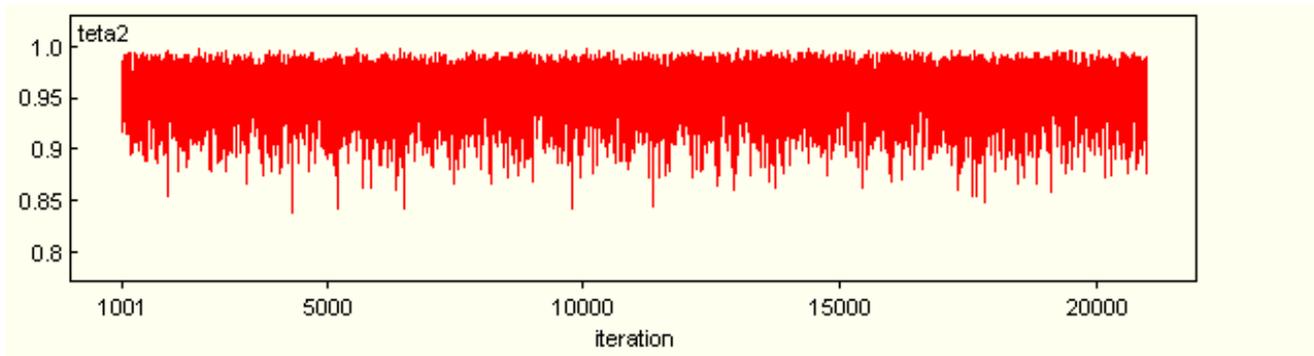
A partir da Tabela 4.12 e da Figura 4.7 tem-se que aproximadamente 37 refugiados dos 38 com resultados positivos em ambos os testes tenham resultados verdadeiros positivos. É provável que aproximadamente 73 refugiados dos 87 com resultados positivos para o teste sorológico mas com exames de fezes negativos tenham resultados verdadeiros positivos. A maioria dos refugiados com teste positivo no exame de fezes mas negativo para o teste sorológico são prováveis terem resultados verdadeiros positivos. Finalmente, do dado latente e histograma para  $Y_4$ , parece provável que aproximadamente 11 refugiados dos 35 com resultados negativos em ambos os testes tenham resultados verdadeiros positivos.



**Figura 4.8.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta_1$ ,  $\theta_1$ ,  $\eta_2$  e  $\theta_2$  da combinação dos testes.

Como é evidente da Tabela 4.11 e da Figura 4.8, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da prevalência usando dados da combinação de ambos os testes foi 75,0%, embora o intervalo de credibilidade de 95% foi de (51,6%, 91,9%), esta mesma situação também ocorre com os outros parâmetros.





**Figura 4.9.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos dois testes.

Através da Figura 4.9, tem-se uma indicação de convergência para todos os parâmetros da combinação do teste sorológico e exame de fezes.

## 4.4 A Indicação de Convergência

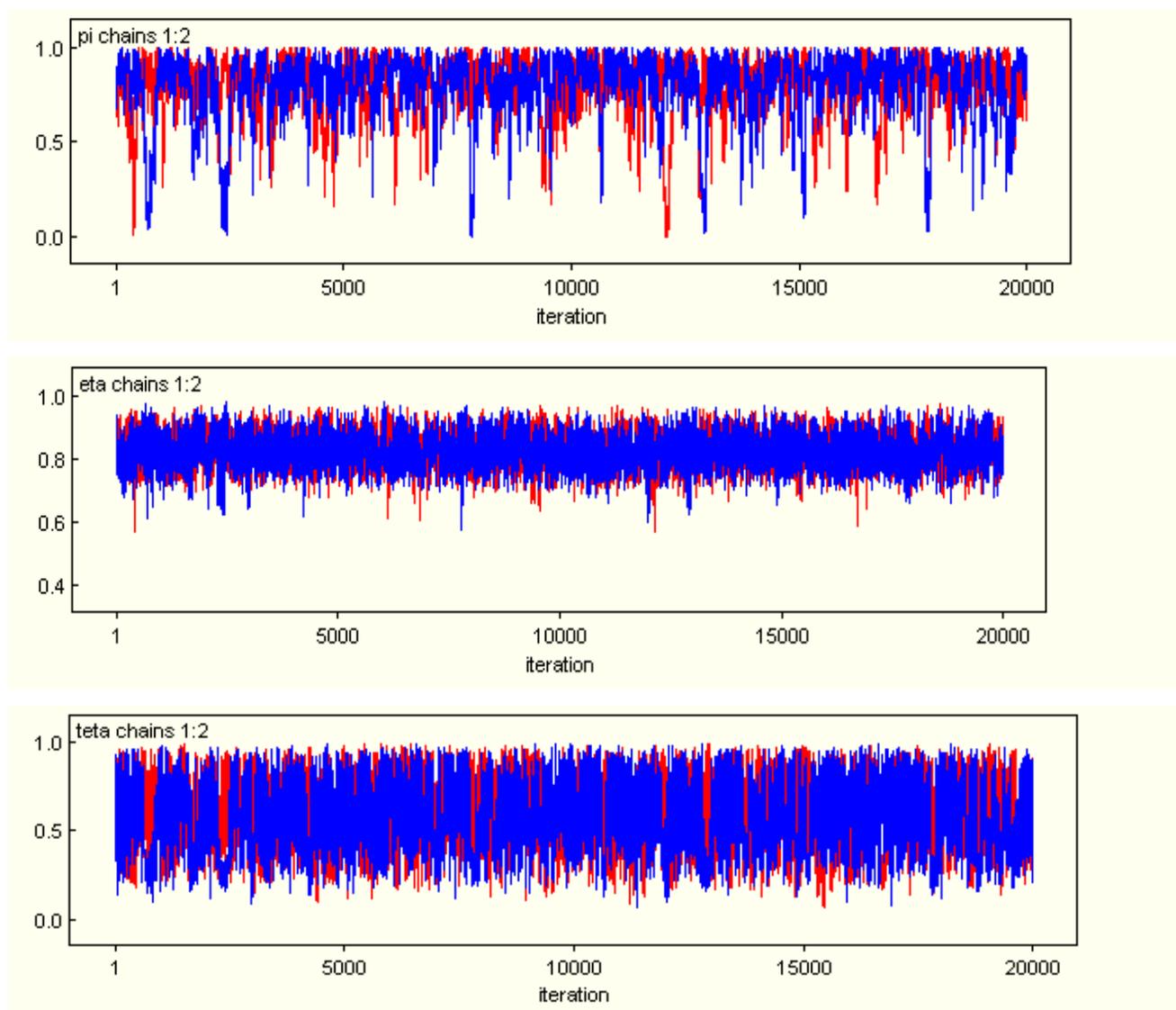
Como exemplo, para análise da indicação de convergência, os resultados da aplicação do diagnóstico de Gelman & Rubin (1992), descrito no Apêndice D, são apresentados para os dados da infecção causada pelo parasita *strongyloides* em 162 refugiados do Camboja que chegaram em Montreal, Canadá entre julho de 1982 a fevereiro de 1983. Estes resultados são apresentados para o teste sorológico, exame de fezes e para a combinação destes dois testes.

Para esta análise foram geradas duas cadeias, com 20000 valores cada uma.

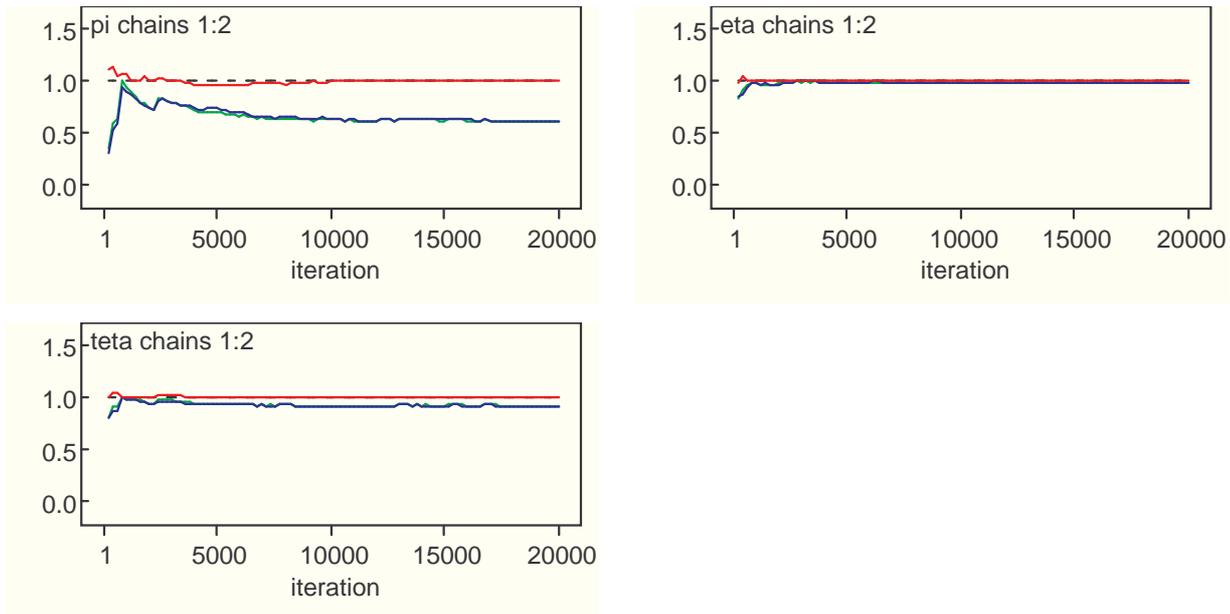
- **Teste Sorológico**

A Figura 4.10 ilustra o comportamento das trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta$  e  $\theta$  do teste sorológico, com as iterações iniciais não descartadas. Através desta Figura, verifica-se que a indicação de convergência é quase

imediate, ou seja, mesmo antes de 1000 iterações já existe a indicação de convergência. A Figura 4.11 mostra os gráficos fornecidos pelo *Software WinBUGS* para o diagnóstico de Gelman & Rubin.



**Figura 4.10.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do teste sorológico.



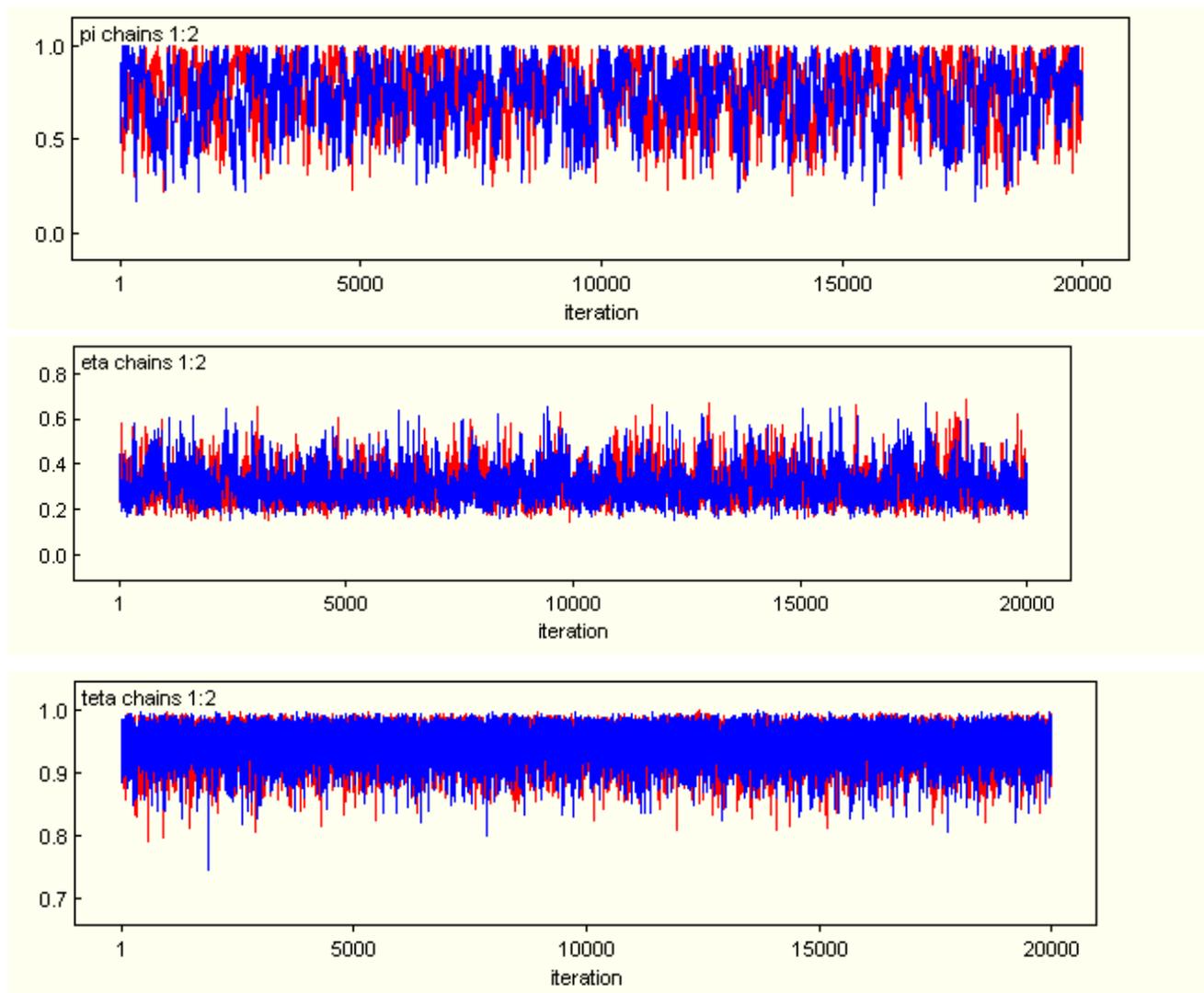
**Figura 4.11.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do teste sorológico. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas, a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = \text{combinada}/\text{individual}$ .

Através da Figura 4.11, verifica-se que a largura dos intervalos nas cadeias combinadas e a largura média dos intervalos dentro de cada cadeia individual estão praticamente sobrepostas. O fator de redução  $R$  converge para 1 à medida que aumenta o número de iterações, tendo assim a indicação de convergência para todos os parâmetros do teste sorológico.

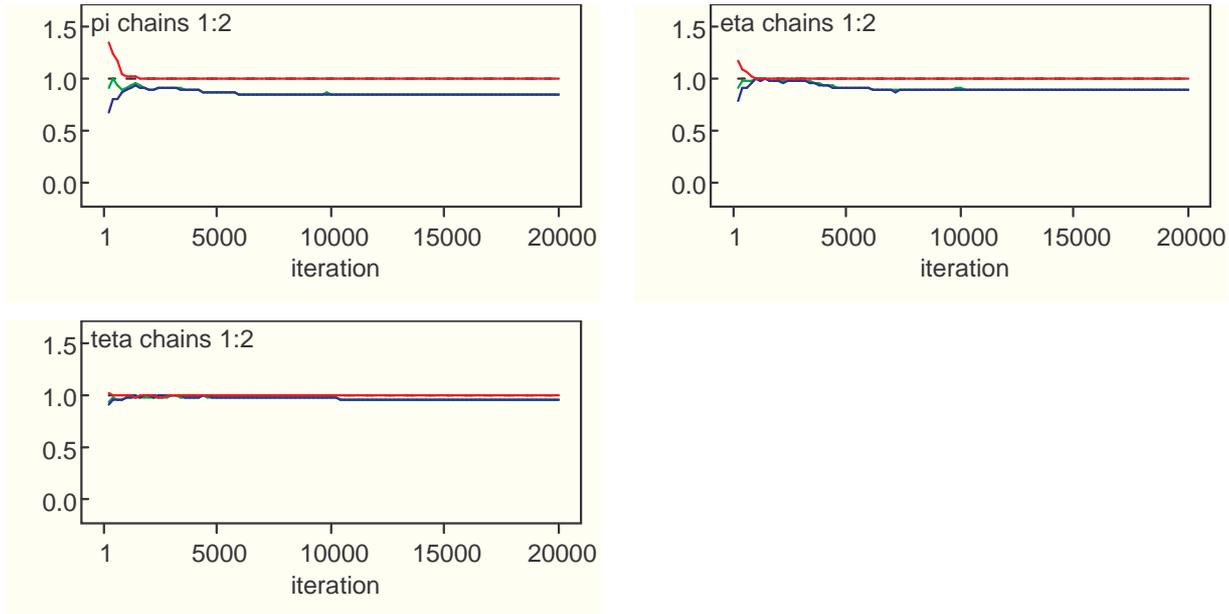
- **Exame de Fezes**

A Figura 4.12 ilustra o comportamento das trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do exame de fezes, com as iterações iniciais não descartadas. Através desta Figura, verifica-se que a indicação de convergência é quase imediata, ou seja, mesmo antes de 1000 iterações já existe a indicação de convergência.

A Figura 4.13 mostra os gráficos fornecidos pelo *Software WinBUGS* para o diagnóstico de Gelman & Rubin.



**Figura 4.12.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do exame de fezes.



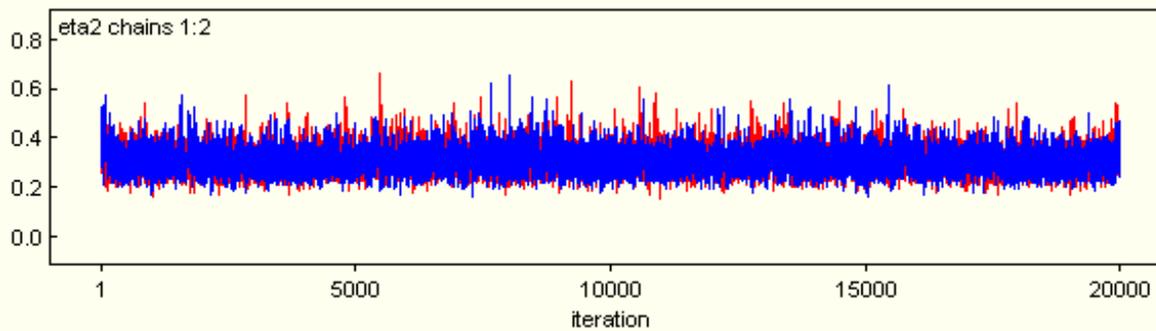
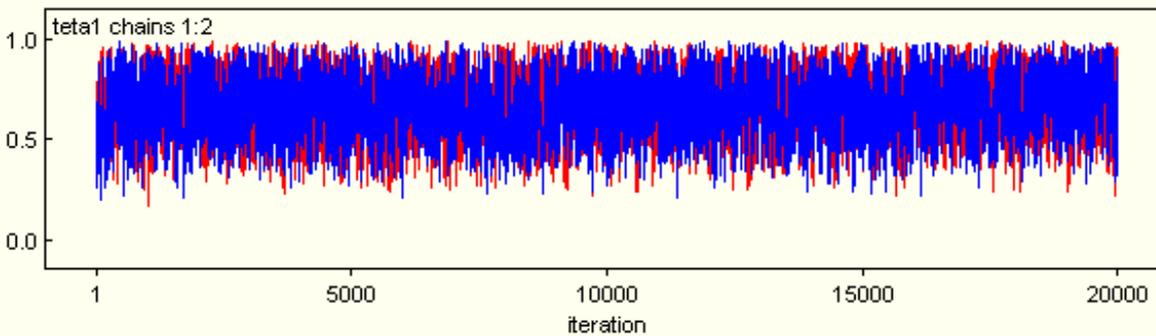
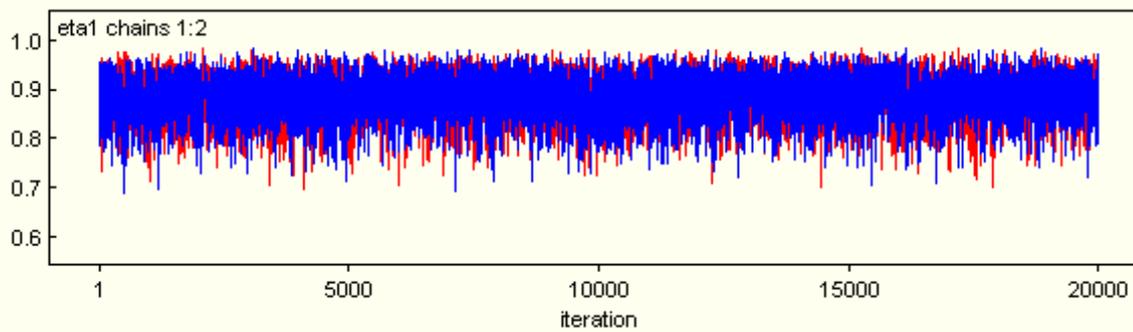
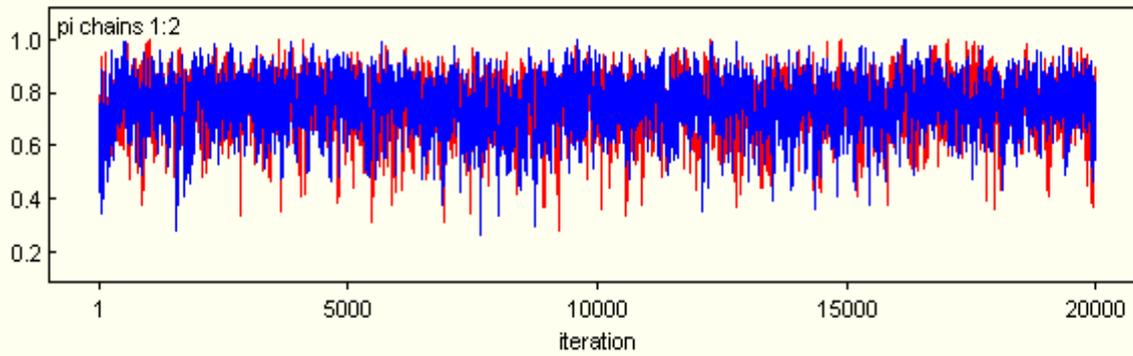
**Figura 4.13.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do exame de fezes. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas, a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = \text{combinada}/\text{individual}$ .

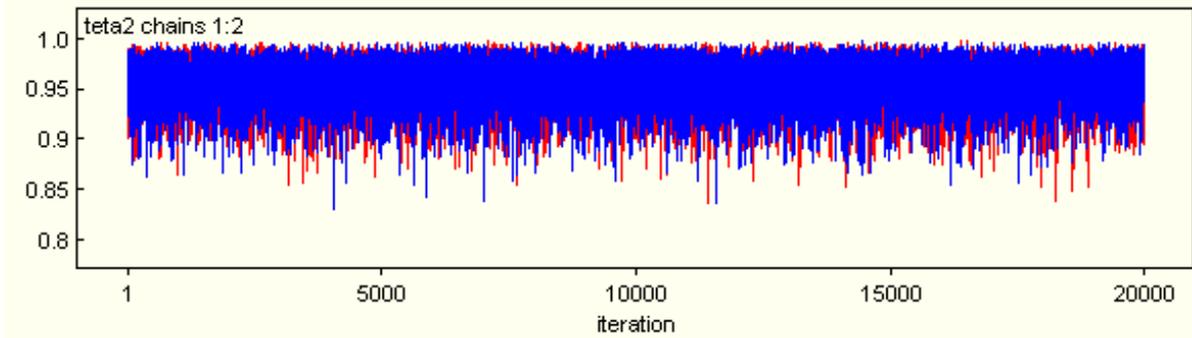
Através da Figura 4.13, verifica-se que a largura dos intervalos nas cadeias combinadas e a largura média dos intervalos dentro de cada cadeia individual estão praticamente sobrepostas. O fator de redução  $R$  converge para 1 à medida que aumenta o número de iterações, tendo assim a indicação de convergência para todos os parâmetros do exame de fezes.

- **Combinação do teste sorológico e exame fezes**

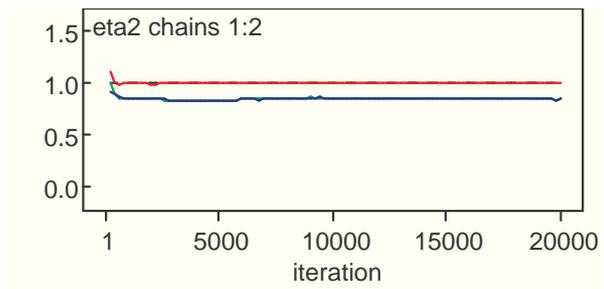
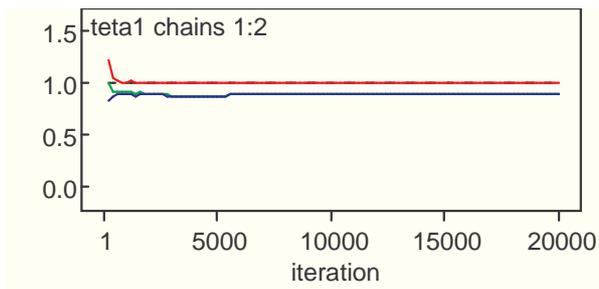
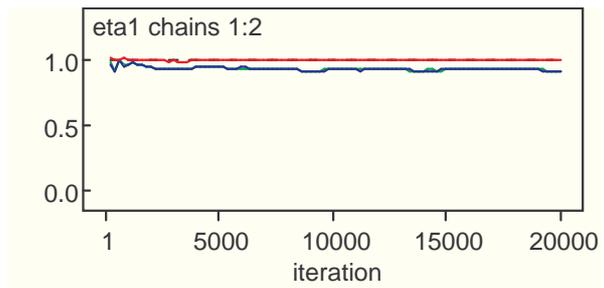
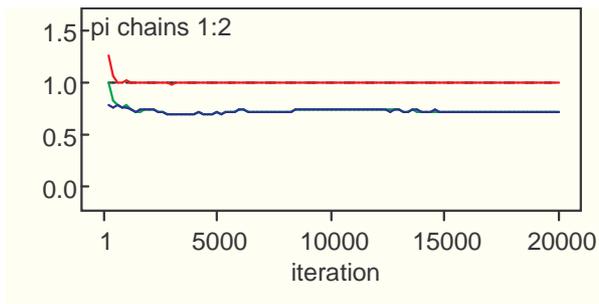
A Figura 4.14 ilustra o comportamento das trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi$ ,  $\eta_1$ ,  $\theta_1$ ,  $\eta_2$  e  $\theta_2$  da combinação do teste sorológico e exame de fezes, com as iterações iniciais não descartadas. Através desta Figura, verifica-se que a indicação de convergência é quase imediata, ou seja, mesmo antes de 1000 iterações

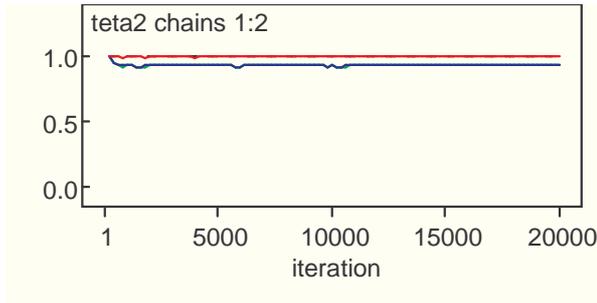
já existe a indicação de convergência. A Figura 4.15 mostra os gráficos fornecidos pelo *Software WinBUGS* para o diagnóstico de Gelman & Rubin.





**Figura 4.14.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos testes.





**Figura 4.15.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos dois testes. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas, a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = combinada/individual$ .

Através da Figura 4.15, verifica-se que a largura dos intervalos nas cadeias combinadas e a largura média dos intervalos dentro de cada cadeia individual estão praticamente sobrepostas. O fator de redução  $R$  converge para 1 à medida que aumenta o número de iterações, tendo assim a indicação de convergência para todos os parâmetros da combinação dos dois testes.

Como para todos os parâmetros estimados via implementação do algoritmo *Gibbs Sampling* baseado no procedimento de Joseph *et al.* (1995), para o teste sorológico, exame de fezes e combinação de ambos, realizou-se a indicação de convergência, pode-se utilizar as estatísticas fornecidas pelo *Software WinBUGS* para interpretar a presente aplicação.

## 4.5 Comentários

A estrutura bayesiana desenvolvida no Capítulo 4 baseou-se no método proposto por Joseph *et al.* (1995), para obter o algoritmo *Gibbs Sampling* de um único teste diagnóstico, assim como de dois testes diagnósticos. Estes algoritmos foram implementados para a

estimação dos parâmetros de cada teste diagnóstico (teste sorológico e exame de fezes) separadamente, e da combinação dos dois testes. Esta metodologia foi aplicada aos dados de Joseph *et al.* (1995) relacionados com 162 refugiados do Camboja que chegaram em Montreal, Canadá entre julho de 1982 a fevereiro de 1983 para detecção da infecção causada pelo parasita *strongyloides*, com distribuição *a priori* não informativa para a prevalência da infecção e *prioris* informativas para as sensibilidades e especificidades do teste sorológico e exame de fezes. A partir das estimativas dos parâmetros de interesse para o teste sorológico e exame de fezes separadamente, como para a combinação destes dois testes observou-se que:

1. Para os refugiados do Camboja, considerando cada teste separadamente, as prevalências da infecção causada pelo parasita *strongyloides* encontradas pelo teste sorológico e exame de fezes foram de 79,7% e 74,3%; sensibilidades de 82,8% e 31,3%, especificidades de 58,8% e 94,8%, valores preditivos positivos de 86,2% e 93,8%, valores preditivos negativos de 40,8% e 31,8%, taxas de erros dos resultados falsos negativos de 17,2% e 68,7%, e taxas de erros dos resultados falsos positivos de 41,2% e 5,2%, respectivamente. As amplitudes dos intervalos de credibilidade dos refugiados para as prevalências da infecção, segundo cada teste diagnóstico, foram de 72,5% e 58,2%; para as sensibilidades de 18,9% e 26,4%, especificidades de 67,4% e 10,1%, valores preditivos positivos de 70,3% e 22,1%, valores preditivos negativos de 84,3% e 69,9%, taxas de erros de resultados falsos negativos de 18,9% e 26,4%, e taxas de erros dos resultados falsos positivos de 67,4% e 10,1%, respectivamente.

Nota-se que quando cada teste diagnóstico é considerado separadamente, o sorológico estima prevalência, sensibilidade, valor preditivo negativo e taxa de erro do resultado falso positivo da infecção superior ao exame de fezes, mas com especificidade, valor preditivo positivo e taxa de erro do resultado falso negativo inferior.

Em relação aos refugiados, o teste sorológico mostrou-se mais sensível (82,8%) para diagnosticar a infecção do que o exame de fezes (31,3%), e menos específico (58,8%) para não diagnosticar a infecção do que o exame de fezes (94,8%). Com isso, estes valores

apoiam a afirmação de que os exames de fezes subestimam a prevalência da população, produzindo baixa sensibilidade, alta especificidade e melhor valor preditivo positivo, enquanto o teste sorológica resulta em superestimação, exibindo alta sensibilidade, baixa especificidade e melhor valor preditivo negativo.

Com respeito aos intervalos de credibilidade, nota-se que os intervalos obtidos pelo exame de fezes para a prevalência da infecção, especificidade, valor preditivo positivo e negativo, como para a taxa de erro do resultado falso positivo são mais precisos, pois possuem menor amplitude, do que o teste sorológico; mas com intervalos menos precisos para a sensibilidade e taxa de erro do resultado falso negativo, como mostra a Tabela 4.13.

**Tabela 4.13.** Intervalos de credibilidade e suas amplitudes para o teste sorológico e exame de fezes considerados separadamente.

	Teste sorológico		Exame de fezes	
	ICr	amplitude	ICr	amplitude
$\pi$	(0,266, 0,991)	0,725	(0,407, 0,989)	0,582
$\eta$	(0,734, 0,923)	0,189	(0,209, 0,473)	0,264
$\theta$	(0,240, 0,914)	0,674	(0,885, 0,986)	0,101
$\gamma$	(0,294, 0,997)	0,703	(0,777, 0,998)	0,221
$\delta$	(0,025, 0,868)	0,843	(0,014, 0,713)	0,699
$\eta^*$	(0,077, 0,266)	0,189	(0,527, 0,791)	0,264
$\theta^*$	(0,086, 0,760)	0,674	(0,014, 0,115)	0,101

ICr: Intervalo de credibilidade.

2. Combinando os dois testes, a prevalência da infecção encontrada para os refugiados foi de 75,0%; sensibilidades de 88,3% e 31,3%, especificidades de 66,2% e 95,7%, valores preditivos positivos de 87,5% e 95,5%, valores preditivos negativos de 63,2% e 31,5%, taxas de erros dos resultados falsos negativos de 11,7% e 68,7%, e taxas de erros dos resultados falsos positivos de 33,8% e 4,3%, respectivamente.

Nota-se que quando se combina os testes, o sorológico estima sensibilidade, valor preditivo negativo e taxa de erro do resultado falso positivo da infecção superior ao exame de fezes, mas com especificidade, valor preditivo positivo e taxa de erro do resultado falso negativo inferior.

Em relação aos refugiados, o teste sorológico mostrou-se mais sensível (88,3%) para diagnosticar a infecção do que o exame de fezes (31,3%), e menos específico (66,2%) para não diagnosticar a infecção do que o exame de fezes (95,7%). Destes valores obtém-se que o valor preditivo negativo é melhor no teste sorológico, e que o valor preditivo positivo é melhor no exame de fezes.

As amplitudes dos intervalos para a prevalência da infecção, segundo os dois testes foi de 40,3%; para as sensibilidades de 16,4% e 20,9%, especificidades de 55,7% e 8,3%, valores preditivos positivos de 36,3% e 12,5%, valores preditivos negativos de 63,2% e 49,7%, taxas de erros dos resultados falsos negativos de 16,4% e 20,9%, e taxas de erros dos resultados falsos positivos de 55,7% e 8,3%, respectivamente.

Com respeito aos intervalos de credibilidade, tem-se que os intervalos obtidos pelo exame de fezes para a especificidade, valor preditivo positivo e negativo, como para a taxa de erro do resultado falso positivo são mais precisos, pois possuem menor amplitude, do que o teste sorológico; mas com intervalos menos precisos para a sensibilidade e taxa de erro do resultado falso negativo, como exhibe a Tabela 4.14.

**Tabela 4.14.** Intervalos de credibilidade e suas amplitudes para o teste sorológico e exame de fezes considerados separadamente, assim como para a combinação de ambos.

	Teste sorológico		Exame de fezes		Combinação dos teste	
	ICr	amplitude	ICr	amplitude	ICr	amplitude
$\pi$	(0,266, 0,991)	0,725	(0,407, 0,989)	0,582	(0,516, 0,919)	0,403
$\eta_1$	(0,734, 0,923)	0,189			(0,789, 0,953)	0,164
$\theta_1$	(0,240, 0,914)	0,674			(0,371, 0,928)	0,557
$\gamma_1$	(0,294, 0,997)	0,703			(0,625, 0,988)	0,363
$\delta_1$	(0,025, 0,868)	0,843			(0,245, 0,874)	0,629
$\eta_1^*$	(0,077, 0,266)	0,189			(0,047, 0,211)	0,164
$\theta_1^*$	(0,086, 0,760)	0,674			(0,072, 0,629)	0,557
$\eta_2$			(0,209, 0,473)	0,264	(0,224, 0,433)	0,209
$\theta_2$			(0,885, 0,986)	0,101	(0,906, 0,989)	0,083
$\gamma_2$			(0,777, 0,998)	0,221	(0,868, 0,993)	0,125
$\delta_2$			(0,014, 0,713)	0,699	(0,103, 0,600)	0,497
$\eta_2^*$			(0,527, 0,791)	0,264	(0,567, 0,776)	0,209
$\theta_2^*$			(0,014, 0,115)	0,101	(0,011, 0,094)	0,083

ICr: Intervalo de credibilidade.

Baseando-se na Tabela 4.14 tem-se que, comparando a prevalência da infecção causada pelo parasita *strongyloides* do teste sorológico e exame de fezes separadamente, com a prevalência da infecção combinando ambos os testes nota-se uma redução de 32,2% e 17,9% nas amplitudes do teste sorológico e exame de fezes, quando comparadas com a amplitude da prevalência combinando os dois testes. Com redução nas amplitudes de 2,5%, 11,7%, 34,0%, 21,4%, 2,5% e 11,7% para a sensibilidade, especificidade, valores preditivos positivo e negativo, taxa de erro de um resultado falso negativo e falso positivo do teste sorológico, respectivamente, quando comparadas com as obtidas para a combinação de ambos os testes. Para o exame de fezes a redução nas amplitudes foi de 5,5%, 1,8%, 9,6%, 20,2%, 5,5% e 1,8% para a sensibilidade, especificidade, valores preditivos positivo e negativo,

taxa de erro de um resultado falso negativo e falso positivo, respectivamente, comparando com as obtidas para a combinação de ambos os testes.

Destes resultados pode-se concluir que inferências mais definidas para os refugiados do Camboja sobre a prevalência da infecção causada pelo parasita *strongyloides*, as sensibilidades, especificidades, valores preditivos positivos e negativos, assim como para as taxas de erros dos resultados falsos negativos e falsos positivos, foram obtidas da combinação de ambos os testes comparadas àquelas do teste sorológico e exame de fezes separadamente.

Então, nota-se que quando os métodos bayesianos propostos são utilizados, combinando os dois testes, obtém-se melhores estimativas para os parâmetros dos testes diagnóstico, pois com a combinação de ambos os testes tem-se mais informação.

3. A indicação de convergência para todos os parâmetros de interesse nesta aplicação foi verificada através da análise do comportamento das trajetórias das cadeias, e também a partir do diagnóstico de Gelman & Rubin.

4. O *Software WinBUGS*, utilizado para a implementação dos algoritmos na aplicação apresentada, mostrou-se uma importante ferramenta computacional, pois forneceu automaticamente os resumos decorrentes da amostra obtida como média, mediana, desvio padrão, intervalo de credibilidade e, também, gráficos para análise das trajetórias das cadeias geradas, densidades e diagnósticos de convergência de Gelman & Rubin de cada parâmetro.

5. Os resultados estimados para a infecção causada pelo parasita *strongyloides* apresentados nesta aplicação, para cada teste separadamente, assim como para a combinação de ambos, estão de acordo com os apresentados em Joseph *et al.* (1995).

## Capítulo 5

# Aplicação: Diagnóstico de Obesidade

Neste capítulo, será apresentada uma aplicação com dados reais sobre a obesidade em escolares do sexo feminino e masculino, com o objetivo de estimar os parâmetros de um único critério antropométrico, assim como dos dois critérios combinados na ausência de um padrão ouro. Na Seção 5.1 será apresentada uma introdução sobre a Obesidade e Critérios Antropométricos.

Na Seção 5.2, segue os resultados obtidos pelo critério de Must e Cole para a obesidade em escolares do sexo feminino e masculino. A Seção 5.3 apresenta os resultados obtidos a partir da implementação do algoritmo *Gibbs Sampling* baseado no procedimento descrito em Joseph *et al.* (1995) para cada critério antropométrico separadamente, assim como para a combinação dos critérios em ambos os sexos.

Uma discussão sobre a indicação de convergência dos parâmetros envolvidos na aplicação será apresentada na Seção 5.4. Na Seção 5.5, segue os comentários sobre os valores estimados nesta aplicação.

## 5.1 Obesidade e Critérios Antropométricos

Conforme Guillaume (1999) e Zlochevsky (1996) *apud* Sotelo (2003) a obesidade é uma doença crônica, definida como excesso de gordura corporal, em que ocorre concomitância de fatores de risco genéticos e ambientais (Blair, 1996; Egger, 1996). Os fatores genéticos aparecem como os mais determinantes da Massa Corporal; no entanto, as situações ambientais podem diminuir ou aumentar a influência desses fatores (Stunkard, 1990; Perry *et al.*, 1990).

A avaliação da prevalência de sobrepeso e obesidade em inquéritos nacionais de saúde baseia-se totalmente em critérios antropométricos, mais especificamente em medidas de Peso e Estatura.

O critério de Must *et al.* (Must, 1991) utilizado para diagnosticar estado nutricional é a distribuição percentilar para Índice de Massa Corporal (IMC), que foi elaborada para classificação de crianças a partir de 6 anos e adultos segundo sexo, faixa etária e grupo étnico, baseado nos dados da *National Health and Nutrition Examination I* (NHANES I). A partir do cálculo do IMC (peso em quilos dividido pela estatura em metros ao quadrado), utilizam-se os pontos de corte proposto por Must que definem como sobrepeso crianças com IMC entre os percentis 85 e 95 e obesas crianças com IMC acima do percentil 95.

As curvas de Cole *et al.* (2000) para IMC foram desenvolvidas baseadas em dados de 6 países (Brasil, Estados Unidos, Grã Bretanha, Hong Kong, Holanda e Cingapura), cada amostra com mais de 10.000 indivíduos. Elas foram estimadas de forma que os pontos das curvas ajustadas dos percentis 85 e 95 de IMC aos 18 anos fossem obrigatoriamente os pontos de corte para sobrepeso e obesidade utilizados para adultos (25 e 30 Kg/m<sup>2</sup>, respectivamente). As curvas contemplam a distribuição percentilar por faixa etária dentro de cada sexo.

Na ausência de um padrão ouro como indicador de obesidade, em crianças e adolescentes, os métodos estatísticos bayesianos apresentados nas Subseções 4.1.1 e 4.1.2 foram

propostos para estimar os parâmetros de cada um destes critérios antropométricos (Must e Cole), assim como para suas combinações. As densidades marginais *a posteriori* de todos os parâmetros foram estimadas usando o *Gibbs Sampling*.

As técnicas foram aplicadas para estimar a prevalência de obesidade e para investigar as propriedades dos critérios antropométricos, Must e Cole, respectivamente, usando dados de um estudo realizado em uma escola particular da cidade de São Paulo, cursada por 2787 adolescentes de famílias bem afortunadas. Os adolescentes com doença aguda ou crônica foram excluídos. A amostra estudada incluiu 418 adolescentes aleatoriamente selecionados. O consentimento informado foi assinado pelos pais dos adolescentes. O protocolo de estudo foi submetido e aprovado pelo Comitê Ético do Hospital de São Paulo.

A distribuição referente ao sexo foi de 47,61% ( $n = 199$ ) feminino e 52,39% ( $n = 219$ ) masculino, de um total de 418 escolares.

## 5.2 Obesidade em Escolares

Esta aplicação, teve por objetivo estimar os parâmetros (sensibilidade, especificidade, valores preditivos positivo e negativo, assim como a prevalência de obesidade) de cada critério antropométrico (Must e Cole) e da suas combinações em uma situação onde não há padrão ouro.

Os dados consistem de 418 adolescentes aleatoriamente selecionados, onde 199 são do sexo feminino e 219 do sexo masculino. Os resultados obtidos pelos critérios de Must e Cole para a obesidade no sexo feminino e masculino, apresentam-se nas Tabelas 5.1 e 5.2, respectivamente.

**Tabela 5.1.** Resultados dos critérios antropométricos Must e Cole para a obesidade em 199 escolares do sexo feminino.

		Cole		
		Positivo	Negativo	Total
Must	Positivo	9	2	11
	Negativo	0	188	188
Total		9	190	199

**Tabela 5.2.** Resultados dos critérios antropométricos Must e Cole para a obesidade em 219 escolares do sexo masculino.

		Cole		
		Positivo	Negativo	Total
Must	Positivo	20	8	28
	Negativo	0	191	191
Total		20	199	219

O objetivo nesta seção é estimar a prevalência ( $\pi$ ) de obesidade em escolares, a sensibilidade ( $\eta$ ), especificidade ( $\theta$ ), e os valores preditivos positivo ( $\gamma$ ) e negativo ( $\delta$ ), assim como as taxas de erros, taxa de um resultado falso negativo ( $\eta^*$ ) e falso positivo ( $\theta^*$ ), para os resultados de cada critério separadamente (Must e Cole) e para a combinação dos resultados de ambos os critérios antropométricos.

Para a estimação destes parâmetros os algoritmos *Gibbs Sampling* apresentados nas Subseções 4.1.1 e 4.1.2 da Seção 4.1 foram implementados utilizando o *Software WinBUGS*. As implementações computacionais destes algoritmos em forma de programa, encontram-se no Apêndice G.

Para o presente estudo, foram consideradas distribuições *a priori* não informativas para as sensibilidades e especificidades de obesidade dos dois critérios antropométricos tanto para o sexo feminino quanto para o masculino, com o intuito de expressar a falta de informação. Para a prevalência de obesidade, estipulou-se uma distribuição *a priori* informativa, a partir da informação do especialista da Disciplina de Nutrição e Metabolismo,

UNIFESP/EPM, de que a prevalência de obesidade na população estudada está entre 10% e 20%, esta informação é válida tanto para o sexo feminino quanto para o sexo masculino, como também para os dois critérios antropométricos (Must e Cole). Através do procedimento para obtenção dos hiperparâmetros da distribuição *a priori*, apresentado na Subseção 4.1.3 do Capítulo 4, tem-se que os hiperparâmetros da *priori* beta para a prevalência de obesidade são (87,375, 611,625) e os hiperparâmetros da *priori* beta para as sensibilidades e especificidades de cada critério antropométrico são (1,1) e (1,1), respectivamente.

Os resultados foram obtidos a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas. Os resultados para médias, desvios padrões, medianas e intervalos de credibilidade são apresentados na Seção 5.3.

## 5.3 Resultados

Esta seção apresenta os resultados obtidos pelos critérios antropométricos Must e Cole separadamente, assim como para a combinação destes dois critérios. Estes resultados foram obtidos a partir da implementação dos algoritmos *Gibbs Sampling* apresentados nas Subseções 4.1.1 e 4.1.2, respectivamente.

### 5.3.1 Critérios Antropométricos Must e Cole para o Sexo Feminino

Esta aplicação com o critério de Must, envolve  $a = 11$  escolares com resultados positivos e  $b = 188$  escolares com resultados negativos, e com o critério de Cole, envolve  $a = 9$  escolares com resultados positivos e  $b = 190$  escolares com resultados negativos. Os resultados para o critério de Must e de Cole, obtidos a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas, são

apresentados nas Tabela 5.3 e Tabela 5.4. As Tabelas 5.5 e 5.6 sumarizam os resultados obtidos para os dados latentes  $Y_1$  e  $Y_2$  do critério de Must e Cole, respectivamente.

As Figuras 5.1 e 5.2 mostram as distribuições estimadas via *Gibbs Sampling* para os dados latentes usando a informação do critério de Must e Cole separadamente. As Figuras 5.3, 5.4, 5.5 e 5.6 apresentam, respectivamente, as densidades estimadas via *Gibbs Sampling* e o comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta$  e  $\theta$  usando dados do critério de Must e Cole separadamente.

**Tabela 5.3.** Valores estimados via *Gibbs Sampling* para os parâmetros do critério de Must, com  $a = 11$  e  $b = 188$  para o sexo feminino.

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,124	0,012	0,100	0,124	0,149
$\eta$	0,267	0,173	0,012	0,248	0,646
$\theta$	0,964	0,024	0,911	0,967	0,998
$\gamma$	0,508	0,287	0,026	0,512	0,976
$\delta$	0,903	0,025	0,858	0,902	0,955
$\eta^*$	0,733	0,173	0,354	0,752	0,988
$\theta^*$	0,036	0,024	0,002	0,033	0,089

**Tabela 5.4.** Valores estimados via *Gibbs Sampling* para os parâmetros do critério de Cole, com  $a = 9$  e  $b = 190$  para o sexo feminino.

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,124	0,012	0,100	0,123	0,148
$\eta$	0,225	0,152	0,010	0,205	0,562
$\theta$	0,969	0,021	0,922	0,972	0,998
$\gamma$	0,504	0,289	0,025	0,508	0,974
$\delta$	0,898	0,023	0,858	0,897	0,946
$\eta^*$	0,775	0,152	0,438	0,795	0,99
$\theta^*$	0,031	0,021	0,001	0,028	0,077

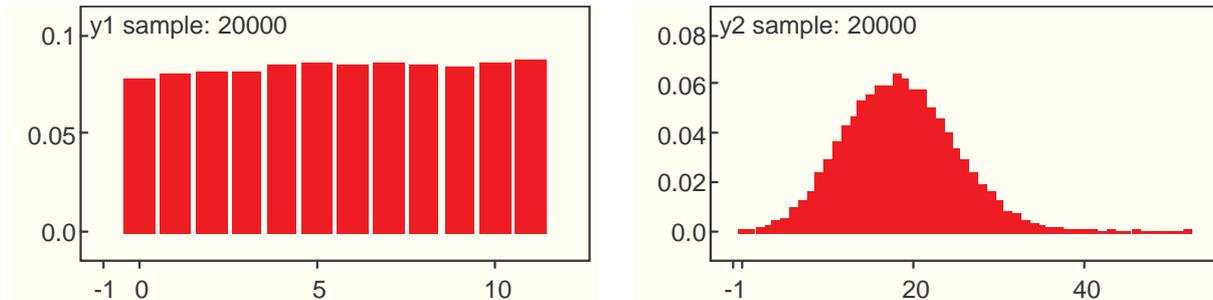
As Tabelas 5.3 e 5.4 apresentam as médias, desvios padrões, medianas e intervalos de credibilidade para a prevalência de obesidade, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo para o sexo feminino, segundo os dois critérios antropométricos. Para o critério de Must a prevalência, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade foram em média 12,4%, 26,7%, 96,4%, 50,8%, 90,3%, 73,3% e 3,6%, respectivamente. Os intervalos de credibilidade foram (10,0%, 14,9%) para prevalência, (1,2%, 64,6%) para sensibilidade, (91,1%, 99,8%) para especificidade, (2,6%, 97,6%) para o valor preditivo positivo, (85,8%, 95,5%) para o valor preditivo negativo, (35,4%, 98,8%) para a taxa de erro do resultado falso negativo e (0,2%, 8,9%) para a taxa de erro do resultado falso positivo. Já o critério de Cole apresentou prevalência, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade de 12,4%, 22,5%, 96,9%, 50,4%, 89,8%, 77,5% e 3,1%, respectivamente. Os intervalos de credibilidade foram (10,0%, 14,8%) para prevalência, (1,0%, 56,2%) para sensibilidade, (92,2%, 99,8%) para especificidade, (2,5%, 97,4%) para o valor preditivo positivo, (85,8%, 94,6%) para o valor preditivo negativo, (43,8%, 99,0%) para a taxa de erro do resultado falso negativo e (0,1%, 7,7%) para a taxa de erro do resultado falso positivo. Para as meninas classificadas como obesas, considerando os dois critérios separadamente, encontram-se diferenças entre estes critérios para as sensibilidades e especificidades. O critério de Must para obesidade estima sensibilidade superior ao critério de Cole, mas especificidade inferior.

Com respeito aos intervalos de credibilidade, as amplitudes dos intervalos das prevalências foram de 4,9% para o critério de Must e 4,8% para o critério de Cole, das sensibilidades foram de 63,4% para o critério de Must e 55,2% para o de Cole. E as amplitudes das especificidades foram 8,7% para Must e 7,6% para Cole, dos valores preditivos positivos foram de 95% para Must e 94,9% para Cole, dos valores preditivos negativos foram de 9,7% para Must e 8,8% para Cole, das taxas de erros do resultado falso negativo foram de

63,4% para Must e 55,2% para Cole e das taxas de erros do resultado falso positivo foram de 8,7% para Must e 7,6% para Cole. Isto significa que os intervalos obtidos para todos os parâmetros estimados pelo critério de Cole são mais precisos quando comparados aos de Must.

**Tabela 5.5.** Valores estimados via *Gibbs Sampling* para os dados latentes do critério de

Must, com $a = 11$ e $b = 188$ para o sexo feminino.					
parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	5,601	3,432	0,0	6,0	11,0
$Y_2$	18,3	6,274	7,0	18,0	31,0

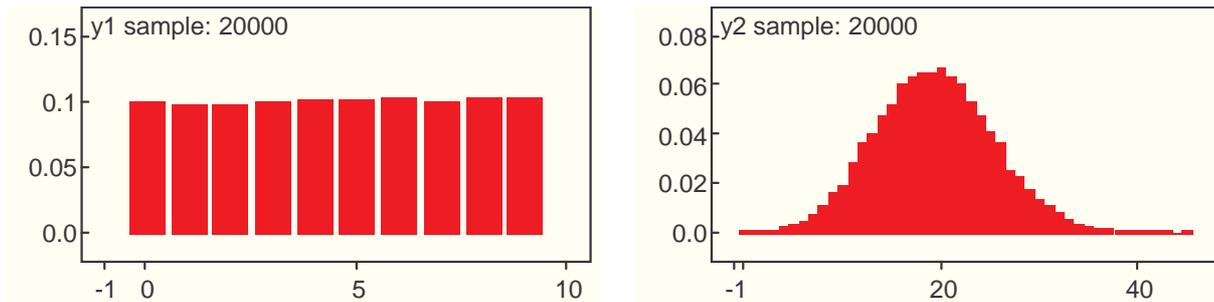


**Figura 5.1.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1$  e  $Y_2$  do critério de Must para o sexo feminino.

A partir da Tabela 5.5 e da Figura 5.1 tem-se que aproximadamente 6 escolares do sexo feminino dos 11 com resultados positivos para o critério de Must tenham resultados verdadeiros positivos, e que aproximadamente 18 escolares do sexo feminino dos 188 com resultados negativos para o critério de Must tenham resultados verdadeiros positivos.

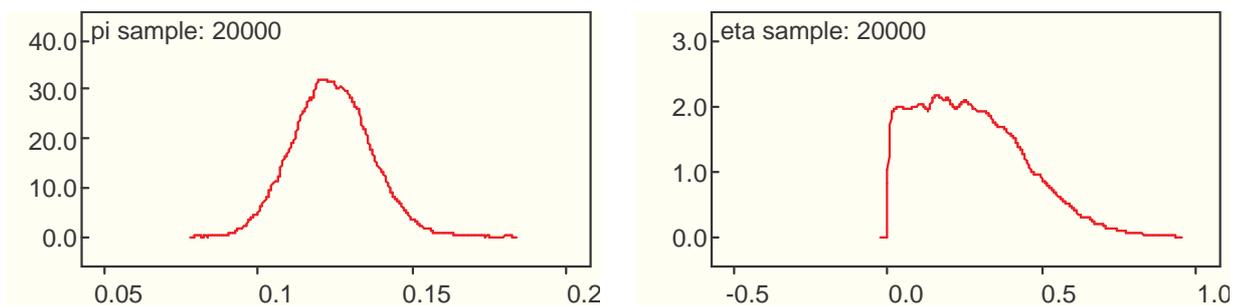
**Tabela 5.6.** Valores estimados via *Gibbs Sampling* para os dados latentes do critério de Cole, com  $a = 9$  e  $b = 190$  para o sexo feminino.

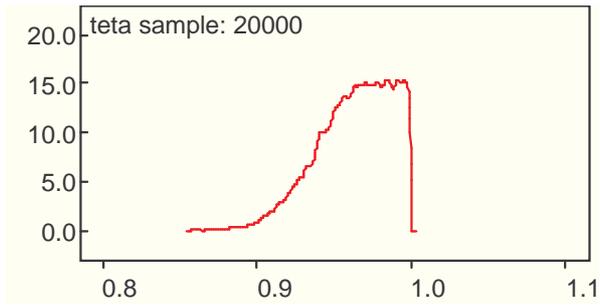
parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	4,542	2,871	0,0	5,0	9,0
$Y_2$	19,28	5,94	8,0	19,0	31,0



**Figura 5.2.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1$  e  $Y_2$  do critério de Cole para o sexo feminino.

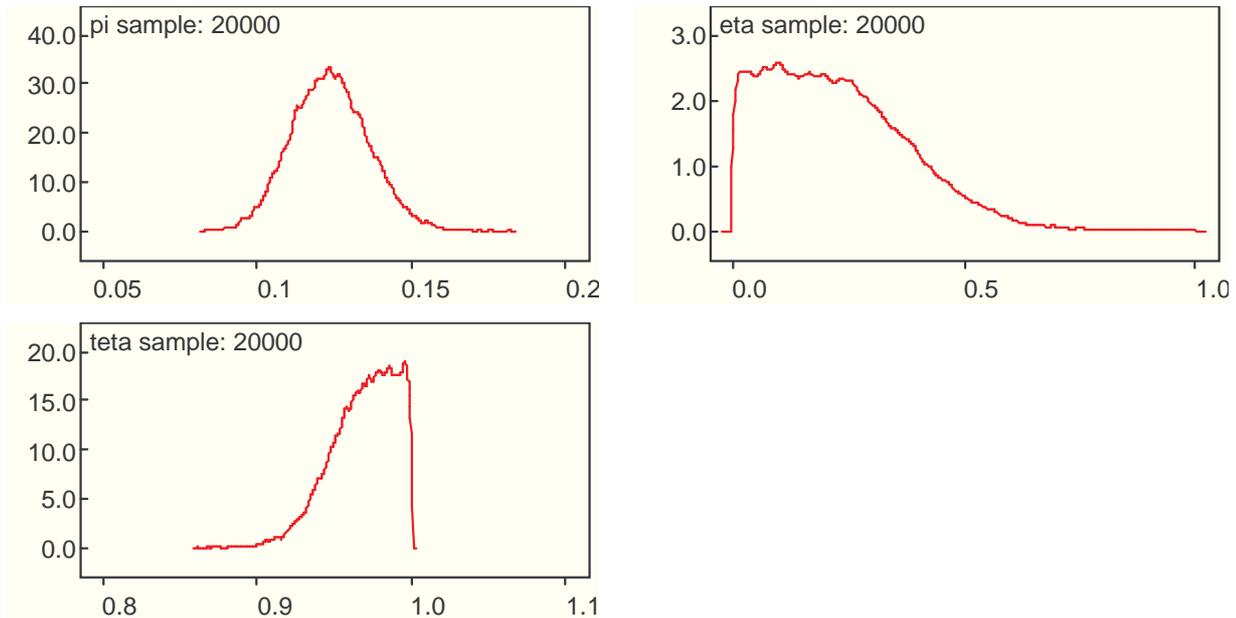
A partir da Tabela 5.6 e da Figura 5.2 tem-se que aproximadamente 5 escolares do sexo feminino dos 9 com resultados positivos para o critério de Cole tenham resultados verdadeiros positivos, e que aproximadamente 19 escolares do sexo feminino dos 190 com resultados negativos para o critério de Cole tenham resultados verdadeiros positivos.





**Figura 5.3.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério de Must para o sexo feminino.

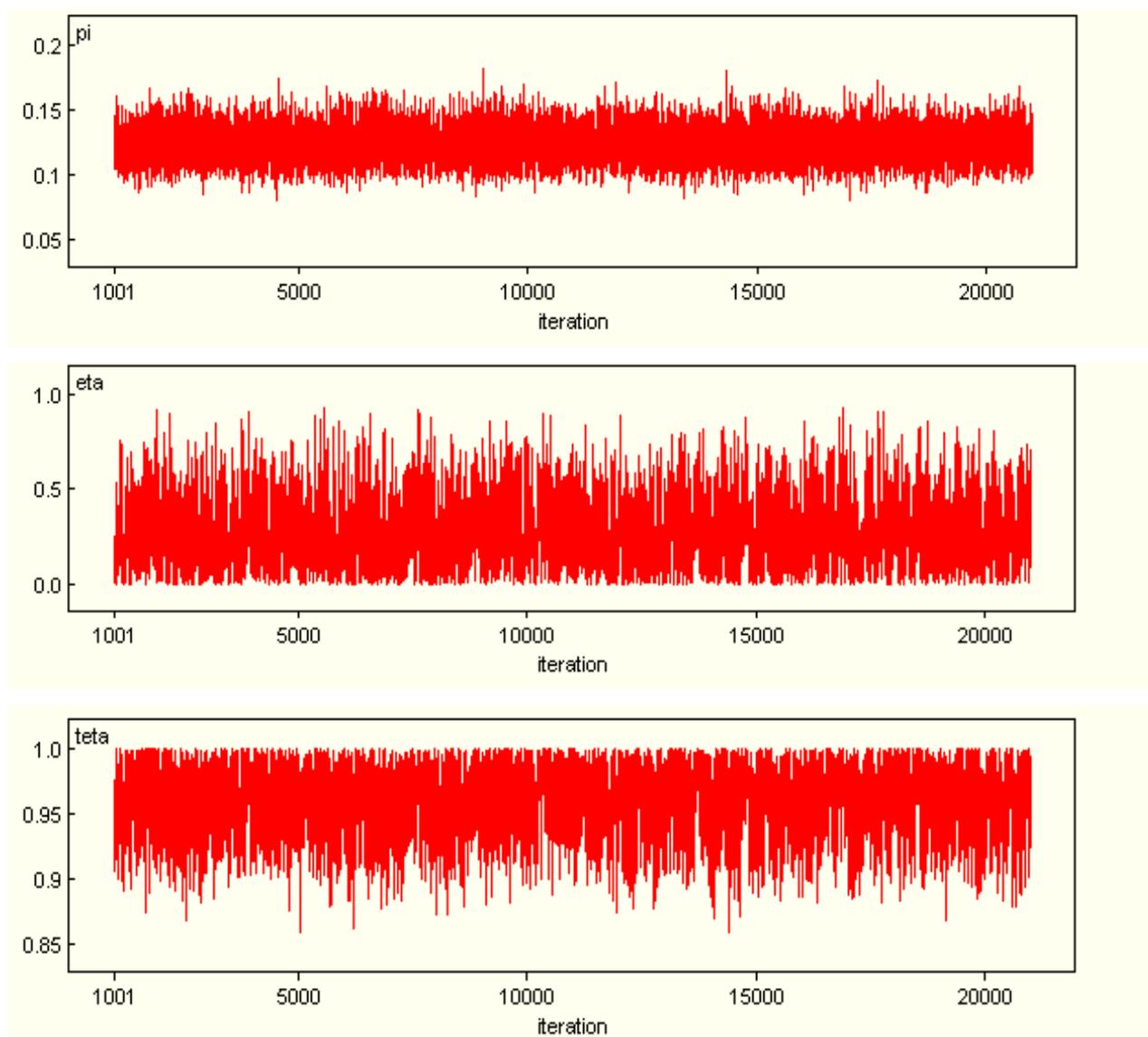
Como é evidente da Tabela 5.3 e da Figura 5.3, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da sensibilidade usando dados do critério de Must foi 26,7%, embora o intervalo de credibilidade de 95% foi de (1,20%, 64,6%), esta mesma situação também ocorre com os outros parâmetros.



**Figura 5.4.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério de Cole para o sexo feminino.

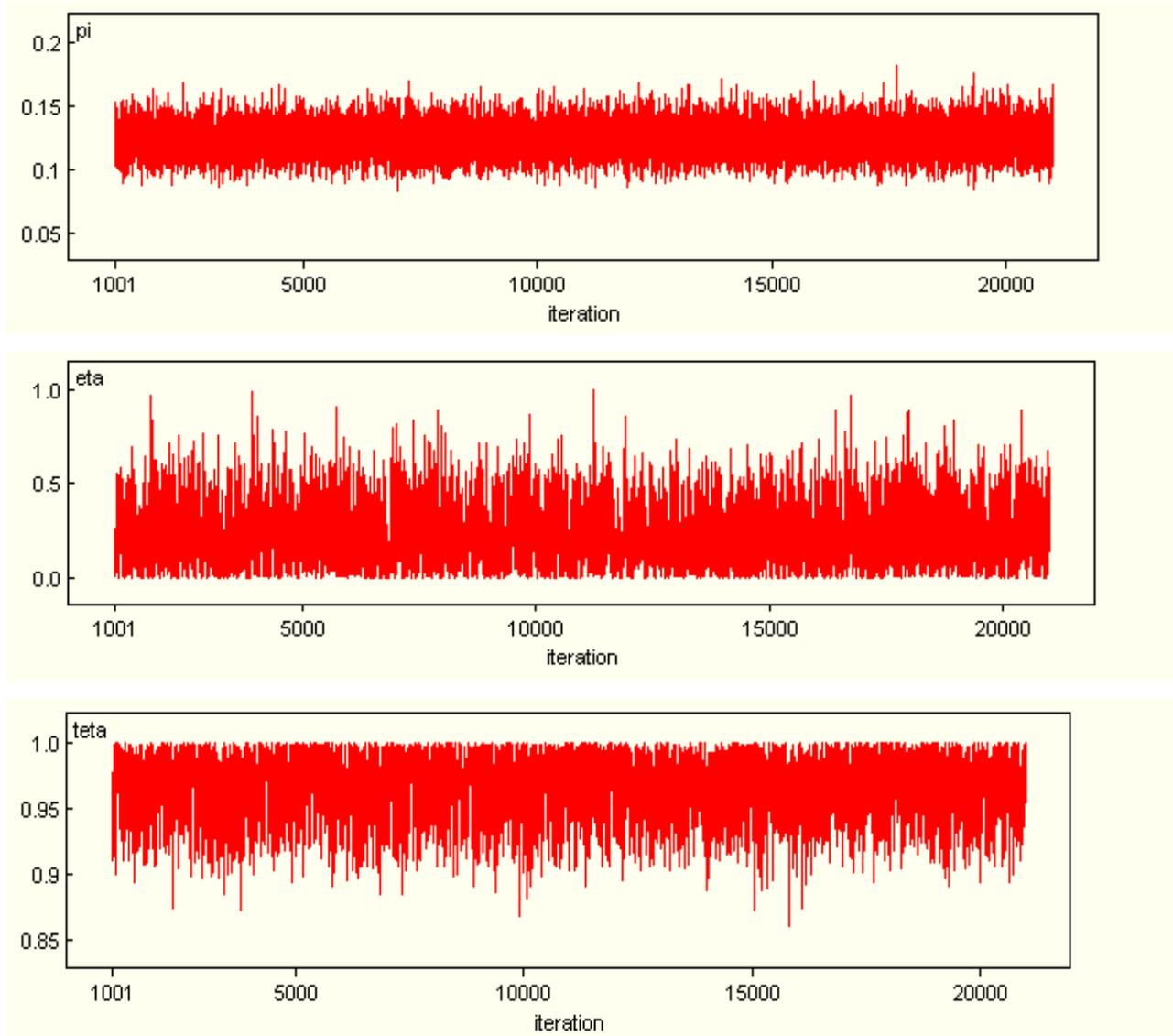
Como é evidente da Tabela 5.4 e da Figura 5.4, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da sensibilidade usando dados

do critério de Cole foi 22,5%, embora o intervalo de credibilidade de 95% foi de (1,0%, 56,2%), esta mesma situação também ocorre com os outros parâmetros.



**Figura 5.5.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério de Must para o sexo feminino.

Através da Figura 5.5, tem-se uma indicação de convergência para todos os parâmetros do critério de Must.



**Figura 5.6.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério de Cole para o sexo feminino.

Através da Figura 5.6, tem-se uma indicação de convergência para todos os parâmetros do critério de Cole.

### 5.3.2 Critérios Antropométricos Must e Cole para o Sexo Masculino

Esta aplicação com o critério de Must, envolve  $a = 28$  escolares com resultados positivos e  $b = 191$  escolares com resultados negativos, e com o critério de Cole, envolve  $a = 20$  escolares com resultados positivos e  $b = 199$  escolares com resultados negativos. Os resultados para o critério de Must e de Cole, obtidos a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas, são apresentados nas Tabela 5.7 e Tabela 5.8. As Tabelas 5.9 e 5.10 sumarizam os resultados obtidos para os dados latentes  $Y_1$  e  $Y_2$  do critério de Must e Cole, respectivamente.

As Figuras 5.7 e 5.8 mostram as distribuições estimadas via *Gibbs Sampling* para os dados latentes usando a informação do critério de Must e Cole separadamente. As Figuras 5.9, 5.10, 5.11 e 5.12 apresentam, respectivamente, as densidades estimadas via *Gibbs Sampling* e o comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta$  e  $\theta$  usando dados do critério de Must e Cole separadamente.

**Tabela 5.7.** Valores estimados via *Gibbs Sampling* para os parâmetros do critério de Must, com  $a = 28$  e  $b = 191$  para o sexo masculino.

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,125	0,012	0,100	0,124	0,149
$\eta$	0,478	0,281	0,023	0,471	0,964
$\theta$	0,916	0,045	0,829	0,917	0,993
$\gamma$	0,455	0,272	0,022	0,448	0,948
$\delta$	0,925	0,041	0,851	0,924	0,995
$\eta^*$	0,522	0,281	0,036	0,529	0,977
$\theta^*$	0,084	0,045	0,007	0,083	0,171

**Tabela 5.8.** Valores estimados via *Gibbs Sampling* para os parâmetros do critério de Cole, com  $a = 20$  e  $b = 199$  para o sexo masculino.

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,124	0,012	0,101	0,124	0,149
$\eta$	0,389	0,244	0,018	0,369	0,886
$\theta$	0,942	0,035	0,871	0,944	0,997
$\gamma$	0,489	0,287	0,024	0,483	0,969
$\delta$	0,916	0,035	0,854	0,914	0,985
$\eta^*$	0,611	0,244	0,115	0,630	0,983
$\theta^*$	0,058	0,035	0,003	0,056	0,129

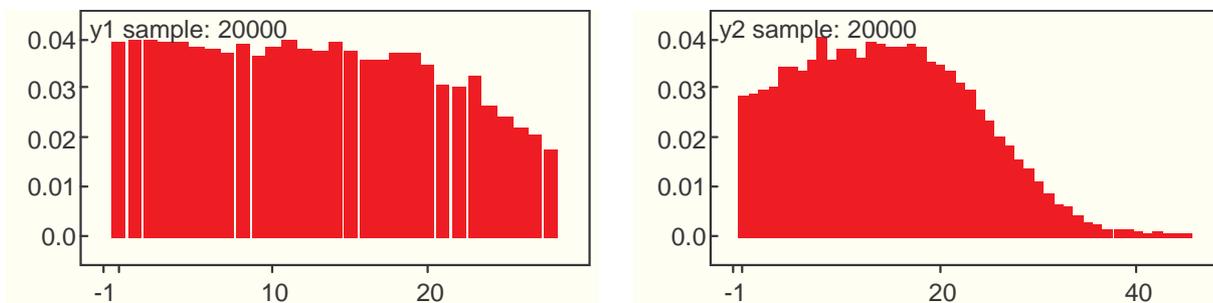
As Tabelas 5.7 e 5.8 apresentam as médias, desvios padrões, medianas e intervalos de credibilidade para a prevalência de obesidade, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo para o sexo masculino, segundo os dois critérios antropométricos. Para o critério de Must a prevalência, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade foram em média 12,5%, 47,8%, 91,6%, 45,5%, 92,5%, 52,2% e 4,5%, respectivamente. Os intervalos de credibilidade foram (10,0%, 14,9%) para prevalência, (2,3%, 96,4%) para sensibilidade, (82,9%, 99,3%) para especificidade, (2,2%, 94,8%) para o valor preditivo positivo, (85,1%, 99,5%) para o valor preditivo negativo, (3,6%, 97,7%) para a taxa de erro do resultado falso negativo e (0,7%, 17,1%) para a taxa de erro do resultado falso positivo. Já o critério de Cole apresentou prevalência, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade de 12,4%, 38,9%, 94,2%, 48,9%, 91,6%, 61,1% e 5,8%, respectivamente. Os intervalos de credibilidade foram (10,1%, 14,9%) para prevalência, (1,8%, 88,6%) para sensibilidade, (87,1%, 99,7%) para especificidade, (2,4%, 96,9%) para o valor preditivo positivo, (85,4%, 98,5%) para o valor preditivo negativo, (11,5%, 98,3%) para a taxa de erro do resultado falso negativo e (0,3%, 12,9%) para a taxa de erro do resultado

falso positivo. Para os meninos classificados como obesos, considerando os dois critérios separadamente, encontram-se diferenças entre estes critérios para as sensibilidades e especificidades. O critério de Must para obesidade estima sensibilidade superior ao critério de Cole, mas especificidade inferior.

Com respeito aos intervalos de credibilidade, as amplitudes dos intervalos das prevalências foram de 4,9% para o critério de Must e 4,8% para o critério de Cole, das sensibilidades foram de 94,1% para o critério de Must e 86,8% para o de Cole. E as amplitudes das especificidades foram 16,4% para Must e 12,6% para Cole, dos valores preditivos positivos foram de 92,6% para Must e 94,5% para Cole, dos valores preditivos negativos foram de 14,1% para Must e 13,1% para Cole, das taxas de erros do resultado falso negativo foram de 94,1% para Must e 86,8% para Cole e das taxas de erros do resultado falso positivo foram de 16,4% para Must e 12,6% para Cole. Isto significa que os intervalos obtidos para a prevalência, sensibilidade, especificidade, valor preditivo negativo, taxa de erro do resultado falso negativo e falso positivo, pelo critério de Cole são mais precisos, mas com intervalo para o valor preditivo positivo menos preciso.

**Tabela 5.9.** Valores estimados via *Gibbs Sampling* para os dados latentes do critério de Must, com  $a = 28$  e  $b = 191$  para o sexo masculino.

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	12,72	7,934	0,0	12,0	27,0
$Y_2$	14,37	8,662	0,0	14,0	32,0

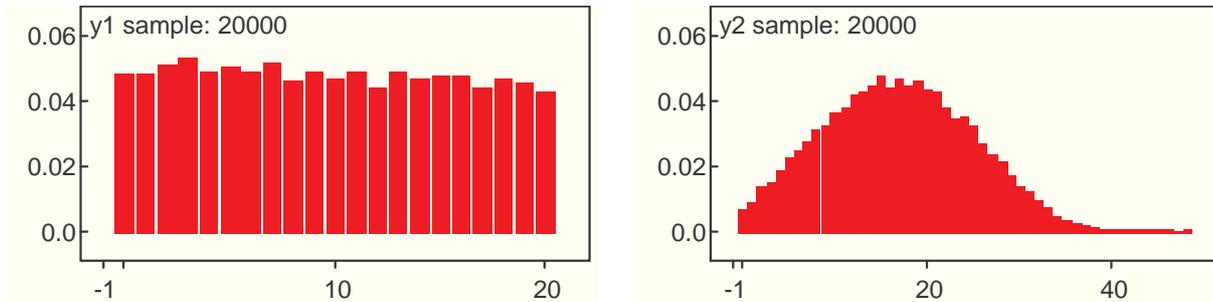


**Figura 5.7.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1$  e  $Y_2$  do critério de Must para o sexo masculino.

A partir da Tabela 5.9 e da Figura 5.7 tem-se que aproximadamente 13 escolares do sexo masculino dos 28 com resultados positivos para o critério de Must tenham resultados verdadeiros positivos, e que aproximadamente 14 escolares do sexo masculino dos 191 com resultados negativos para o critério de Must tenham resultados verdadeiros positivos.

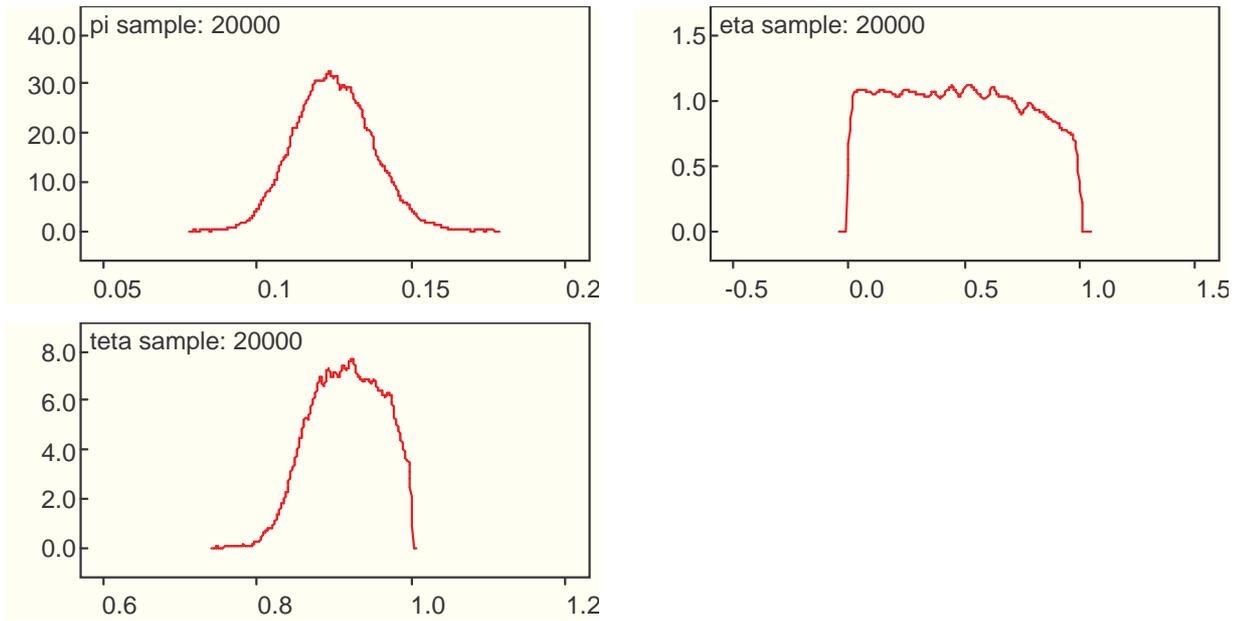
**Tabela 5.10.** Valores estimados via *Gibbs Sampling* para os dados latentes do critério de Cole, com  $a = 20$  e  $b = 199$  para o sexo masculino.

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	9,774	6,022	0,0	10,0	20,0
$Y_2$	16,77	7,955	2,0	17,0	32,0



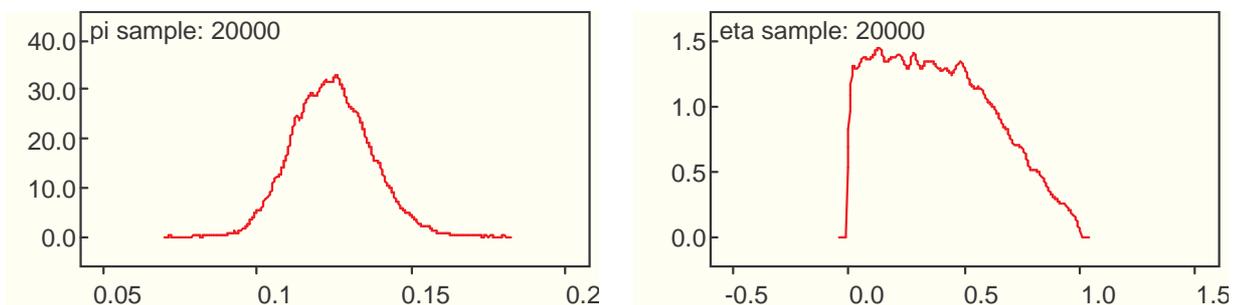
**Figura 5.8.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1$  e  $Y_2$  do critério de Cole para o sexo masculino.

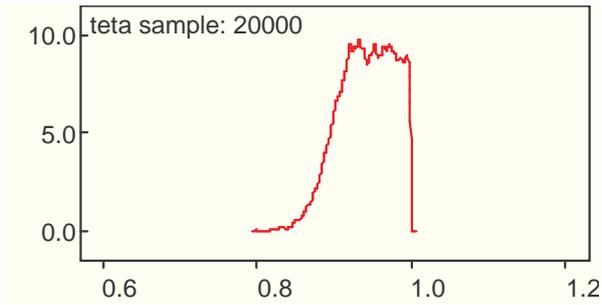
A partir da Tabela 5.10 e da Figura 5.8 tem-se que aproximadamente 10 escolares do sexo masculino dos 20 com resultados positivos para o critério de Cole tenham resultados verdadeiros positivos, e que aproximadamente 17 escolares do sexo masculino dos 199 com resultados negativos para o critério de Cole tenham resultados verdadeiros positivos.



**Figura 5.9.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério de Must para o sexo masculino.

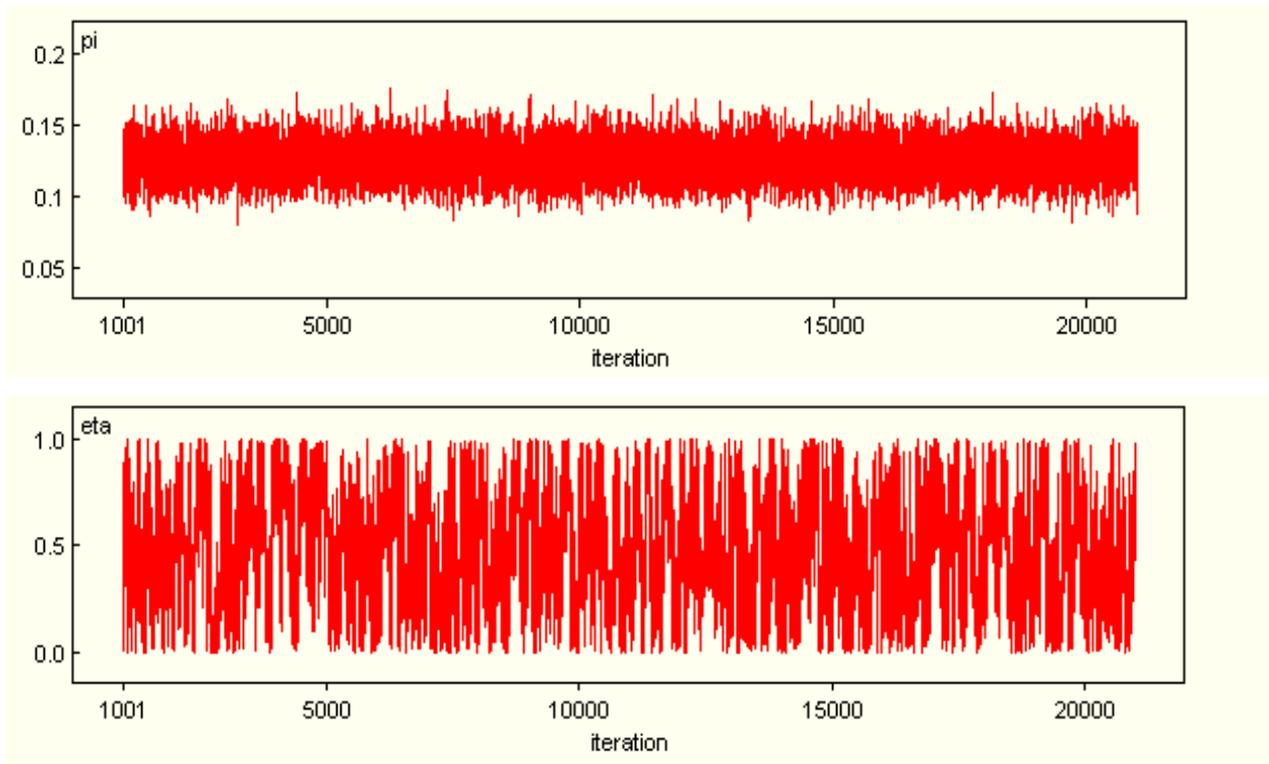
Como é evidente da Tabela 5.7 e da Figura 5.9, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da sensibilidade usando dados do critério de Must foi 47,8%, embora o intervalo de credibilidade de 95% foi de (2,3%, 96,4%), esta mesma situação também ocorre com os outros parâmetros.

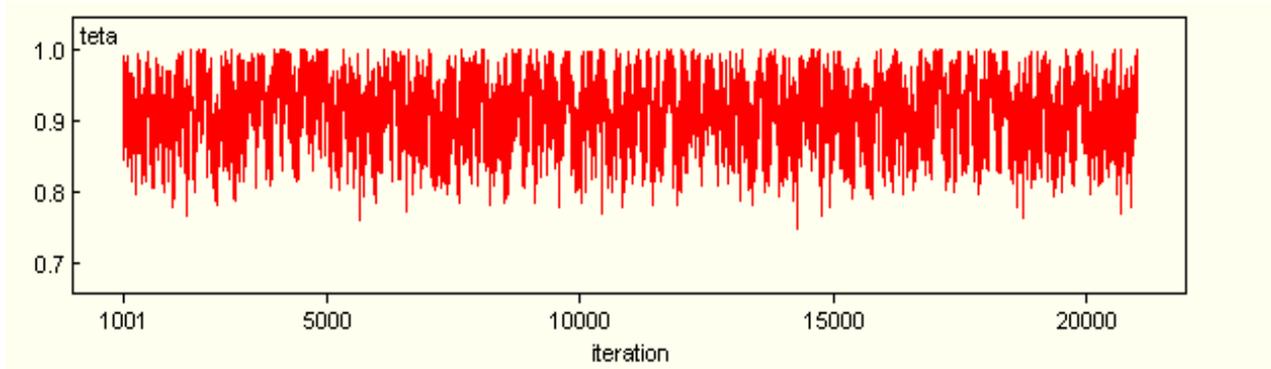




**Figura 5.10.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério de Cole para o sexo masculino.

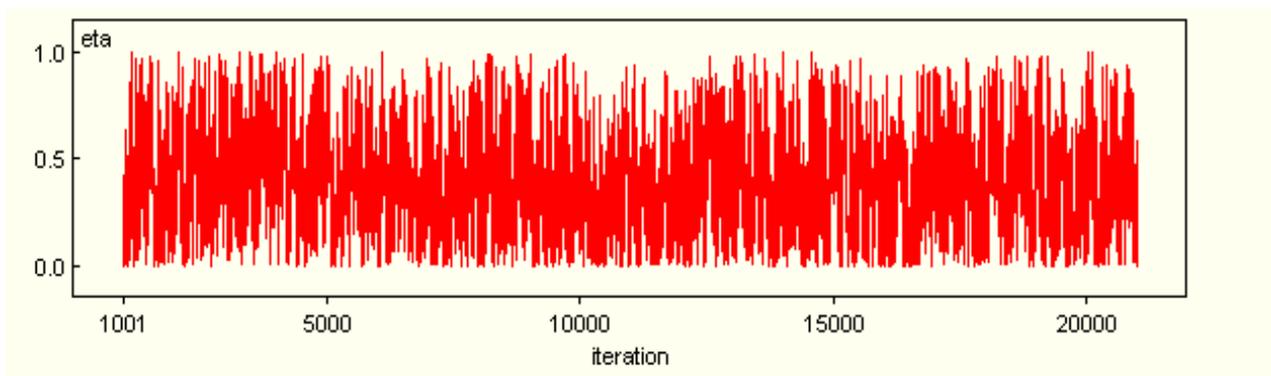
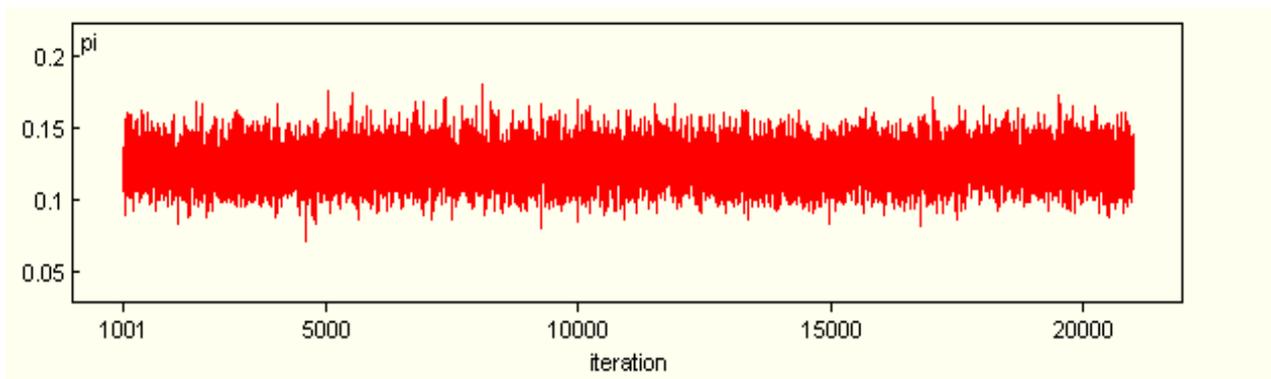
Como é evidente da Tabela 5.8 e da Figura 5.10, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da sensibilidade usando dados do critério de Cole foi 24,4%, embora o intervalo de credibilidade de 95% foi de (1,8%, 88,6%), esta mesma situação também ocorre com os outros parâmetros.

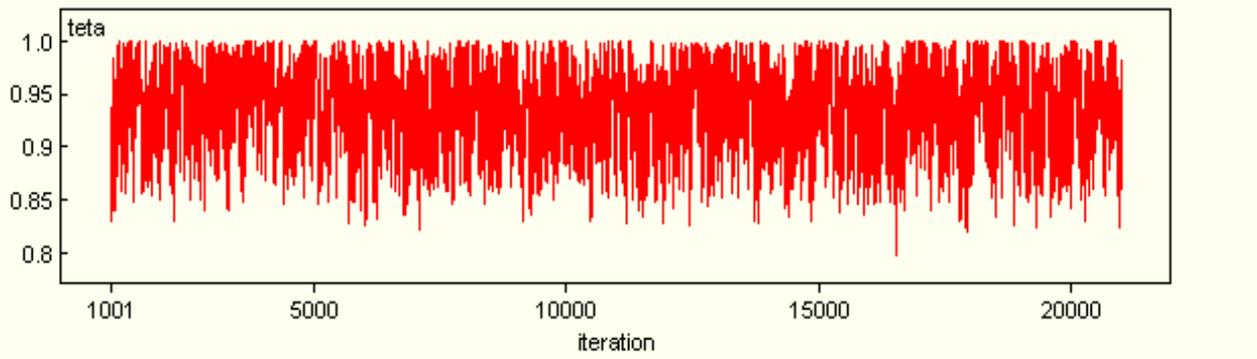




**Figura 5.11.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério de Must para o sexo masculino.

Através da Figura 5.11, tem-se uma indicação de convergência para todos os parâmetros do critério de Must.





**Figura 5.12.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta$  e  $\theta$  do critério de Cole para o sexo masculino.

Através da Figura 5.12, tem-se uma indicação de convergência para todos os parâmetros do critério de Cole.

### 5.3.3 Combinação dos Critério Must e Cole para ambos os Sexos

Esta aplicação com a combinação do critério antropométrico de Must e Cole para o sexo feminino, envolve  $u = 9$  escolares com resultados positivos em ambos os critérios,  $v = 2$  escolares com resultados positivos no de Must mas negativos no de Cole,  $w = 0$  escolar com resultado negativo no de Must mas positivo no de Cole, e  $x = 188$  escolares com resultados negativos em ambos os critérios, e com a combinação do critério antropométrico de Must e Cole para o sexo masculino, envolve  $u = 20$  escolares com resultados positivos em ambos os critérios,  $v = 8$  escolares com resultados positivos no de Must mas negativos no de Cole,  $w = 0$  escolar com resultado negativo no de Must mas positivo no de Cole, e  $x = 191$  escolares com resultados negativos em ambos os critérios.

Os resultados obtidos para o sexo feminino a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas, são apresentados na Tabela 5.11. A Tabela 5.12 sumariza os resultados obtidos para os dados latentes  $Y_1, Y_2, Y_3$  e  $Y_4$  da combinação destes dois critérios para o sexo feminino. E os resultados

obtidos para o sexo masculino a partir das cadeias geradas, com um período de aquecimento de 1000 iterações e 20000 posteriores armazenadas, são apresentados na Tabela 5.13. A Tabela 5.14 sumariza os resultados obtidos para os dados latentes  $Y_1, Y_2, Y_3$  e  $Y_4$  da combinação destes dois critérios para o sexo masculino.

A Figura 5.12 mostra as distribuições estimadas via *Gibbs Sampling* para os dados latentes usando a informação da combinação dos critérios para o sexo feminino. As Figuras 5.13 e 5.14 apresentam, respectivamente, as densidades estimadas via *Gibbs Sampling* e o comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  usando dados da combinação dos dois critérios antropométricos para o sexo feminino. E a Figura 5.15 mostra as distribuições estimadas via *Gibbs Sampling* para os dados latentes usando a informação da combinação dos critérios para o sexo masculino. As Figuras 5.16 e 5.17 apresentam, respectivamente, as densidades estimadas via *Gibbs Sampling* e o comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  usando dados da combinação dos dois critérios antropométricos para o sexo masculino.

É importante salientar que quando os critérios são combinados, os parâmetros (sensibilidade, especificidade, valor preditivo positivo e negativo), por exemplo, do critério Must são estimados considerando os dados do critério Cole, e vice-versa. Para a presente aplicação da combinação de ambos os critérios os parâmetros  $\eta_1, \theta_1, \gamma_1$  e  $\delta_1$  são do critério Must e  $\eta_2, \theta_2, \gamma_2$  e  $\delta_2$  do Cole.

É relevante notar que o dado observado do total de resultados negativos no critério de Must mas positivos no de Cole ( $w$ ), não possui nenhuma informação, isto é,  $w = 0$  tanto para o sexo feminino quanto para o masculino. Dessa forma, o dado latente  $Y_3$  não será fornecido.

• **Sexo Feminino**

**Tabela 5.11.** Valores estimados via *Gibbs Sampling* para os parâmetros da combinação do critério de Must e Cole, com  $u = 9$ ,  $v = 2$ ,  $w = 0$  e  $x = 188$  para o sexo feminino.

parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,110	0,010	0,089	0,109	0,131
$\eta_1$	0,869	0,115	0,574	0,902	0,996
$\theta_1$	0,992	0,008	0,972	0,994	0,999
$\gamma_1$	0,931	0,059	0,781	0,947	0,998
$\delta_1$	0,984	0,014	0,947	0,988	0,999
$\eta_2$	0,764	0,138	0,455	0,781	0,978
$\theta_2$	0,995	0,005	0,981	0,996	0,999
$\gamma_2$	0,948	0,049	0,815	0,962	0,998
$\delta_2$	0,972	0,017	0,934	0,974	0,997
$\eta_1^*$	0,131	0,115	0,004	0,098	0,426
$\theta_1^*$	0,008	0,008	0,000	0,006	0,027
$\eta_2^*$	0,236	0,138	0,022	0,219	0,545
$\theta_2^*$	0,005	0,005	0,000	0,004	0,018

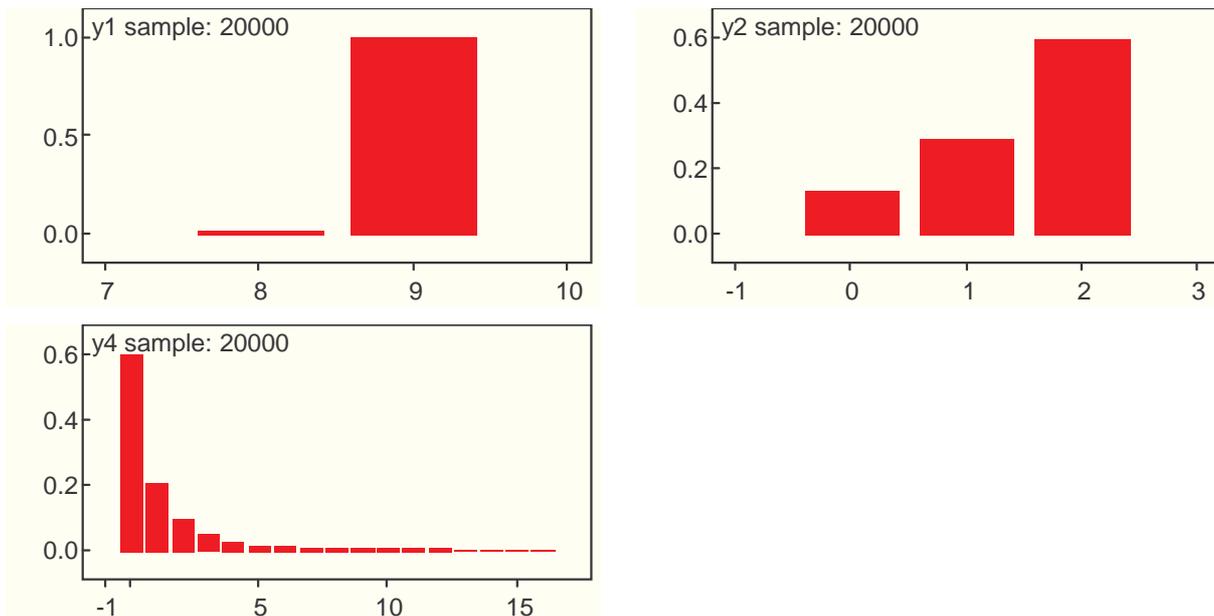
A Tabelas 5.11 apresenta as médias, desvios padrões, medianas e intervalos de credibilidade para a prevalência de obesidade, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo para o sexo feminino, segundo os dois critérios antropométricos. Para os dois critérios combinados a prevalência de obesidade foi em média 11,0%, e o intervalo de credibilidade (8,9%, 13,1%) com amplitude de 4,2%. Para o critério de Must as estimativas das sensibilidades, especificidades, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade foram em média 86,9%, 99,2%, 93,1%, 98,4%, 13,1% e 0,8%, com intervalos de credibilidade (57,4%, 99,6%), (97,2%, 99,9%), (78,1%, 99,8%), (94,7%, 99,9%), (0,4%, 42,6%) e (0,0%,

2,7%) de amplitudes 42,2%, 2,7%, 21,7%, 5,2%, 42,2% e 2,7%, respectivamente. Já o critério de Cole apresentou sensibilidades, especificidades, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade de 76,4%, 99,5%, 94,8%, 97,2%, 23,6% e 0,5%, com intervalos de credibilidade (45,5%, 97,8%), (98,1%, 99,9%), (81,5%, 99,8%), (93,4%, 99,7%), (2,2%, 54,5%) e (0,0%, 1,8%) de amplitudes 52,3%, 1,8%, 18,3%, 6,3%, 52,3% e 1,8%, respectivamente. Para as meninas classificadas como obesas, considerando a combinação dos dois critérios, encontram-se diferenças entre estes critérios para as sensibilidades e especificidades. O critério de Must para obesidade estima sensibilidade superior ao critério de Cole, mas especificidade inferior.

Com respeito aos intervalos de credibilidade, tem-se que os obtidos para a especificidade, valor preditivo positivo e taxa de erro do resultado falso positivo pelo critério de Cole são mais precisos, mas para a sensibilidade, valor preditivo negativo e taxa de erro do resultado falso negativo menos precisos.

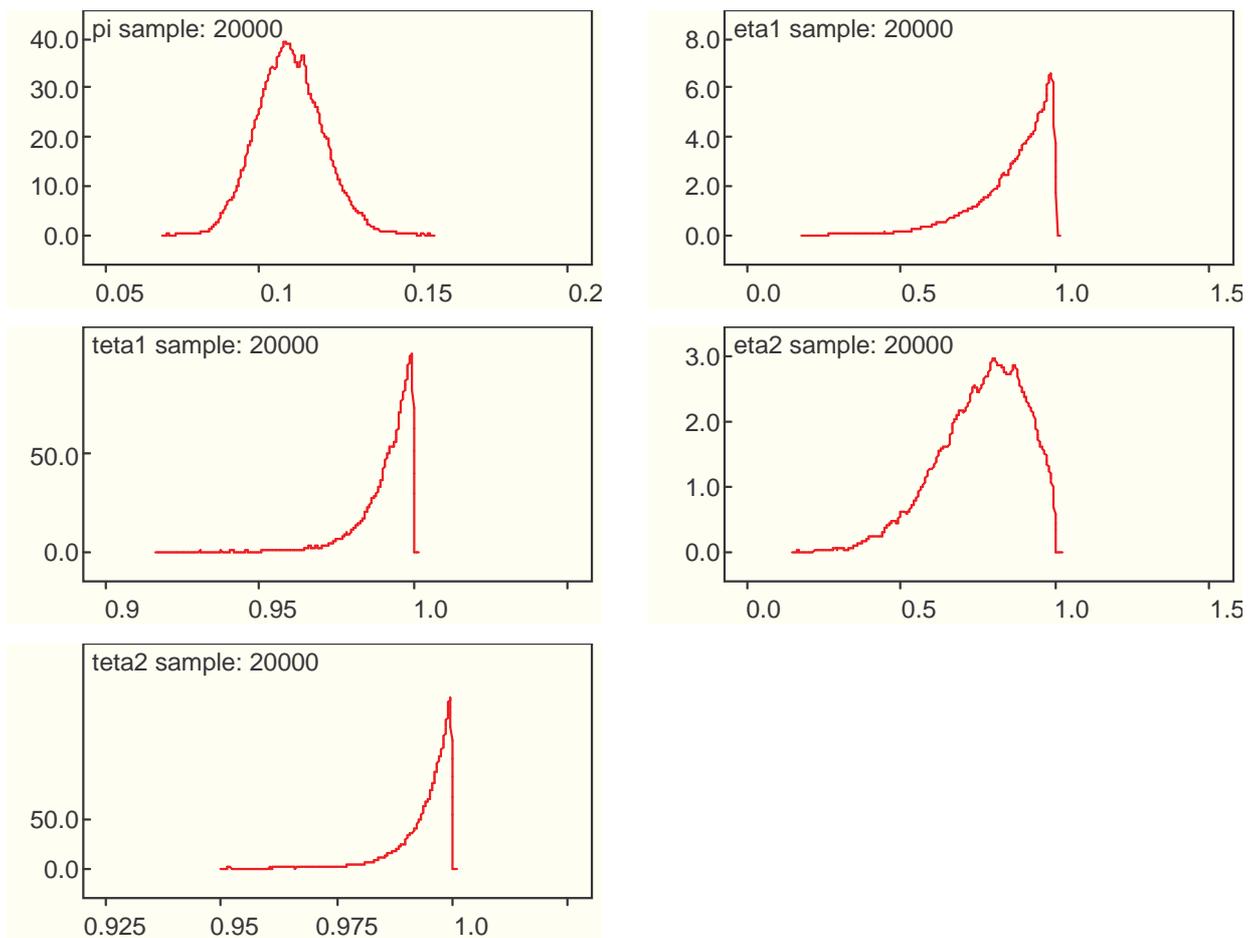
**Tabela 5.12.** Valores estimados via *Gibbs Sampling* para os dados latentes da combinação do critério de Must e Cole, com  $u = 9$ ,  $v = 2$ ,  $w = 0$  e  $x = 188$  para o sexo

feminino.					
parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	8,995	0,072	9,0	9,0	9,0
$Y_2$	1,462	0,7067	0,0	2,0	2,0
$Y_3$	---	---	---	---	---
$Y_4$	0,8537	1,484	0,0	0,0	5,0



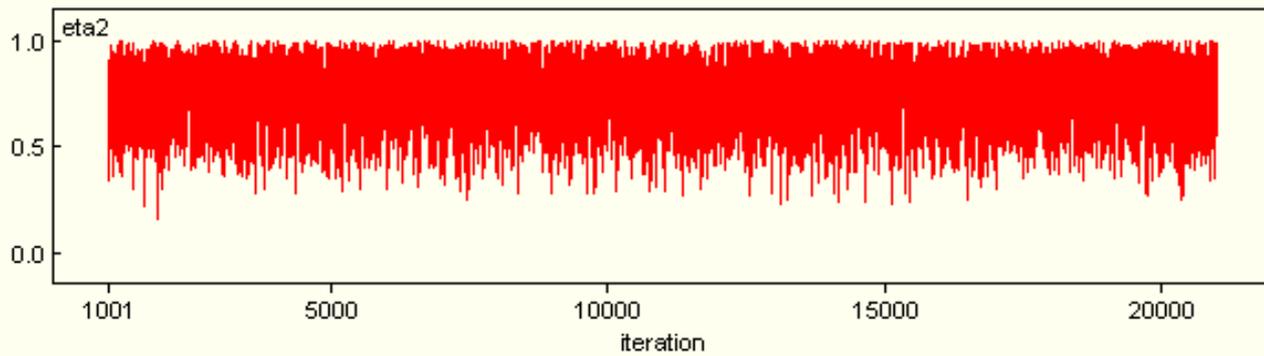
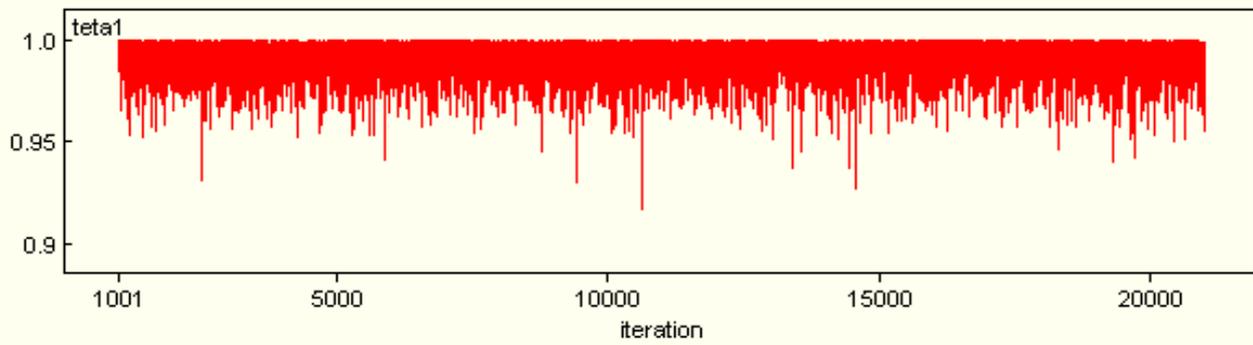
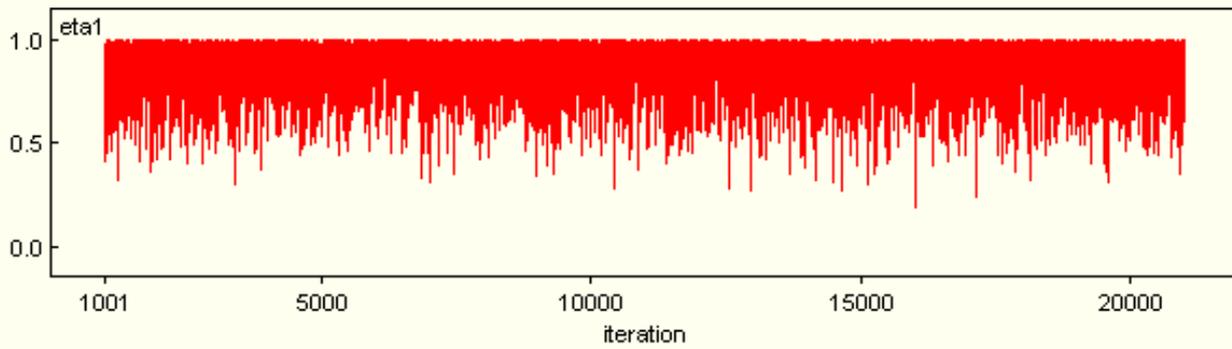
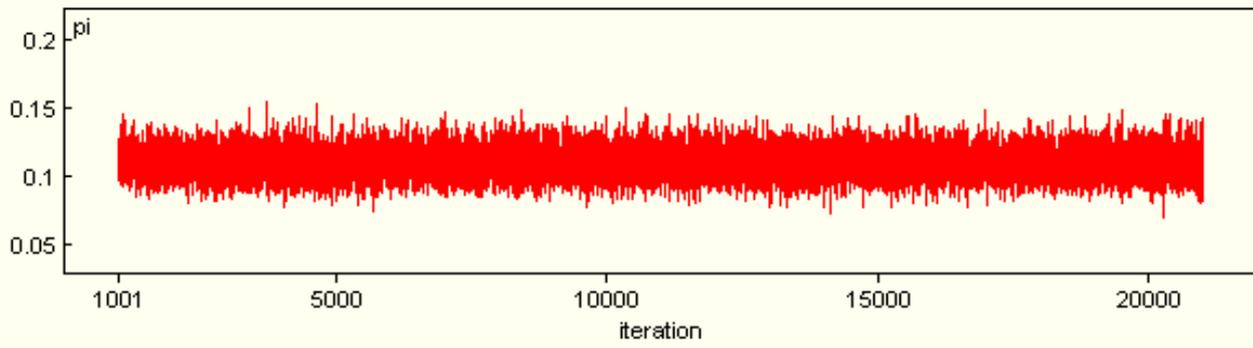
**Figura 5.12.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1, Y_2,$  e  $Y_4$  da combinação dos dois critérios para o sexo feminino.

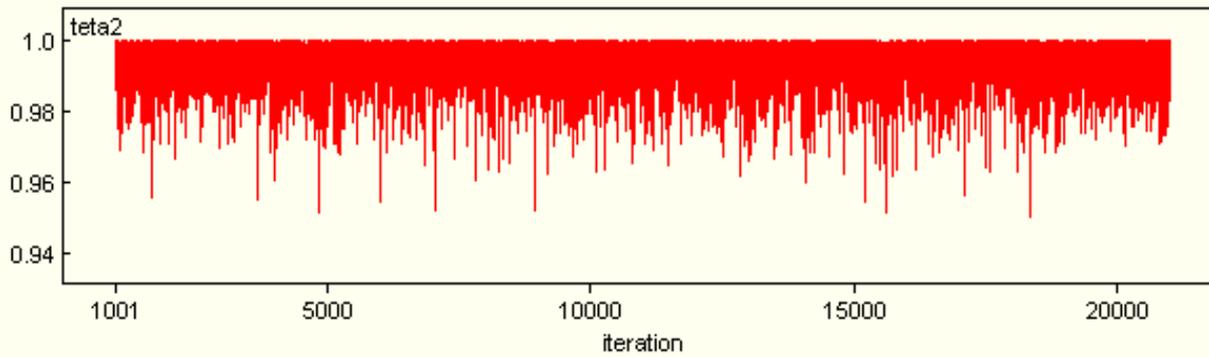
A partir da Tabela 5.12 e da Figura 5.12 tem-se que aproximadamente 9 escolares do sexo feminino com resultados positivos em ambos os critérios tenham resultados verdadeiros positivos. É provável que aproximadamente 1 escolar do sexo feminino dos 11 com resultado positivo para o critério de Must mas com Cole negativo tenha resultado verdadeiro positivo. Finalmente, do dado latente e histograma para  $Y_4$ , parece provável que aproximadamente 1 escolar do sexo feminino com resultado negativo em ambos os critérios tenha resultado verdadeiro positivo.



**Figura 5.13.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta_1$ ,  $\theta_1$ ,  $\eta_2$  e  $\theta_2$  da combinação dos critérios antropométricos para o sexo feminino.

Como é evidente da Tabela 5.11 e da Figura 5.13, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da sensibilidade de Must usando dados da combinação de ambos os critérios foi 86,9%, embora o intervalo de credibilidade de 95% foi de (57,4%, 99,6%), esta mesma situação também ocorre com os outros parâmetros.





**Figura 5.14.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos dois critérios para o sexo feminino.

Através da Figura 5.14, tem-se uma indicação de convergência para todos os parâmetros da combinação do critério de Must e Cole.

- **Sexo Masculino**

**Tabela 5.13.** Valores estimados via *Gibbs Sampling* para os parâmetros da combinação do critério de Must e Cole, com  $u = 20, v = 8, w = 0$  e  $x = 191$  para o sexo masculino.

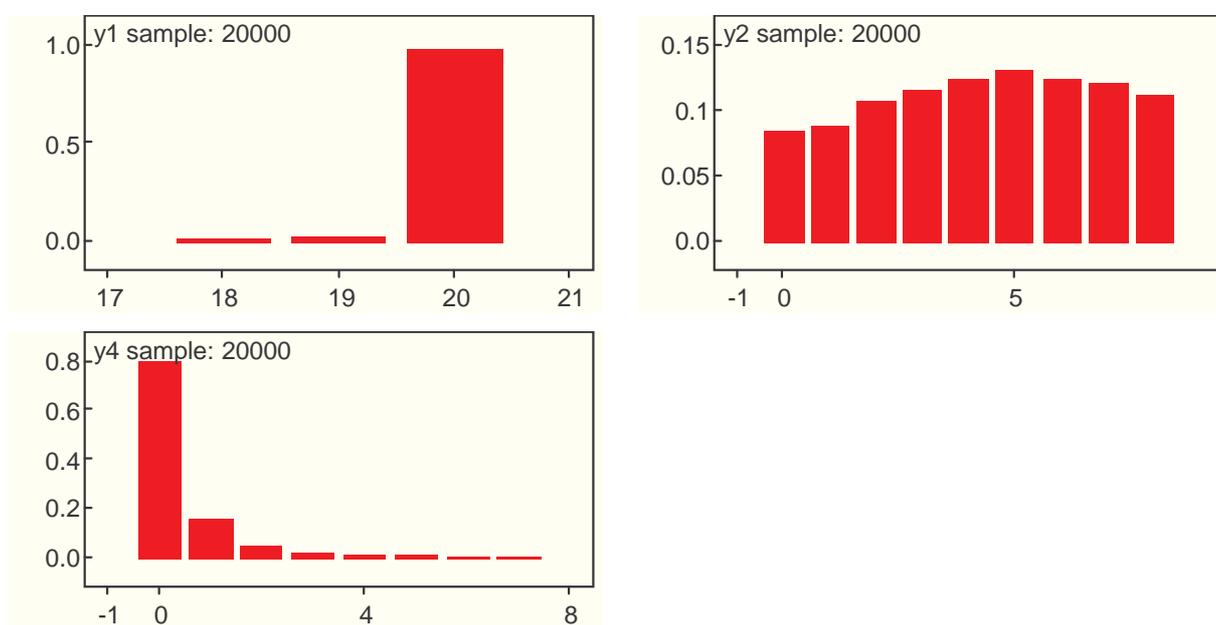
parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$\pi$	0,122	0,011	0,101	0,121	0,144
$\eta_1$	0,952	0,046	0,829	0,966	0,998
$\theta_1$	0,976	0,016	0,938	0,978	0,999
$\gamma_1$	0,852	0,089	0,667	0,859	0,991
$\delta_1$	0,993	0,006	0,976	0,995	0,999
$\eta_2$	0,799	0,110	0,576	0,802	0,984
$\theta_2$	0,995	0,005	0,980	0,996	0,999
$\gamma_2$	0,957	0,042	0,844	0,968	0,999
$\delta_2$	0,973	0,015	0,941	0,973	0,998
$\eta_1^*$	0,048	0,046	0,001	0,034	0,170
$\theta_1^*$	0,024	0,016	0,001	0,022	0,062
$\eta_2^*$	0,201	0,110	0,014	0,198	0,422
$\theta_2^*$	0,005	0,005	0,000	0,004	0,019

A Tabela 5.13 apresenta as médias, desvios padrões, medianas e intervalos de credibilidade para a prevalência de obesidade, sensibilidade, especificidade, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo para o sexo masculino, segundo os dois critérios antropométricos. Para os dois critérios combinados a prevalência de obesidade foi em média 12,2%, e o intervalo de credibilidade (10,1%, 14,4%) com amplitude de 4,3%. Para o critério de Must as estimativas das sensibilidades, especificidades, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade foram em média 95,2%, 97,6%, 85,2%, 99,3%, 4,8% e 2,4%, com intervalos de credibilidade (82,9%, 99,8%), (93,8%, 99,9%), (66,7%, 99,1%), (97,6%, 99,9%), (0,1%, 17,0%) e (0,1%, 6,2%) de amplitudes 16,9%, 6,1%, 32,4%, 2,3%, 16,9% e 6,1%, respectivamente. Já o critério de Cole apresentou sensibilidades, especificidades, valores preditivos positivo e negativo, assim como as taxas de erros dos resultados falso negativo e falso positivo de obesidade de 79,9%, 99,5%, 95,7%, 97,3%, 20,1% e 0,5%, com intervalos de credibilidade (57,6%, 98,4%), (98,0%, 99,9%), (84,4%, 99,9%), (94,1%, 99,8%), (1,4%, 42,2%) e (0,0%, 1,9%) de amplitudes 40,8%, 1,9%, 15,5%, 5,7%, 40,8% e 1,9%, respectivamente. Para os meninos classificados como obesos, considerando a combinação dos dois critérios, encontram-se diferenças entre estes critérios para as sensibilidades e especificidades. O critério de Must para obesidade estima sensibilidade superior ao critério de Cole, mas especificidade inferior.

Com respeito aos intervalos de credibilidade, tem-se que os obtidos para a especificidade, valor preditivo positivo e taxa de erro do resultado falso positivo pelo critério de Cole são mais precisos, mas para a sensibilidade, valor preditivo negativo e taxa de erro do resultado falso negativo menos precisos.

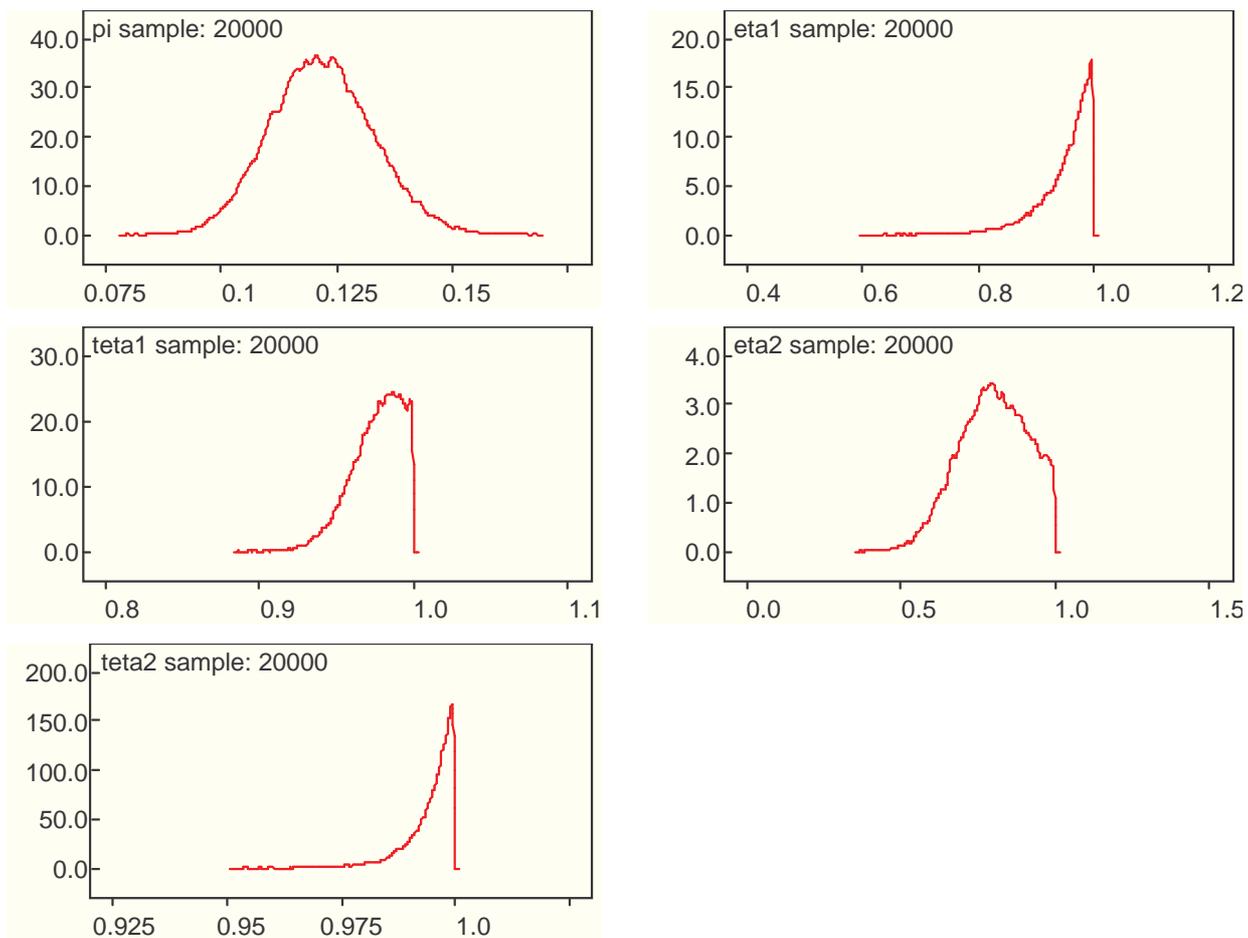
**Tabela 5.14.** Valores estimados via *Gibbs Sampling* para os dados latentes da combinação do critério de Must e Cole, com  $u = 20$ ,  $v = 8$ ,  $w = 0$  e  $x = 191$  para o sexo

masculino.					
parâmetro	média	desvio padrão	2,5%	mediana	97,5%
$Y_1$	19,98	0,155	20,0	20,0	20,0
$Y_2$	4,263	2,465	0,0	4,0	8,0
$Y_3$	---	---	---	---	---
$Y_4$	0,275	0,615	0,0	0,0	2,0



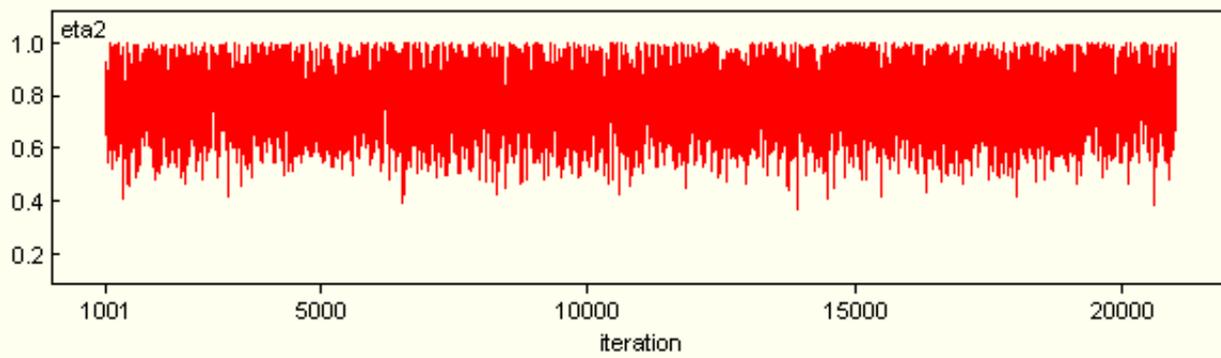
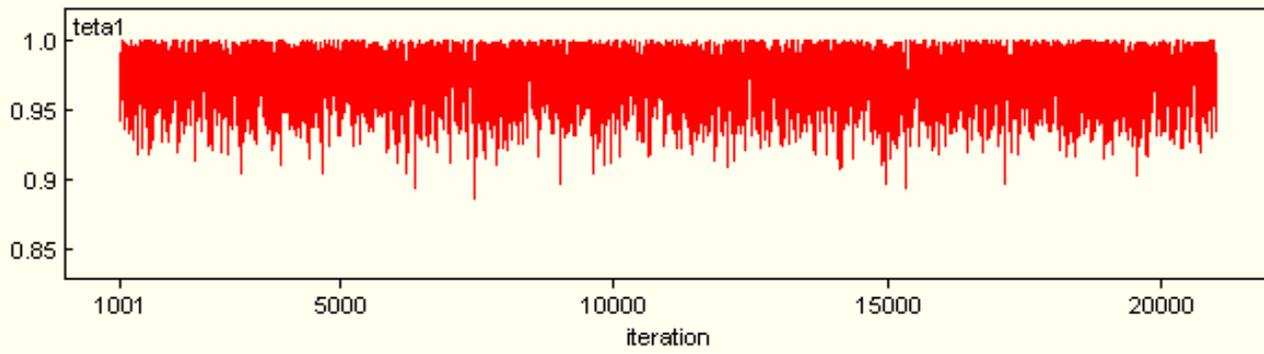
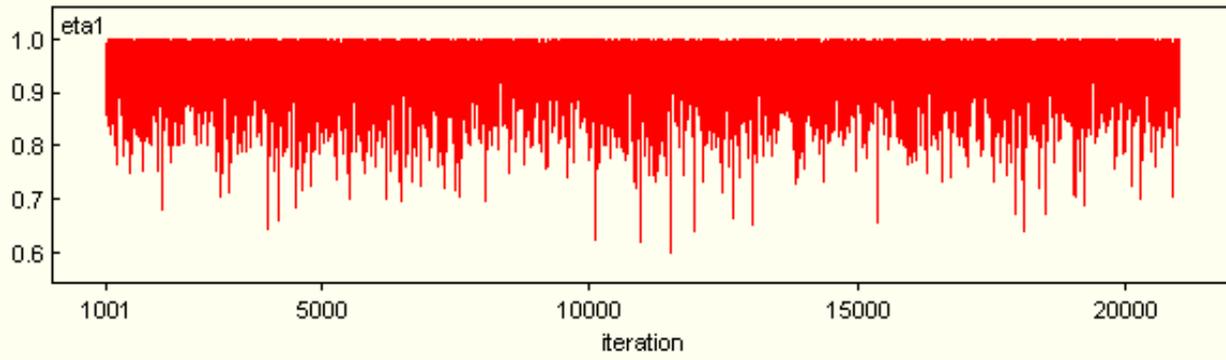
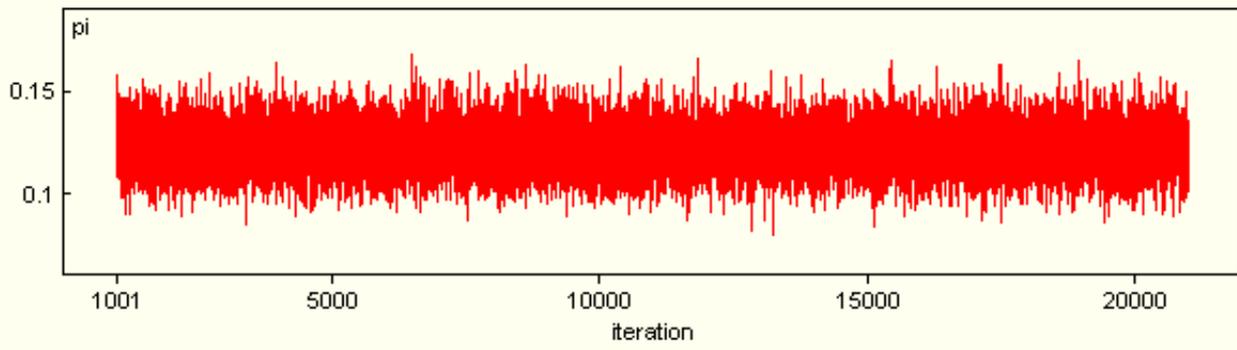
**Figura 5.15.** Distribuições estimadas via *Gibbs Sampling* para os dados latentes  $Y_1, Y_2$  e  $Y_4$  da combinação dos dois critérios para o sexo masculino.

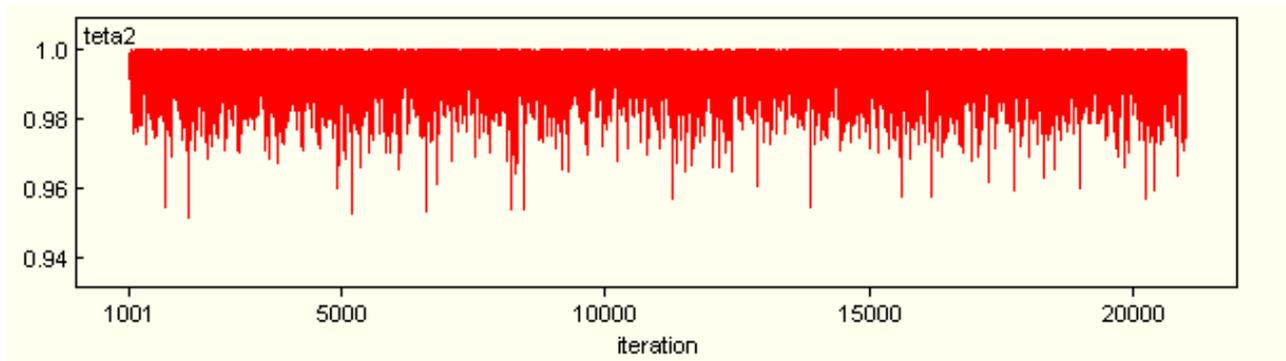
A partir da Tabela 5.14 e da Figura 5.15 tem-se que aproximadamente 20 escolares do sexo masculino com resultados positivos em ambos os critérios tenham resultados verdadeiros positivos. É provável que aproximadamente 4 escolares do sexo masculino dos 28 com resultado positivo para o critério de Must mas com Cole negativo tenham resultados verdadeiros positivos. Finalmente, do dado latente e histograma para  $Y_4$ , parece provável que nenhum escolar do sexo masculino com resultado negativo em ambos os critérios tenha resultado verdadeiro positivo.



**Figura 5.16.** Densidades estimadas via *Gibbs Sampling* para os parâmetros  $\pi$ ,  $\eta_1$ ,  $\theta_1$ ,  $\eta_2$  e  $\theta_2$  da combinação dos critérios antropométricos para o sexo masculino.

Como é evidente da Tabela 5.13 e da Figura 5.16, as densidades são assimétricas. Por exemplo, a média da distribuição marginal *a posteriori* da sensibilidade para o critério de Must usando dados da combinação de ambos os critérios foi 95,2%, embora o intervalo de credibilidade de 95% foi de (82,9%, 99,8%), esta mesma situação também ocorre com os outros parâmetros.





**Figura 5.17.** Comportamento das trajetórias das cadeias ao longo das iterações para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos dois critérios para o sexo masculino.

Através da Figura 5.17, tem-se uma indicação de convergência para todos os parâmetros da combinação do critério de Must e Cole para o sexo masculino.

## 5.4 A Indicação de Convergência

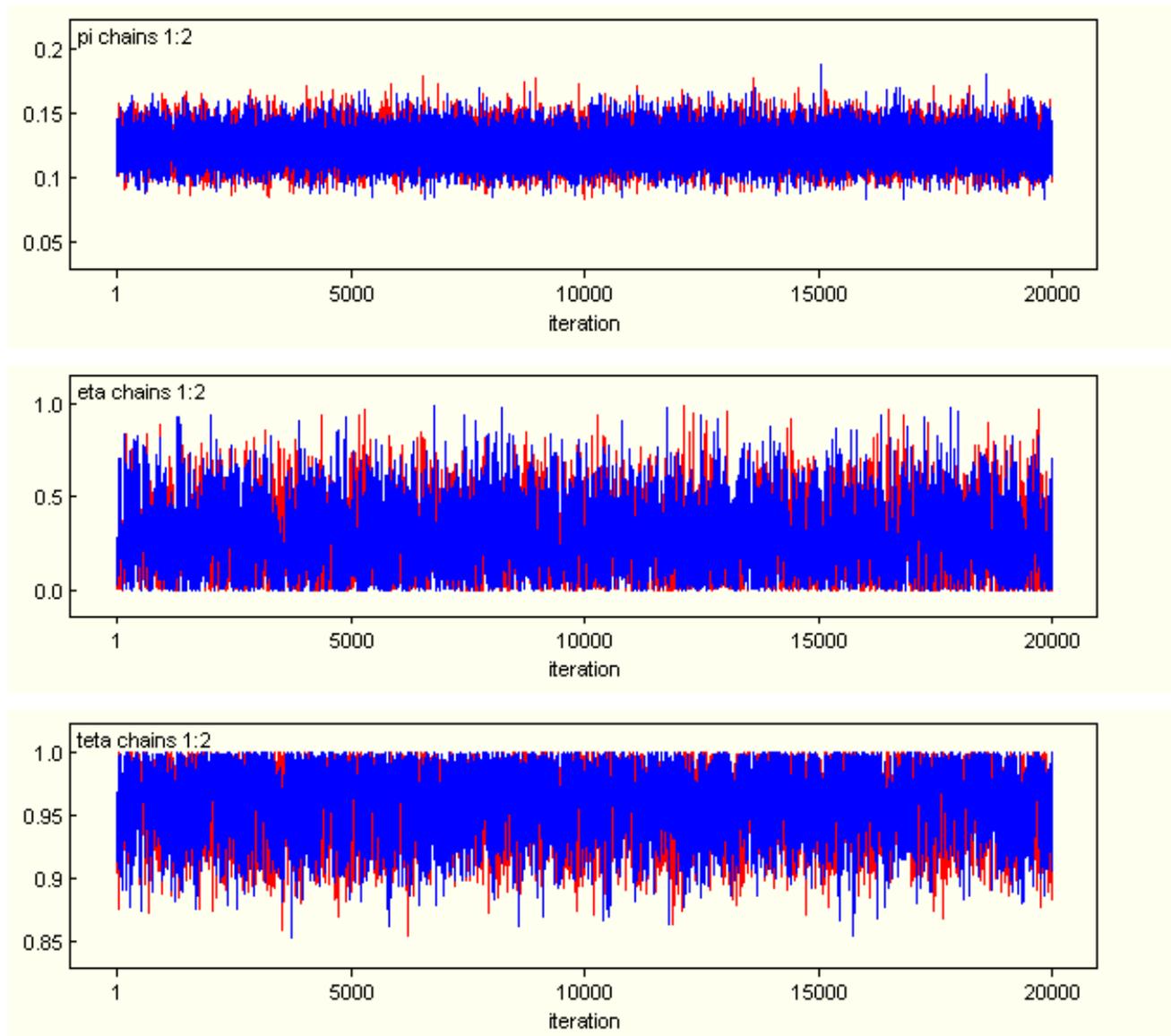
Como exemplo, para análise da indicação de convergência, os resultados da aplicação do diagnóstico de Gelman & Rubin (1992), descrito no Apêndice D, são apresentados para os dados de obesidade em 418 escolares. Estes resultados são apresentados para o critério de Must, Cole e para a combinação destes dois critérios.

Para esta análise foram geradas duas cadeias, com 20000 valores cada uma.

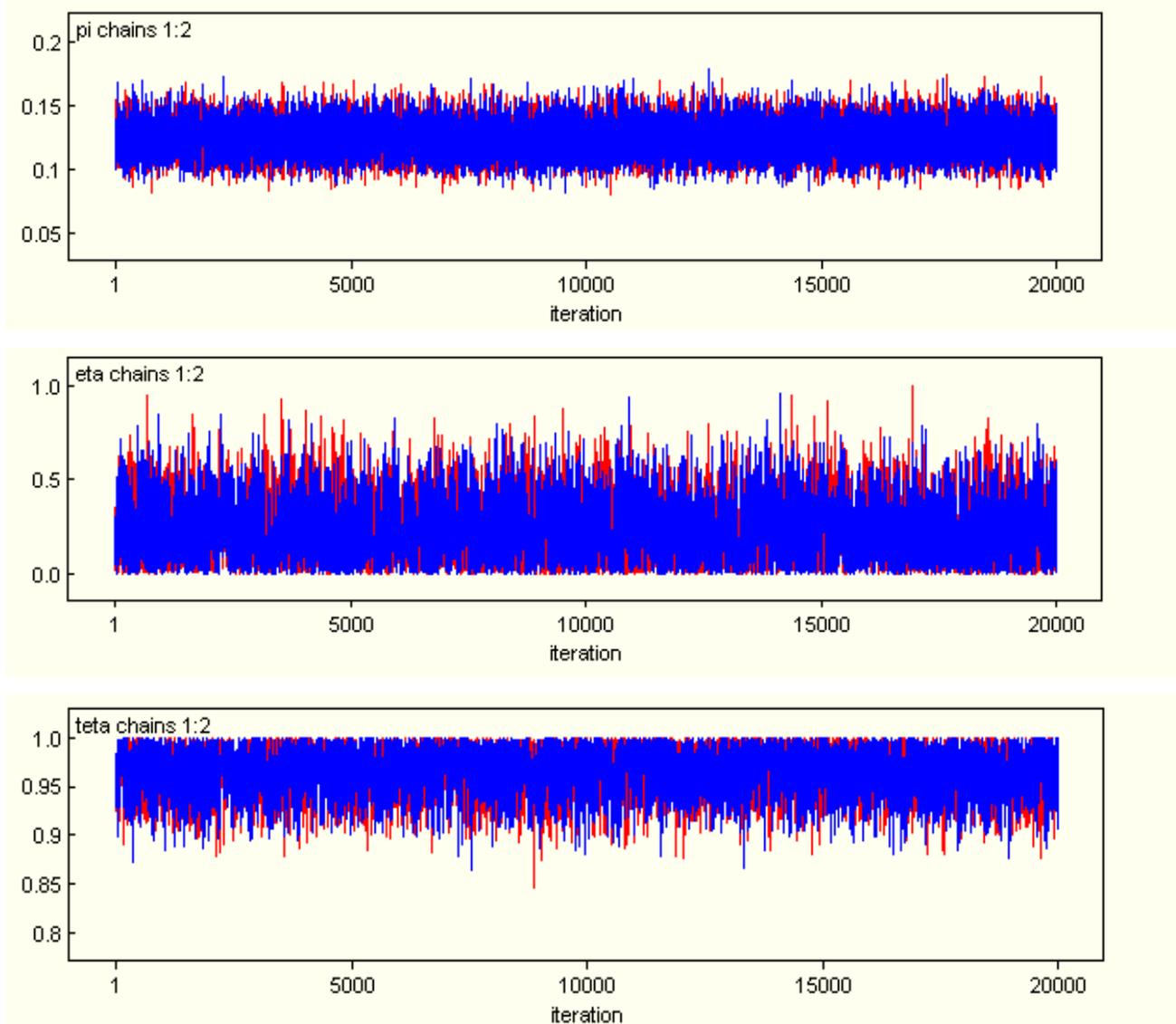
### 5.4.1 Critérios Antropométricos Must e Cole para o Sexo Feminino

As Figuras 5.18 e 5.19 ilustram o comportamento das trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta$  e  $\theta$  do critério Must e Cole, respectivamente, com as iterações iniciais não descartadas. Através destas Figuras, verifica-se

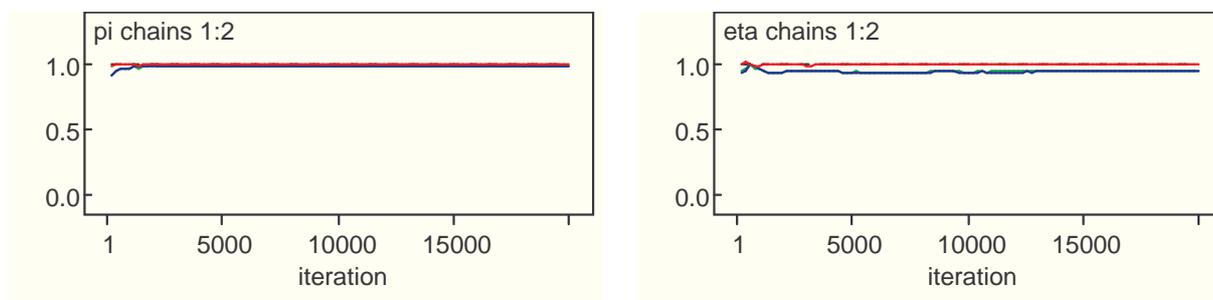
que a indicação de convergência é quase imediata, ou seja, mesmo antes de 1000 iterações já existe a indicação de convergência. As Figura 5.20 e 5.21 mostram os gráficos fornecidos pelo *Software WinBUGS* para o diagnóstico de Gelman & Rubin.

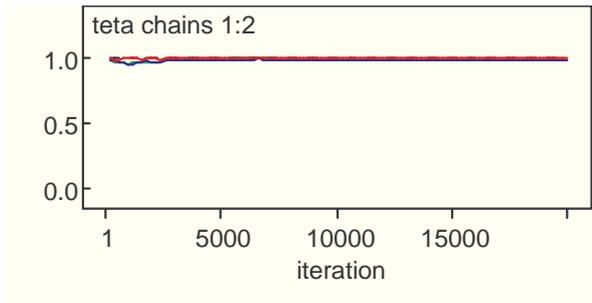


**Figura 5.18.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério Must para o sexo feminino.

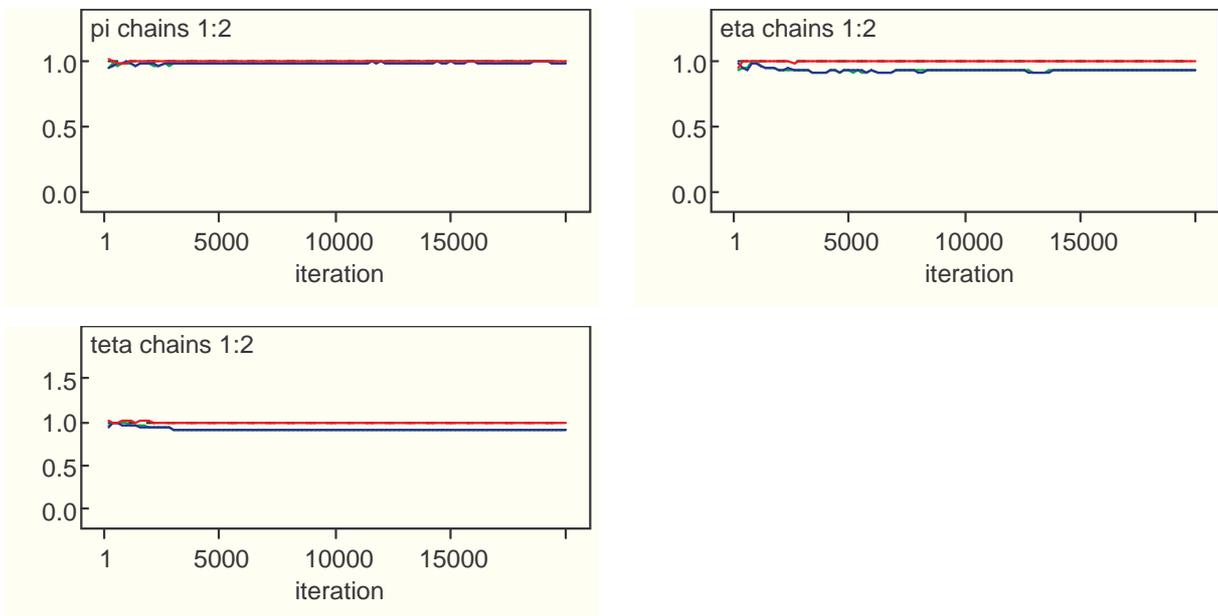


**Figura 5.19.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta$  e  $\theta$  do critério Cole para o sexo feminino.





**Figura 5.20.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi, \eta$  e  $\theta$  do critério Must para o sexo feminino. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas, a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = \text{combinada}/\text{individual}$ .



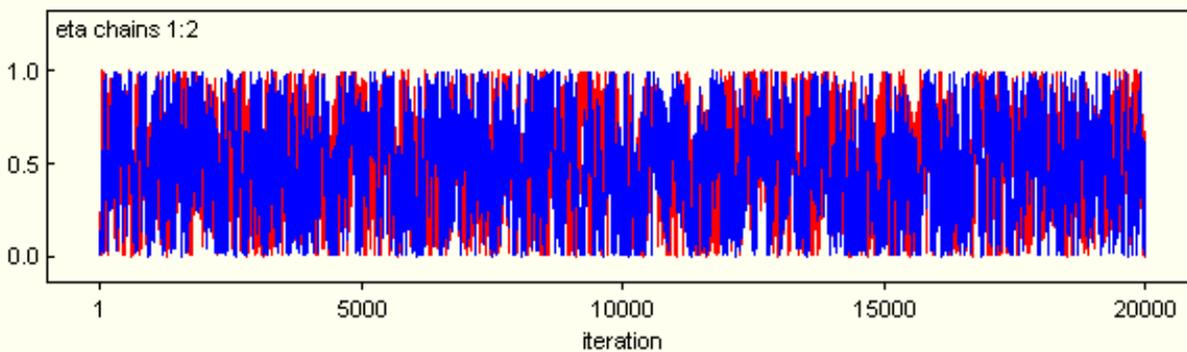
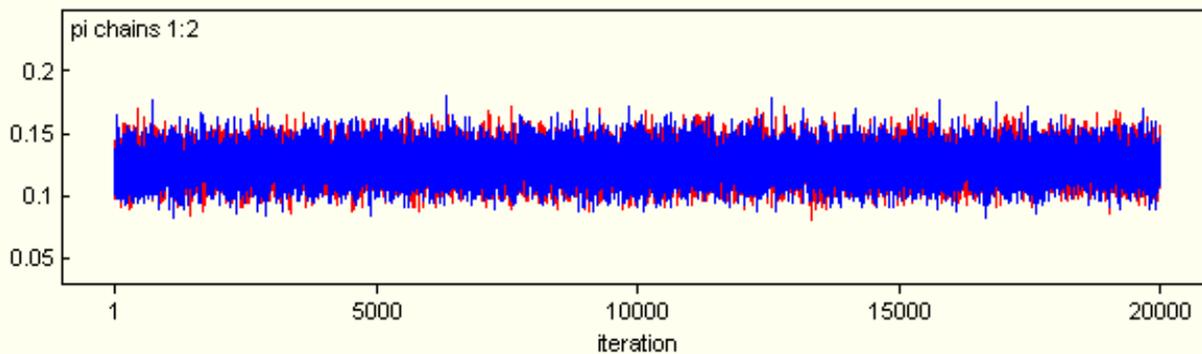
**Figura 5.21.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi, \eta$  e  $\theta$  do critério Cole para o sexo feminino. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas, a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = \text{combinada}/\text{individual}$ .

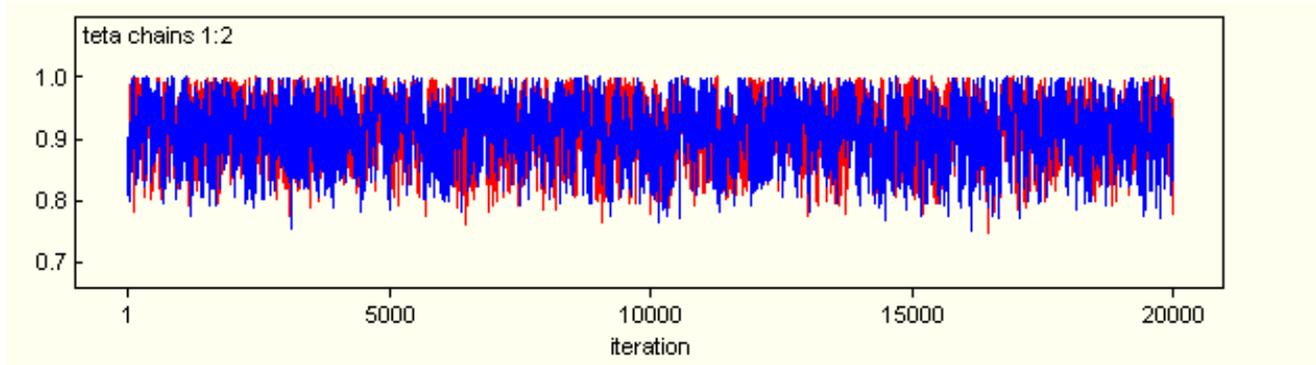
Através das Figuras 5.20 e 5.21, verifica-se que a largura dos intervalos nas cadeias combinadas e a largura média dos intervalos dentro de cada cadeia individual estão praticamente sobrepostas. O fator de redução  $R$  converge para 1 à medida que aumenta o número de iterações, tendo assim a indicação de convergência para todos os parâmetros

do critério de Must e Cole para o sexo feminino, respectivamente.

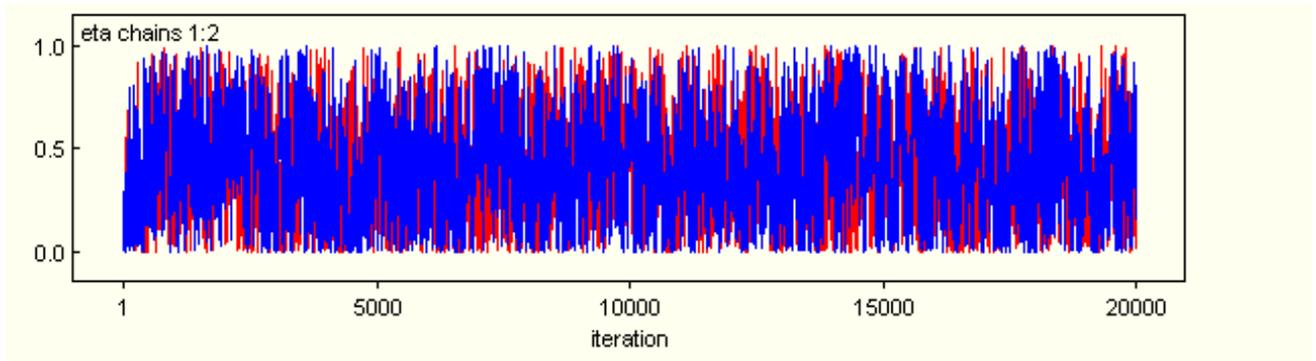
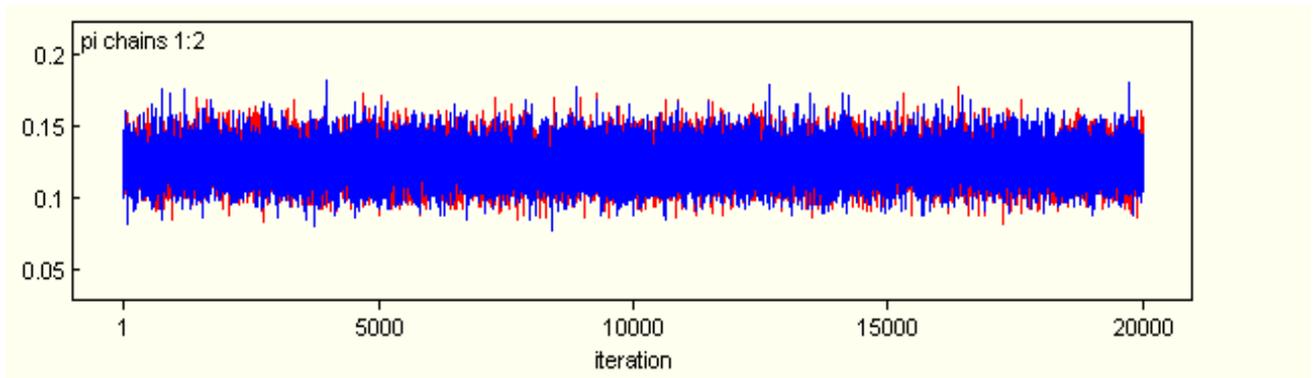
## 5.4.2 Critérios Antropométricos Must e Cole para o Sexo Masculino

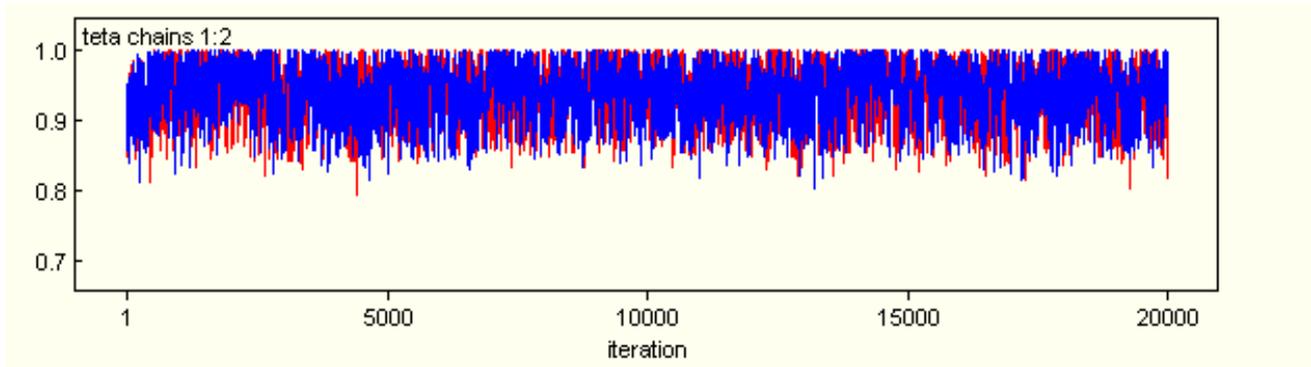
As Figuras 5.22 e 5.23 ilustram o comportamento das trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta$  e  $\theta$  do critério Must e Cole, respectivamente, com as iterações iniciais não descartadas. Através destas Figuras, verifica-se que a indicação de convergência é quase imediata, ou seja, mesmo antes de 1000 iterações já existe a indicação de convergência. As Figura 5.24 e 5.25 mostram os gráficos fornecidos pelo *Software WinBUGS* para o diagnóstico de Gelman & Rubin.



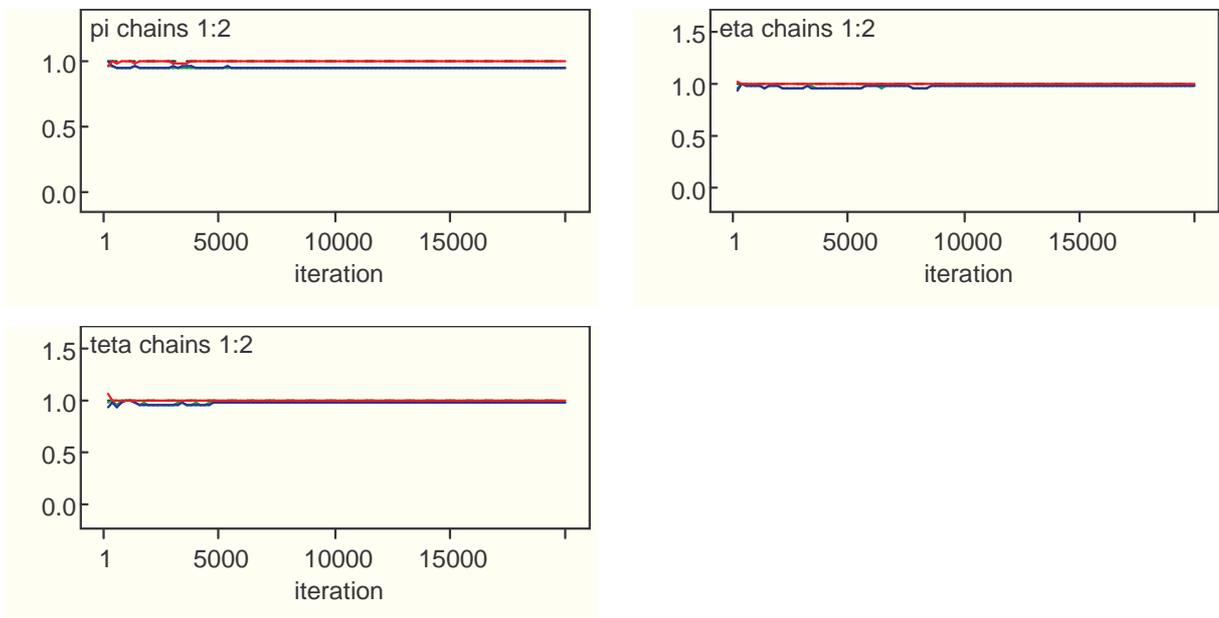


**Figura 5.22.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério Must para o sexo masculino.

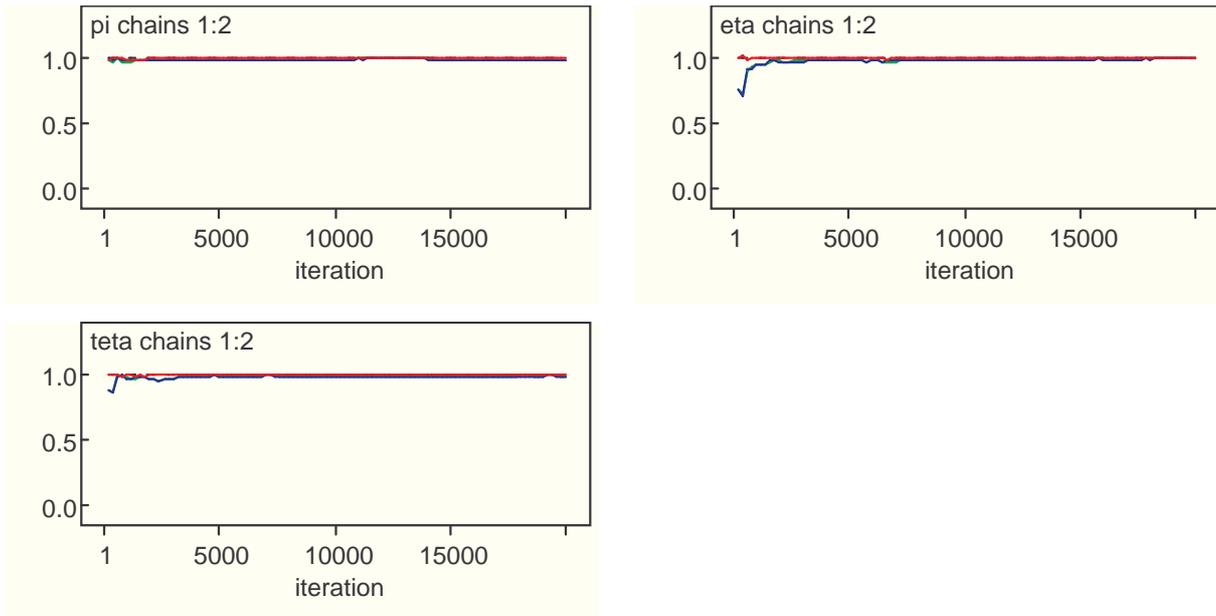




**Figura 5.23.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta$  e  $\theta$  do critério Cole para o sexo masculino.



**Figura 5.24.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi, \eta$  e  $\theta$  do critério Must para o sexo masculino. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas, a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = combinada/individual$ .



**Figura 5.25.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi$ ,  $\eta$  e  $\theta$  do critério Cole para o sexo masculino. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas, a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = combinada/individual$ .

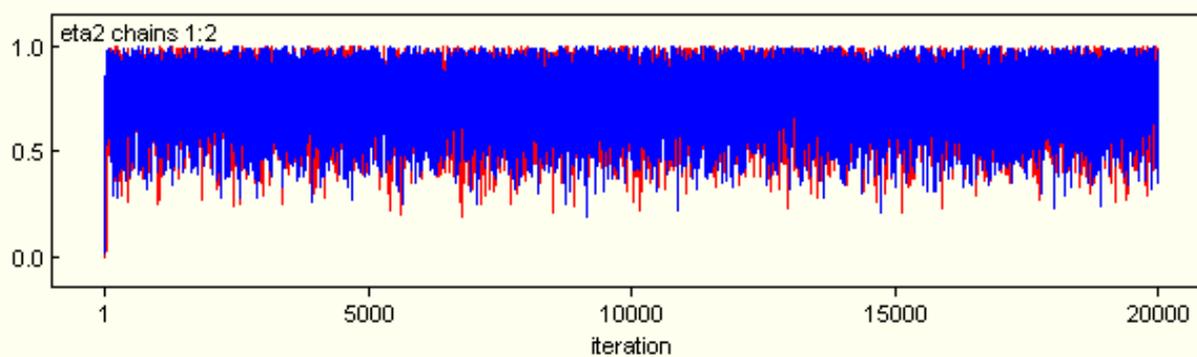
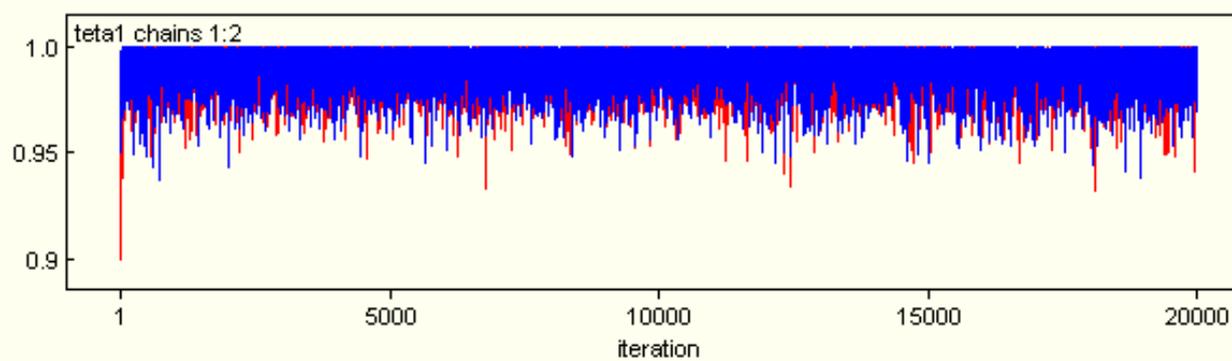
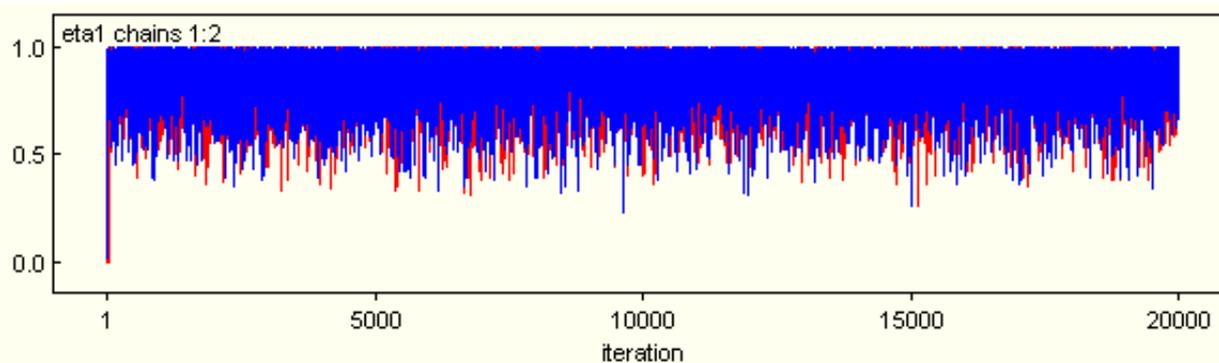
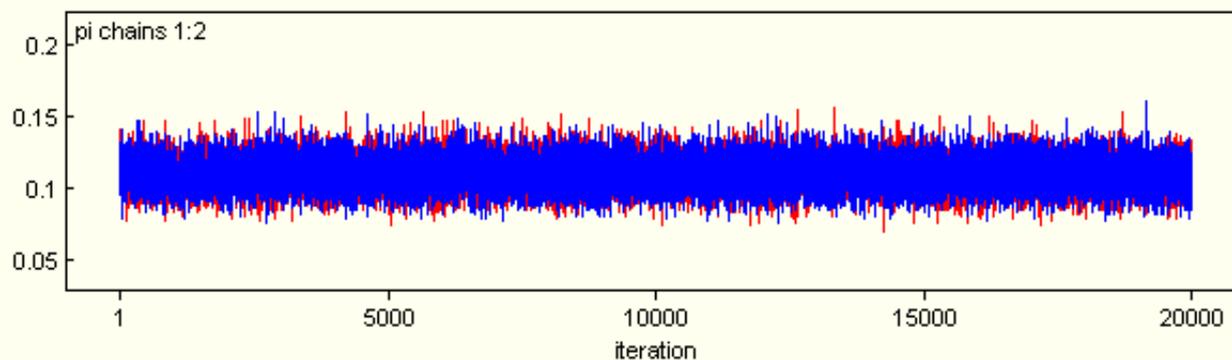
Através das Figuras 5.24 e 5.25, verifica-se que a largura dos intervalos nas cadeias combinadas e a largura média dos intervalos dentro de cada cadeia individual estão praticamente sobrepostas. O fator de redução  $R$  converge para 1 à medida que aumenta o número de iterações, tendo assim a indicação de convergência para todos os parâmetros do critério de Must e Cole para o sexo masculino, respectivamente.

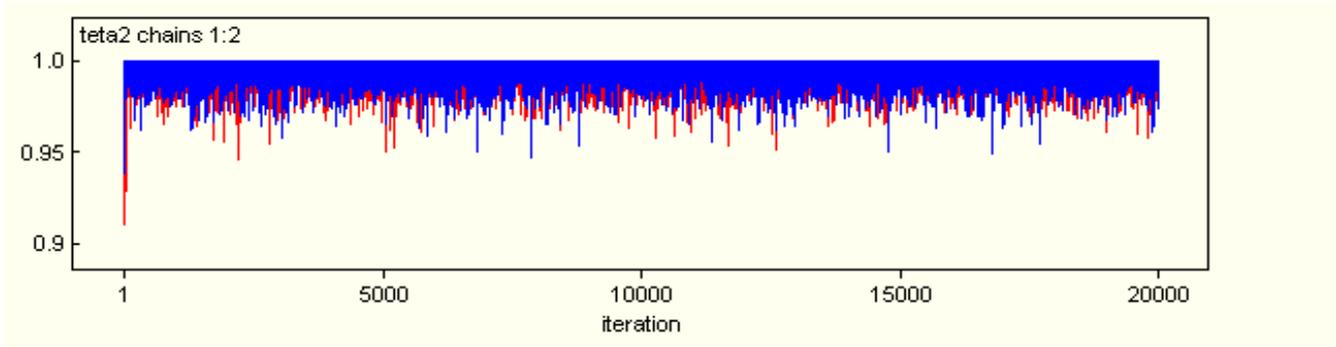
### 5.4.3 Combinação dos Critério Must e Cole para ambos os Sexos

- **Sexo Feminino**

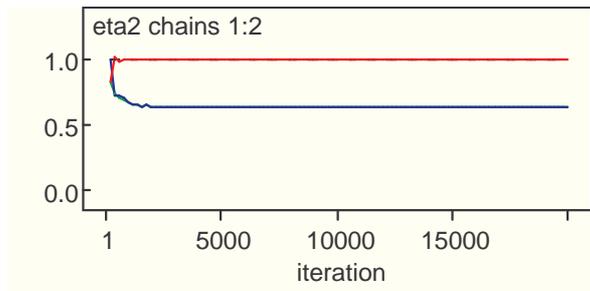
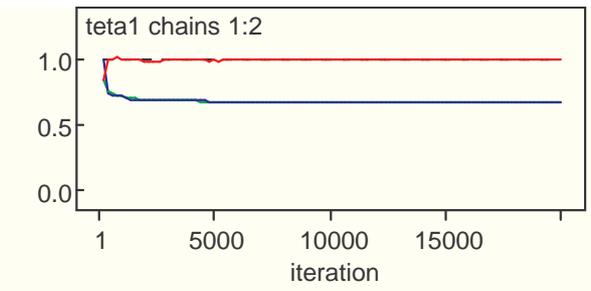
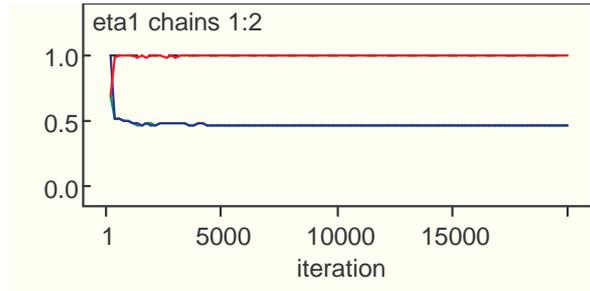
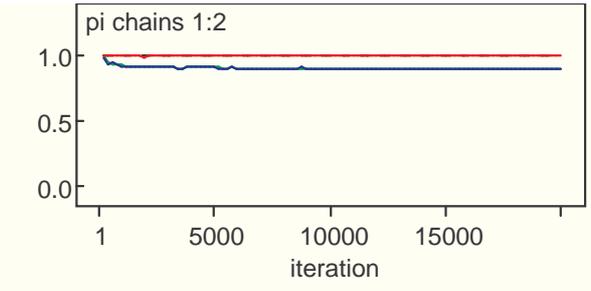
A Figura 5.26 ilustra o comportamento das trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi$ ,  $\eta_1$ ,  $\theta_1$ ,  $\eta_2$  e  $\theta_2$  da combinação dos critérios Must e Cole para o sexo feminino, com as iterações iniciais não descartadas. Através desta Figura, verifica-se que a indicação de convergência é quase imediata, ou seja, mesmo

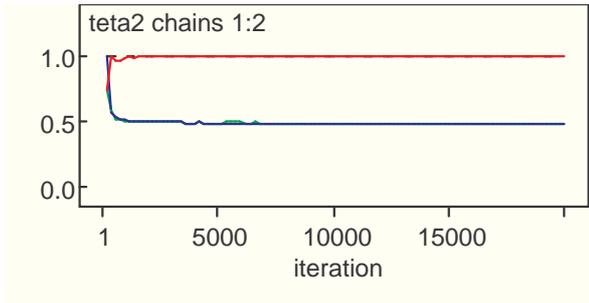
antes de 1000 iterações já existe a indicação de convergência. A Figura 5.27 mostra os gráficos fornecidos pelo *Software WinBUGS* para o diagnóstico de Gelman & Rubin.





**Figura 5.26.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos critérios para o sexo feminino.



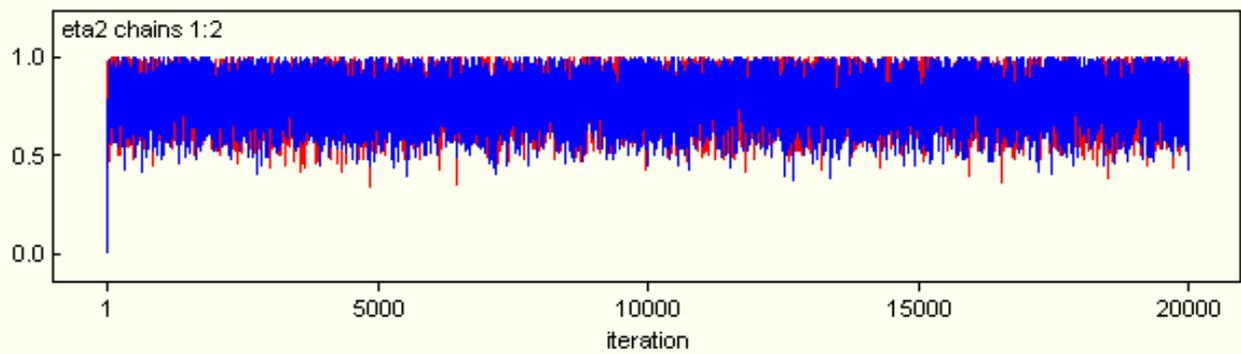
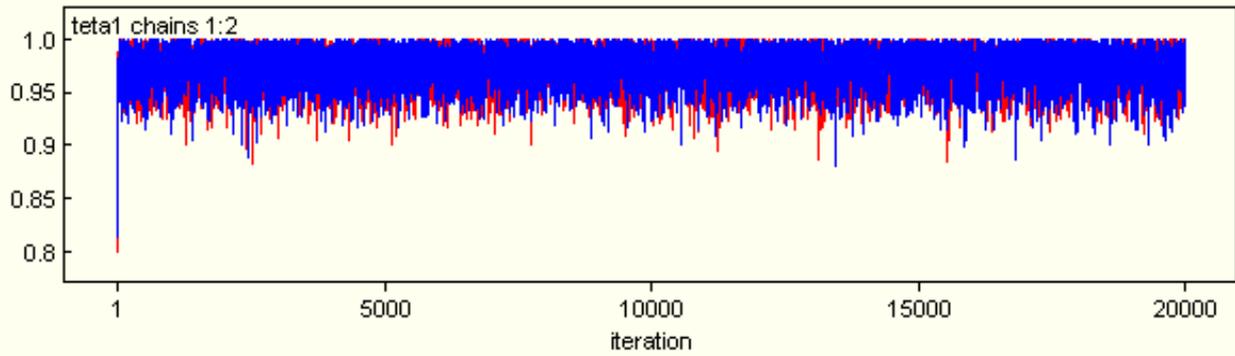
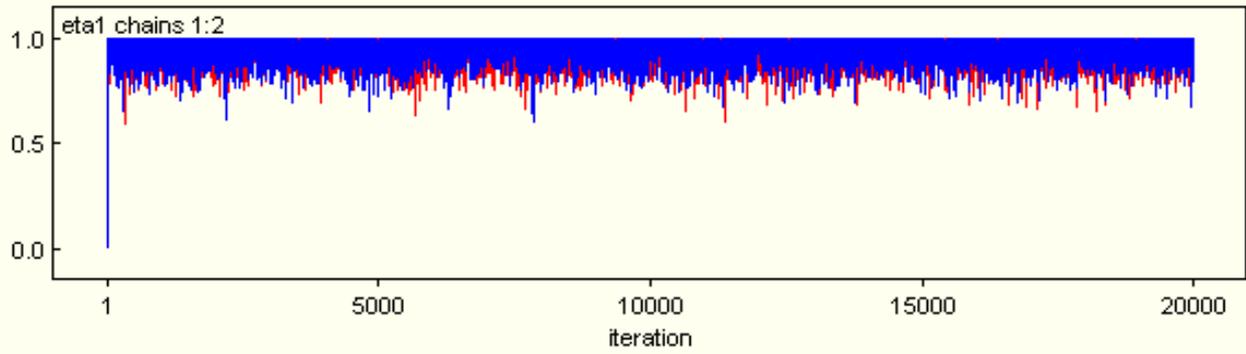
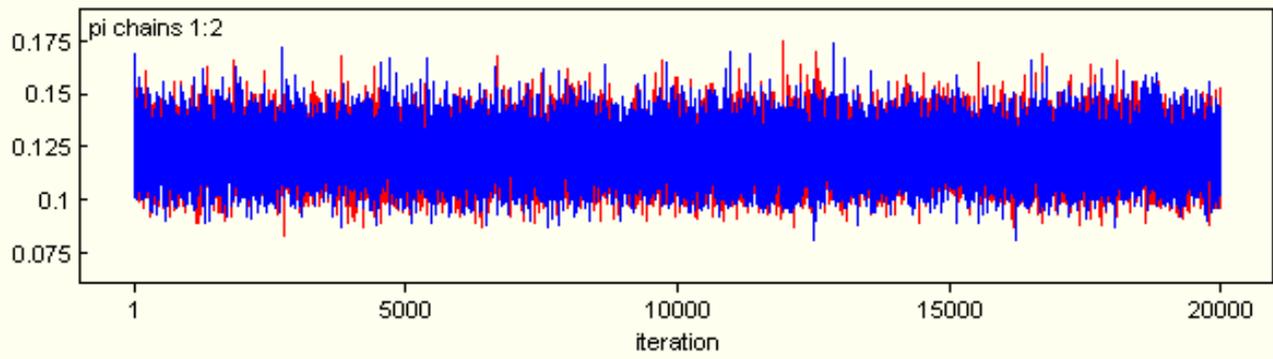


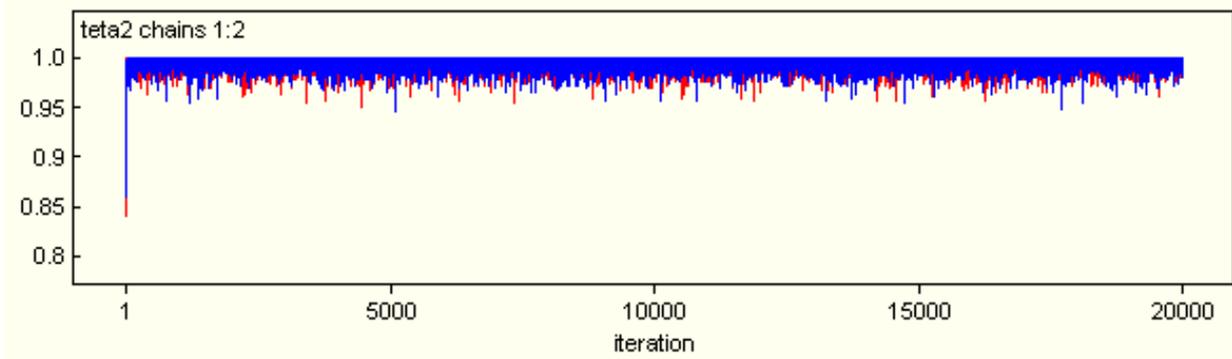
**Figura 5.27.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos dois critérios para o sexo feminino. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = combinada/individual$ .

Através da Figura 5.27, verifica-se que a largura dos intervalos nas cadeias combinadas e a largura média dos intervalos dentro de cada cadeia individual estão praticamente sobrepostas. O fator de redução  $R$  converge para 1 à medida que aumenta o número de iterações, tendo assim a indicação de convergência para todos os parâmetros da combinação dos dois critérios antropométricos para o sexo feminino.

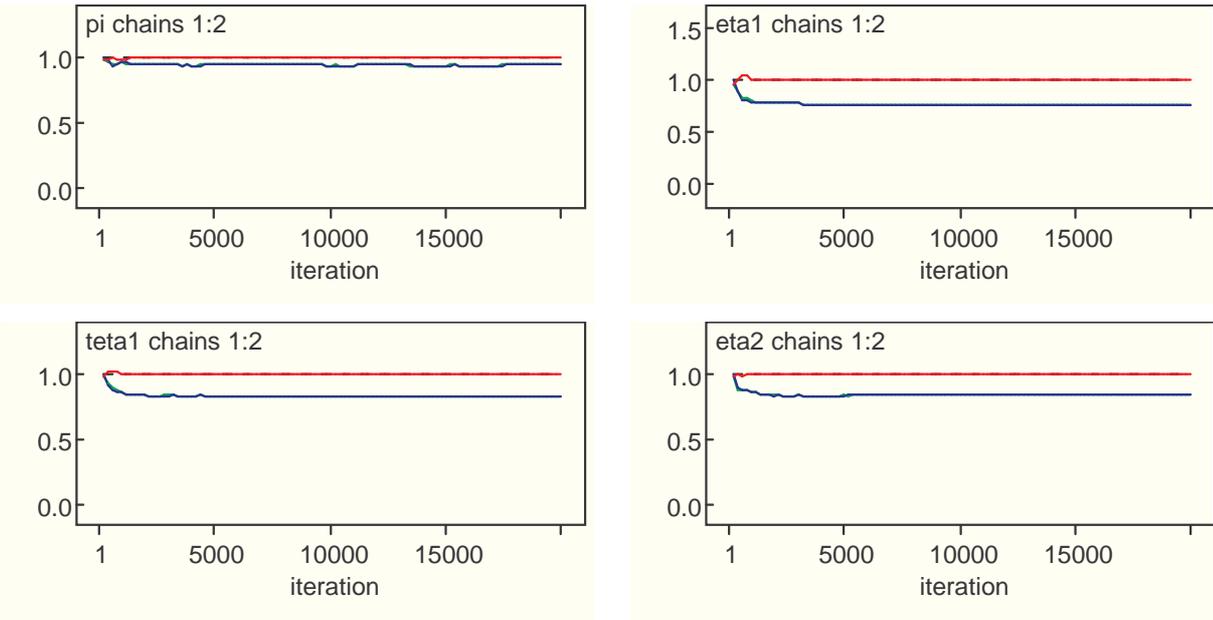
- **Sexo Masculino**

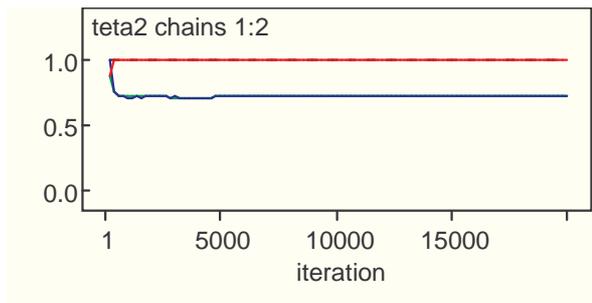
A Figura 5.28 ilustra o comportamento das trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos critérios Must e Cole para o sexo masculino, com as iterações iniciais não descartadas. Através desta Figura, verifica-se que a indicação de convergência é quase imediata, ou seja, mesmo antes de 1000 iterações já existe a indicação de convergência. A Figura 5.29 mostra os gráficos fornecidos pelo *Software WinBUGS* para o diagnóstico de Gelman & Rubin.





**Figura 5.28.** Trajetórias para as duas cadeias geradas via *Gibbs Sampling*, para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos critérios para o sexo masculino.





**Figura 5.29.** Fator de redução na escala ( $R$ ) para os parâmetros  $\pi, \eta_1, \theta_1, \eta_2$  e  $\theta_2$  da combinação dos dois critérios para o sexo masculino. A cor verde representa a largura do intervalo de 80%, nas cadeias combinadas a azul a largura média destes intervalos dentro de cada cadeia individual e o vermelho o fator de redução  $R = combinada/individual$ .

Através da Figura 5.29, verifica-se que a largura dos intervalos nas cadeias combinadas e a largura média dos intervalos dentro de cada cadeia individual estão praticamente sobrepostas. O fator de redução  $R$  converge para 1 à medida que aumenta o número de iterações, tendo assim a indicação de convergência para todos os parâmetros da combinação dos dois critérios antropométricos para o sexo masculino.

Como tanto para os parâmetros estimados para o critério Must, Cole e combinação de ambos, para cada sexo, realizou-se a indicação de convergência, pode-se utilizar as estatísticas fornecidas pelo *Software WinBUGS* para interpretar a presente aplicação.

## 5.5 Comentários

Os algoritmos *Gibbs Sampling* de um único teste diagnóstico, assim como de dois testes diagnósticos desenvolvidos no Capítulo 4 baseado no procedimento proposto por Joseph *et al.* (1995), foram implementados para a estimação dos parâmetros de cada critério antropométrico (Must e Cole) separadamente, e da combinação de ambos. Esta metodologia foi aplicada aos dados de obesidade de 418 escolares, 199 do sexo feminino e 219 do sexo masculino, com distribuição *a priori* informativa para a prevalência de

obesidade e *prioris* não informativas para as sensibilidades e especificidades do critério de Must e Cole, tanto para o sexo feminino quanto para o masculino. A partir das estimativas dos parâmetros de interesse para o critério de Must e Cole separadamente, como para a combinação destes dois critérios observou-se que:

1. Para os escolares do sexo feminino, considerando cada critério separadamente, as prevalências de obesidade encontradas pelo critério de Must e Cole foram de 12,4% para os dois; sensibilidades de 26,7% e 22,5%, especificidades de 96,4% e 96,9%, valores preditivos positivos de 93,1% e 94,8%, valores preditivos negativos de 98,4% e 97,2%, taxas de erros dos resultados falsos negativos de 73,3% e 77,5%, e taxas de erros dos resultados falsos positivos de 3,6% e 3,1%, respectivamente. Para os escolares do sexo masculino, as prevalências de obesidade segundo Must e Cole foram de 12,5% e 12,4%; sensibilidades de 47,8% e 38,9%, especificidades de 91,6% e 94,2%, valores preditivos positivos de 85,2% e 95,7%, valores preditivos negativos de 99,3% e 97,3%, taxas de erros dos resultados falsos negativos de 52,2% e 61,1%, e taxas de erros dos resultados falsos positivos de 8,4% e 5,8%, respectivamente. As amplitudes dos intervalos de credibilidade do sexo feminino para as prevalências de obesidade, segundo cada critério antropométrico, foram de 4,9% e 4,8%; para as sensibilidades de 63,4% e 55,2%, especificidades de 8,7% e 7,6%, valores preditivos positivos de 21,7% e 18,3%, valores preditivos negativos de 5,2% e 6,3%, taxas de erros dos resultados falsos negativos de 63,4% e 55,2%, e taxas de erros dos resultados falsos positivos de 8,7% e 7,6%, respectivamente. E as amplitudes do sexo masculino para as prevalências de obesidade, segundo cada critério antropométrico, foram de 4,9% e 4,8%; para as sensibilidades de 94,1% e 86,8%, especificidades de 16,4% e 12,6%, valores preditivos positivos de 32,4% e 15,5%, valores preditivos negativos de 2,3% e 5,7%, taxas de erros dos resultados falsos negativos de 94,1% e 86,8%, e taxas de erros dos resultados falsos positivos de 16,4% e 12,6%.

Nota-se que quando cada critério antropométrico é considerado separadamente, o de Must estima sensibilidade, valor preditivo negativo e taxa de erro do resultado falso positivo de obesidade superior ao de Cole em ambos os sexos, mas com especificidade, valor

preditivo positivo e taxa de erro do resultado falso negativo inferior. A estimativa da prevalência de obesidade no sexo feminino foi de igual valor para os dois critérios, e no sexo masculino a de Must foi superior ao de Cole. Analisando as prevalências de obesidade, para cada critério separadamente, nota-se que não houve diferença significativa entre os sexos. Em relação as meninas e meninos, o critério de Must mostrou-se mais sensível (26,7% e 47,8%) para classificar obesidade do que o de Cole (22,5% e 38,9%), e menos específico (96,4% e 91,6%) para não classificar obesidade do que o de Cole (96,9% e 94,2%). Com isso, para ambos os sexos o valor preditivo negativo de Must é melhor do que de Cole, e o valor preditivo positivo em Cole é melhor do que de Must.

Com respeito aos intervalos de credibilidade, nota-se que os intervalos obtidos pelo critério de Cole para a prevalência de obesidade, sensibilidade, especificidade, valor preditivo positivo, como para a taxa de erro do resultado falso negativo e taxa de erro do resultado falso positivo, são mais precisos em ambos os sexos. Porém, o intervalo para o valor preditivo negativo é menos preciso, como mostram as Tabelas 5.15 e 5.16, respectivamente.

**Tabela 5.15.** Intervalos de credibilidade e suas amplitudes para o critério de Must e Cole considerados separadamente, para o sexo feminino.

	Must		Cole	
	ICr	amplitude	ICr	amplitude
$\pi$	(0,100, 0,149)	0,049	(0,100, 0,148)	0,048
$\eta$	(0,012, 0,646)	0,634	(0,010, 0,562)	0,552
$\theta$	(0,911, 0,998)	0,087	(0,922, 0,998)	0,076
$\gamma$	(0,026, 0,976)	0,95	(0,025, 0,974)	0,949
$\delta$	(0,858, 0,955)	0,097	(0,858, 0,946)	0,088
$\eta^*$	(0,354, 0,988)	0,634	(0,438, 0,99)	0,552
$\theta^*$	(0,002, 0,089)	0,087	(0,001, 0,077)	0,076

ICr: Intervalo de credibilidade.

**Tabela 5.16.** Intervalos de credibilidade e suas amplitudes para o critério de Must e Cole considerados separadamente, para o sexo masculino.

	Must		Cole	
	ICr	amplitude	ICr	amplitude
$\pi$	(0,100, 0,149)	0,049	(0,101, 0,149)	0,048
$\eta$	(0,023, 0,964)	0,941	(0,018, 0,886)	0,868
$\theta$	(0,829, 0,993)	0,164	(0,871, 0,997)	0,126
$\gamma$	(0,022, 0,948)	0,926	(0,024, 0,969)	0,945
$\delta$	(0,851, 0,995)	0,144	(0,854, 0,985)	0,131
$\eta^*$	(0,036, 0,977)	0,941	(0,115, 0,983)	0,868
$\theta^*$	(0,007, 0,171)	0,164	(0,003, 0,129)	0,126

ICr: Intervalo de credibilidade.

2. Combinando os dois critérios, a prevalência de obesidade encontrada pelos critérios de Must e Cole, para o sexo feminino foi de 11,0%; sensibilidades de 86,9% e 76,4%, especificidades de 99,2% e 99,5%, valores preditivos positivos de 93,1% e 94,8%, valores preditivos negativos de 98,4% e 97,2%, taxas de erros dos resultados falsos negativos de 13,1% e 23,6%, e taxas de erros dos resultados falsos positivos de 0,8% e 0,5%, respectivamente. As amplitudes dos intervalos para a prevalência de obesidade, segundo os dois critérios foi de 4,2%; para as sensibilidades de 42,2% e 52,3%, especificidades de 2,7% e 1,8%, valores preditivos positivos de 21,7% e 18,3%, valores preditivos negativos de 5,2% e 6,3%, taxas de erros dos resultados falsos negativos de 42,2% e 52,3%, e taxas de erros dos resultados falsos positivos de 2,7% e 1,8%, respectivamente.

Para o sexo masculino a prevalência de obesidade foi de 12,2%; sensibilidades de 95,2% e 79,9%, especificidades de 97,6% e 99,5%, valores preditivos positivos de 85,2% e 95,7%, valores preditivos negativos de 99,3% e 97,3%, taxas de erros dos resultados falsos negativos de 4,8% e 20,1%, e taxas de erros dos resultados falsos positivos de 2,4% e 0,5%, respectivamente. As amplitudes dos intervalos para a prevalência de obesidade, segundo os dois critérios foi de 4,3%; para as sensibilidades de 16,9% e 40,8%, especificidades de

6,1% e 1,9%, valores preditivos positivos de 32,4% e 15,5%, valores preditivos negativos de 2,3% e 5,7%, taxas de erros dos resultados falsos negativos de 16,9% e 40,8%, e taxas de erros dos resultados falsos positivos de 6,1% e 1,9%, respectivamente.

Nota-se que quando os critérios são combinados, o de Must estima sensibilidade, valor preditivo negativo e taxa de erro do resultado falso positivo de obesidade superior ao de Cole, mas com especificidade, valor preditivo positivo e taxa de erro do resultado falso negativo inferior, em ambos os sexos.

Interpretando as prevalências encontradas nesse estudo, combinando os critérios antropométricos, nota-se que não houve diferença significativa entre os sexos. Em relação as meninas e meninos, o critério de Must mostrou-se mais sensível (86,9% e 95,2%) para classificar obesidade do que o de Cole (76,4% e 79,9%), e menos específico (99,2% e 97,6%) para não classificar obesidade do que o de Cole (99,5% e 99,5%). Com isso, para ambos os sexos o valor preditivo negativo de Must é melhor do que de Cole, e o valor preditivo positivo em Cole é melhor do que de Must.

Com respeito aos intervalos de credibilidade, tem-se que os intervalos obtidos pelo critério de Cole para a especificidade, valor preditivo positivo, como para a taxa de erro do resultado falso positivo são mais precisos, em ambos os sexos. No entanto, o critério de Must produziu intervalos para a sensibilidade, valor preditivo negativo e taxa de erro do resultado falso negativo mais precisos do que de Cole, como evidenciam as Tabelas 5.17 e 5.18, respectivamente.

**Tabela 5.17.** Intervalos de credibilidade e suas amplitudes para o critério de Must e Cole considerados separadamente, assim como para a combinação de ambos, para o sexo

feminino.						
	Must		Cole		Combinação dos critérios	
	ICr	amplitude	ICr	amplitude	ICr	amplitude
$\pi$	(0,100, 0,149)	0,049	(0,100, 0,148)	0,048	(0,089, 0,131)	0,042
$\eta_1$	(0,012, 0,646)	0,634			(0,574, 0,996)	0,422
$\theta_1$	(0,911, 0,998)	0,087			(0,972, 0,999)	0,027
$\gamma_1$	(0,026, 0,976)	0,95			(0,781, 0,998)	0,217
$\delta_1$	(0,858, 0,955)	0,097			(0,947, 0,999)	0,052
$\eta_1^*$	(0,354, 0,988)	0,634			(0,004, 0,426)	0,422
$\theta_1^*$	(0,002, 0,089)	0,087			(0,000, 0,027)	0,027
$\eta_2$			(0,010, 0,562)	0,552	(0,455, 0,978)	0,523
$\theta_2$			(0,922, 0,998)	0,076	(0,981, 0,999)	0,018
$\gamma_2$			(0,025, 0,974)	0,949	(0,815, 0,998)	0,183
$\delta_2$			(0,858, 0,946)	0,088	(0,934, 0,997)	0,063
$\eta_2^*$			(0,438, 0,99)	0,552	(0,022, 0,545)	0,523
$\theta_2^*$			(0,001, 0,077)	0,076	(0,000, 0,018)	0,018

ICr: Intervalo de credibilidade.

Baseando-se na Tabela 5.17 tem-se que, comparando a prevalência de obesidade do critério de Must e Cole separadamente, com a prevalência de obesidade combinando ambos os critérios antropométricos nota-se uma redução de 7% e 6% nas amplitudes do critério de Must e Cole para o sexo feminino, quando comparadas com a amplitude da prevalência combinando os dois critérios. Com redução nas amplitudes de 21,2%, 6%, 73,3%, 4,5%, 21,2% e 6% para a sensibilidade, especificidade, valores preditivos positivo e negativo, taxa de erro de um resultado falso negativo e falso positivo do critério de Must, respectivamente, quando comparadas com as obtidas para a combinação de ambos os critérios. Para o critério de Cole a redução nas amplitudes foi de 2,9%, 5,8%, 76,6%, 2,5%, 2,9% e 5,8%

para a sensibilidade, especificidade, valores preditivos positivo e negativo, taxa de erro de um resultado falso negativo e falso positivo, respectivamente, comparando com as obtidas para a combinação de ambos os critérios.

**Tabela 5.18.** Intervalos de credibilidade e suas amplitudes para o critério de Must e Cole considerados separadamente, assim como para a combinação de ambos, para o sexo masculino.

	Must		Cole		Combinação dos critérios	
	ICr	amplitude	ICr	amplitude	ICr	amplitude
$\pi$	(0,100, 0,149)	0,049	(0,101, 0,149)	0,048	(0,101, 0,144)	0,043
$\eta_1$	(0,023, 0,964)	0,941			(0,829, 0,998)	0,169
$\theta_1$	(0,829, 0,993)	0,164			(0,938, 0,999)	0,061
$\gamma_1$	(0,022, 0,948)	0,926			(0,667, 0,991)	0,324
$\delta_1$	(0,851, 0,995)	0,144			(0,976, 0,999)	0,023
$\eta_1^*$	(0,036, 0,977)	0,941			(0,001, 0,170)	0,169
$\theta_1^*$	(0,007, 0,171)	0,164			(0,001, 0,062)	0,061
$\eta_2$			(0,018, 0,886)	0,868	(0,576, 0,984)	0,408
$\theta_2$			(0,871, 0,997)	0,126	(0,980, 0,999)	0,019
$\gamma_2$			(0,024, 0,969)	0,945	(0,844, 0,999)	0,155
$\delta_2$			(0,854, 0,985)	0,131	(0,941, 0,998)	0,057
$\eta_2^*$			(0,115, 0,983)	0,868	(0,014, 0,422)	0,408
$\theta_2^*$			(0,003, 0,129)	0,126	(0,000, 0,019)	0,019

ICr: Intervalo de credibilidade.

E baseando-se na Tabela 5.18 tem-se que, comparando a prevalência de obesidade do critério de Must e Cole separadamente, com a prevalência de obesidade combinando ambos os critérios antropométricos nota-se uma redução de 6% e 5% nas amplitudes do critério de Must e Cole para o sexo masculino, quando comparadas com a amplitude da prevalência combinando os dois critérios. Com redução nas amplitudes de 77,2%,

10,3%, 60,2%, 12,1%, 77,2% e 10,3% para a sensibilidade, especificidade, valores preditivos positivo e negativo, taxa de erro de um resultado falso negativo e falso positivo do critério de Must, respectivamente, quando comparadas com as obtidas para a combinação de ambos os critérios. Para o critério de Cole a redução nas amplitudes foi de 46%, 10,7%, 79,0%, 7,4%, 46% e 10,7% para a sensibilidade, especificidade, valores preditivos positivo e negativo, taxa de erro de um resultado falso negativo e falso positivo, respectivamente, comparando com as obtidas para a combinação de ambos os critérios.

Destes resultados pode-se concluir que inferências mais definidas para a população de escolares sobre a prevalência de obesidade, as sensibilidades, especificidades, valores preditivos positivos e negativos, assim como para as taxas de erros dos resultados falsos negativos e falsos positivos, foram obtidas da combinação de ambos os critérios comparadas àquelas de Must e Cole separadamente. Esta situação é devido as amplitudes dos intervalos de credibilidade serem menores combinando os critérios Must e Cole.

Então, nota-se que quando os métodos bayesianos propostos são utilizados, combinando os dois critérios, obtém-se melhores estimativas para os parâmetros de interesse, pois com a combinação de ambos os critérios tem-se mais informação.

3. A indicação de convergência para todos os parâmetros de interesse nesta aplicação foi verificada através da análise do comportamento das trajetórias das cadeias, e também a partir do diagnóstico de Gelman & Rubin.

4. O *Software WinBUGS*, utilizado para a implementação dos algoritmos na aplicação apresentada, mostrou-se uma importante ferramenta computacional, pois forneceu automaticamente os resumos decorrentes da amostra obtida como média, mediana, desvio padrão, intervalo de credibilidade e, também, gráficos para análise das trajetórias das cadeias geradas, densidades e diagnósticos de convergência de Gelman & Rubin de cada parâmetro.

# Capítulo 6

## Conclusões e Perspectivas Futuras

Neste trabalho foi aplicado o *Gibbs Sampling* com variáveis latentes para estimar os parâmetros tanto de um quanto de dois testes diagnósticos na ausência de um diagnóstico padrão ouro. Para o procedimento estudado as distribuições condicionais completas apresentaram formas conhecidas e a implementação usual do algoritmo não foi complicada. A indicação de convergência para os parâmetros envolvidos no teste diagnóstico foi verificada através da análise do comportamento das trajetórias das cadeias, e também através do diagnóstico de Gelman & Rubin.

O *Software WinBUGS*, utilizado para a implementação dos algoritmos nas aplicações apresentadas, mostrou ser uma ferramenta simples e muito eficiente, pois fornece automaticamente os resumos decorrentes da amostra obtida como média, mediana, desvio padrão, intervalo de credibilidade e, também, gráficos para análise das trajetórias das cadeias geradas, densidades e diagnósticos de convergência de Gelman & Rubin de cada parâmetro.

Os resultados obtidos a partir da implementação dos algoritmos *Gibbs Sampling* com variáveis latentes, tanto para a aplicação sobre a infecção causada pelo parasita *strongyloides* em refugiados do Camboja quanto para a aplicação sobre obesidade em escolares do sexo feminino e masculino, foram melhores quando os dois testes diagnósticos, teste

sorológico e exame de fezes, e os dois critérios antropométricos, Must e Cole, foram combinados, comparadas àqueles do teste sorológico, exame de fezes, critério de Must e Cole separadamente.

Dessa forma, tem-se que os métodos bayesianos apresentados neste trabalho mostraram-se úteis na estimação de melhores inferências possíveis de testes diagnósticos na ausência de um padrão ouro.

Em relação as perspectivas futuras, pretende-se desenvolver uma medida de evidência baseada na metodologia proposta por Pereira & Stern (1999), Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses, *Entropy*, **1**, 99-110, para testar a hipótese nula  $H_0 : \pi = 1$  (hipótese do indivíduo possuir a doença) contra a hipótese alternativa  $H_1 : \pi = 0$  (hipótese do indivíduo não possuir a doença). Também, deseja-se encontrar um ponto de corte, através de uma função de perda, para decidir se a hipótese nula será aceita ou não, baseando-se no procedimento de Madruga *et al.* (2000), *On the Bayesianity of Pereira-Stern Tests*, Relatório Técnico do Departamento de Estatística, IME-USP, **10**, 1-9, (RT-MAE-2000-10).

# Apêndice

## A. Transformação de Variável

**Teorema:** Seja,

(a)  $g(\cdot)$ : função diferenciável e estritamente crescente ou decrescente no intervalo  $I \subset \mathfrak{R}$ .

(b)  $g(I)$ : campo de variação da função  $g(\cdot)$ .

(c)  $g^{-1}(\cdot)$ : função inversa de  $g$ .

(4) Seja  $X$  uma variável aleatória contínua com função de densidade de probabilidade  $f_X(x)$ , onde  $f_X(x) = 0$  para  $x \notin I$ .

(5)  $Y = g(X)$  : variável aleatória contínua.

Então, a variável aleatória  $Y$  tem uma função de densidade de probabilidade  $g_Y(y)$  definida da seguinte forma:

$$g(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| & \text{se } y \in g(I) \\ 0 & \text{se } y \notin g(I) \end{cases}$$

Considere uma *priori*  $\eta \sim \text{Beta}(\alpha_1, \beta_1)$  para a sensibilidade  $\eta = 1 - \eta^*$ , e a seguinte função de densidade de probabilidade dada por:

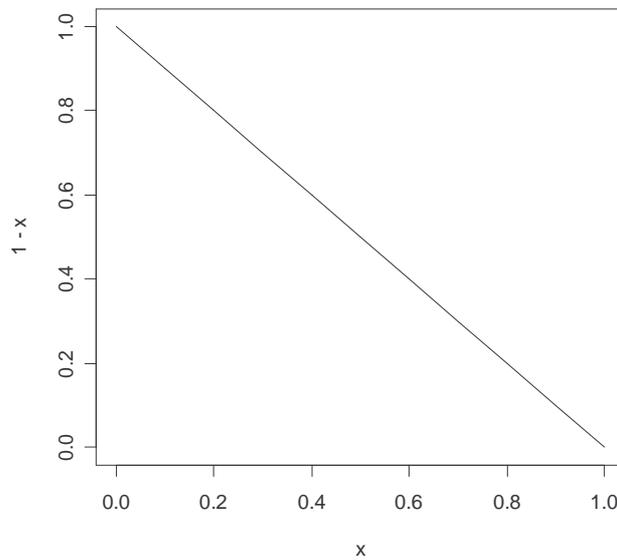
$$f_X(x) = \begin{cases} \frac{(x)^{\alpha_1-1}(1-x)^{\beta_1-1}}{\text{Beta}(\alpha_1, \beta_1)} & \text{se } 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Seja  $g(x) = 1 - x$ , o objetivo é determinar a função de densidade de probabilidade de  $g_Y(y)$  da variável aleatória  $Y = g(X) = 1 - X$ .

*Resolução:* Atribuindo valores para  $x$ , conseqüentemente os valores de  $g(x)$  serão obtidos. Considere a seguinte tabela de valores:

$x$	0	0.1	...	0.9	1
$g(x)$	1	0.9	...	0.1	0

A partir desta tabela, é possível notar que  $g(X)$  é uma função estritamente decrescente. Isto fica mais claro visualizando o gráfico abaixo:



Representação Gráfica de  $g(X)$ .

$I(0, 1)$  : domínio da função de densidade de probabilidade  $\Rightarrow g(I) = (0, 1)$ , pois quando  $x = 0 \Rightarrow y = 1$  e quando  $x = 1 \Rightarrow y = 0$ .

$$y = g(x) = 1 - x \Rightarrow x = g^{-1}(y) = 1 - y$$

$$\frac{dg^{-1}(y)}{dy} = -1$$

Como a função  $g(X)$  é diferenciável e estritamente decrescente em  $I(0, 1)$ , pode-se aplicar o teorema. Onde,

$$\begin{aligned}
f_X(g^{-1}(y)) &= \frac{(1-y)^{\alpha_1-1}(y)^{\beta_1-1}}{\text{Beta}(\alpha_1, \beta_1)} = \frac{\Gamma(\alpha_1 + \beta_1)}{\Gamma(\alpha_1)\Gamma(\beta_1)}(1-y)^{\alpha_1-1}(y)^{\beta_1-1} \\
&= \frac{\Gamma(\alpha_1 + \beta_1)}{\Gamma(\alpha_1)\Gamma(\beta_1)}y^{\beta_1-1}(1-y)^{\alpha_1-1}
\end{aligned}$$

Logo,

$$g_Y(y) = \begin{cases} \frac{\Gamma(\alpha_1+\beta_1)}{\Gamma(\alpha_1)\Gamma(\beta_1)}y^{\beta_1-1}(1-y)^{\alpha_1-1} & \text{para } y \in [0, 1] \\ 0 & \text{para } y \notin [0, 1] \end{cases}$$

$$\begin{cases} \frac{\Gamma(\beta_1+\alpha_1)}{\Gamma(\beta_1)\Gamma(\alpha_1)}y^{\beta_1-1}(1-y)^{\alpha_1-1} & \text{para } y \in [0, 1] \\ 0 & \text{para } y \notin [0, 1] \end{cases}$$

é a função densidade de probabilidade de  $Y \sim \text{Beta}(\beta_1, \alpha_1)$ .

Portanto, pode-se dizer que  $\eta = (1 - \eta^*) \sim \text{Beta}(\alpha_1, \beta_1)$  é equivalente a  $\eta^* = (1 - \eta) \sim \text{Beta}(\beta_1, \alpha_1)$ , e vice-versa.

## B. Gibbs Sampling: Definições e Propriedades

O *Gibbs Sampling* foi desenvolvido por Geman & Geman (1984) no contexto de restauração de imagens. Este método consiste em um esquema markoviano de atualização, que permite a obtenção de amostras de uma distribuição conjunta através de amostragens iterativas das distribuições condicionais completas. Foi popularizado entre os bayesianos por Gelfand & Smith (1990) e a partir de então muitas aplicações têm sido apresentadas. O grande interesse por este método é devido à sua simplicidade conceitual e facilidade de implementação.

Para descrever o algoritmo de *Gibbs*, considere a densidade *a posteriori* conjunta  $p(\boldsymbol{\omega}|t)$ , onde  $t$  representa os dados e  $\boldsymbol{\omega}$  o vetor de parâmetros.

As distribuições de cada componente individual de  $\omega$ , podem ser especificadas, no mínimo proporcional a um determinado núcleo. Denotando estas densidades condicionais completas por

$$p(\omega_i|t, \omega_1, \omega_2, \dots, \omega_k), \quad i = 1, \dots, k$$

e o valor gerado para o  $i$ -ésimo parâmetro na  $j$ -ésima iteração por  $\omega_i^{(j)}$ , o algoritmo procede da seguinte forma:

i) escolha valores iniciais

$$\omega_2^{(0)}, \omega_3^{(0)}, \dots, \omega_{k-1}^{(0)}, \omega_k^{(0)};$$

ii) para  $j = 1$ , gere um valor  $\omega_1^{(j)}$ , da distribuição condicional

$$p(\omega_1|t, \omega_2^{(j-1)}, \omega_3^{(j-1)}, \dots, \omega_k^{(j-1)});$$

gere um valor  $\omega_2^{(j)}$ , da distribuição condicional

$$p(\omega_2|t, \omega_1^{(j)}, \omega_3^{(j-1)}, \dots, \omega_k^{(j-1)});$$

.....

gere um valor  $\omega_k^{(j)}$ , da distribuição condicional

$$p(\omega_k|t, \omega_1^{(j)}, \omega_2^{(j)}, \dots, \omega_{k-1}^{(j)});$$

iii) repita ii) para  $j = 2, 3, \dots, m$ .

Sob condições gerais de regularidade quando  $m \rightarrow \infty$ ,  $(\omega_1^{(m)}, \dots, \omega_k^{(m)}) \rightarrow p(\omega_1, \dots, \omega_k)$  e então  $\omega_i^{(m)} \xrightarrow{D} p(\omega_i)$ , ou seja, a distribuição da cadeia gerada pelo *Gibbs Sampling*, na iteração  $m$ , converge para distribuição de equilíbrio, na norma da variação total, a uma taxa geométrica em  $m$ . Essa propriedade é também conhecida como ergodicidade e uma consequência importante deste resultado é que as médias ergódicas  $\bar{f}_m = \frac{1}{m} \sum_{j=1}^m f(\omega^{(j)})$

convergem quase certamente para  $E[f(\omega)]$ , quando  $m \rightarrow \infty$ , se a esperança calculada sob  $p$ , existir.

Apesar deste resultado ser muito útil, em alguns casos, o algoritmo *Gibbs Sampling* pode apresentar um ritmo de convergência extremamente lento.

### C. Diagnósticos de Convergência

Na implementação dos Métodos de Monte Carlo aplicados a Cadeia de Markov questões importantes tais como a escolha do amostrador, o número de replicações independentes, a escolha de valores iniciais e problemas referentes a estimação e eficiência devem ser considerados.

Na ausência de técnicas gerais para se determinar o comprimento da cadeia, *a priori*, análises estatísticas devem ser realizadas, *a posteriori*, para assegurar a convergência da cadeia. Estes procedimentos são chamados diagnósticos de convergência.

O *Gibbs Sampling* gera cadeias de Markov, de variáveis aleatórias, que convergem em distribuição, para a distribuição de interesse. Diferentes abordagens para se extrair informações das seqüências geradas exploram esta propriedade para determinar o comprimento da fase inicial transiente.

Existem duas formas de monitorar convergência:

- Monitoração informal;
- Monitoração formal.

A monitoração informal da convergência é feita através da análise do comportamento da trajetória de uma cadeia ao longo das iterações, da monitoração dos gráficos das densidades *a posteriori*, estimadas ao longo das iterações, de gráficos das médias ergódicas ao invés de valores gerados.

A monitoração formal da convergência é baseada em propriedades estatísticas das cadeias simuladas. Brooks & Roberts (1995) discutiram vários métodos de diagnósticos

de convergência, apresentados na literatura, com particular ênfase sobre implementação. Os métodos são comparados em termos de sua interpretabilidade e aplicabilidade com recomendações para particulares classes de problemas. Dentre eles, estão os métodos de Raftery & Lewis (1992), na categoria de métodos espectrais, os diagnósticos de Geweke (1992) e de Heidelberger & Welch (1983) e o método de Gelman & Rubin (1992). Sendo este último o de interesse neste trabalho, com discussão no Apêndice D.

#### D. Diagnóstico de Gelman & Rubin (1992)

O método de Gelman & Rubin (1992), para testar convergência, é baseado em 2 ou mais cadeias paralelas, inicializadas em diferentes pontos dispersos em relação a verdadeira distribuição *a posteriori*.

Este método é baseado em uma comparação, para cada uma das variáveis, entre a variância amostral dentro das cadeias e entre as cadeias. Esta comparação é utilizada para se estimar o fator para o qual o parâmetro de escala da distribuição marginal *a posteriori*, pode ser reduzido a medida que o tamanho da amostra cresça.

O método de Geman & Rubin (1992) consiste, essencialmente, em uma análise clássica da variância. Melhores resultados são obtidos para parâmetros cujas densidades são aproximadamente normais.

#### E. Software WinBUGS

Neste apêndice, será introduzido o *Software WinBUGS* devido a sua utilização no desenvolvimento das aplicações deste trabalho.

O *Software WinBUGS* é uma versão interativa do *Windows* para o programa BUGS (Bayesian Inference Using Gibbs Sample) (Spiegelhalter, Thomas, Best e Gilks, 1995), programa este direcionado à aplicação da Inferência Bayesiana, em problemas estatísticos, usando o *Gibbs Sampling*. O BUGS é um sistema desenvolvido na Unidade de

Bioestatística do *Medical Research Council* da Inglaterra. Consiste de um conjunto de funções que permite a especificação do modelo e das distribuições de probabilidade para todos os seus componentes aleatórios (observações e parâmetros).

Para cada conjunto de dados e modelo utilizado, o *Software WinBUGS* fornece os valores amostrados de cada parâmetro, monitorado a cada  $K$  iterações, a partir de uma determinada iteração  $m$ . Ambos os valores de  $K$  e  $m$  bem como os parâmetros a serem monitorados são especificados pelo usuário. Além disso, fornece automaticamente resumos decorrentes da amostra obtida como média, mediana, desvio padrão, intervalo de credibilidade e, também, apresenta gráficos para análise das trajetórias das cadeias geradas, densidades e autocorrelações. Além disso, fornece resultados para o diagnóstico de convergência de Gelman & Rubin (1992) que consiste basicamente em uma análise de variância intra e entre as cadeias geradas. Para este diagnóstico o *Software WinBUGS* fornece gráficos onde a largura do intervalo de 80% para o parâmetro de interesse, considerando as cadeias combinadas, será representada em verde, a largura média dos intervalos de 80% dentro de cada cadeia individual será representada em azul e a razão  $R = combinada/individual$  (fator de redução na escala) será representada em vermelho.

O *Software WinBUGS* é dotado da capacidade de reconhecer formas conjugadas, distribuições discretas ou logconcaidade. Com base nesta classificação o algoritmo de amostragem mais eficiente é selecionado para a simulação. Quando formas conjugadas condicionais são reconhecidas a amostra é eficientemente gerada das distribuições condicionais completas. A linguagem do sistema e a entrada e saída de dados obedecem a mesma sintaxe da linguagem de programação S-PLUS.

F. Programas para o cálculo das estimativas via *Gibbs Sampling* do Teste Sorológico, Exame de Fezes e Combinação dos dois Testes.

*Software* utilizado: *WinBUGS* versão 1.4

**a) Teste Sorológico**

Modelo

```
{  
p1<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
p2<-(pi*(1-eta))/(pi*(1-eta)+(1-pi)*teta)  
for (i in 1:a) {ybern1[i]~dbern(p1)}  
for (i in 1:b) {ybern2[i]~dbern(p2)}  
y1<-sum(ybern1[])  
y2<-sum(ybern2[])  
k1<-y1+y2+alpha  
k2<-a+b-y1-y2+beta  
k3<-y1+alpha1  
k4<-y2+beta1  
k5<-b-y2+alpha2  
k6<-a-y1+beta2  
pi~dbeta(k1,k2)  
eta~dbeta(k3,k4)  
teta~dbeta(k5,k6)  
eta.estrela~dbeta(k4,k3)  
teta.estrela~dbeta(k6,k5)  
gamma<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
delta<-((1-pi)*teta)/(pi*(1-eta)+(1-pi)*teta)  
}
```

Dados

```
list(a=125,b=37,alpha=1,beta=1,alpha1=21.96,beta1=5.49,alpha2=4.93,beta2=2.38)
```

Valore Iniciais

```
list(pi=0.5,eta=0.8,teta=0.675)
```

## b) Exame de Fezes

Modelo

```
{  
p1<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
p2<-(pi*(1-eta))/(pi*(1-eta)+(1-pi)*teta)  
for (i in 1:a) {ybern1[i]~dbern(p1)}  
for (i in 1:b) {ybern2[i]~dbern(p2)}  
y1<-sum(ybern1[])  
y2<-sum(ybern2[])  
k1<-y1+y2+alpha  
k2<-a+b-y1-y2+betak3<-y1+alpha1  
k4<-y2+beta1  
k5<-b-y2+alpha2  
k6<-a-y1+beta2  
pi~dbeta(k1,k2)  
eta~dbeta(k3,k4)  
teta~dbeta(k5,k6)  
eta.estrela~dbeta(k4,k3)  
teta.estrela~dbeta(k6,k5)  
gamma<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
delta<-((1-pi)*teta)/(pi*(1-eta)+(1-pi)*teta)  
}
```

Dados

```
list(a=40,b=122,alpha=1,beta=1,alpha1=4.44,beta1=13.31,alpha2=71.25,beta2=3.75)
```

Valore Iniciais

```
list(pi=0.5,eta=0.25,teta=0.95)
```

### c) Combinação do teste sorológico e exame de fezes

Modelo

```
{
```

```
p1<-(pi*eta1*eta2)/(pi*eta1*eta2+(1-pi)*(1-teta1)*(1-teta2))
```

```
p2<-(pi*eta1*(1-eta2))/(pi*eta1*(1-eta2)+(1-pi)*(1-teta1)*teta2)
```

```
p3<-(pi*(1-eta1)*eta2)/(pi*(1-eta1)*eta2+(1-pi)*teta1*(1-teta2))
```

```
p4<-(pi*(1-eta1)*(1-eta2))/( pi*(1-eta1)*(1-eta2)+(1-pi)*teta1*teta2)
```

```
for (i in 1:u) {ybern1[i]~dbern(p1)}
```

```
for (i in 1:v) {ybern2[i]~dbern(p2)}
```

```
for (i in 1:w) {ybern3[i]~dbern(p3)}
```

```
for (i in 1:x) {ybern4[i]~dbern(p4)}
```

```
y1<-sum(ybern1[])
```

```
y2<-sum(ybern2[])
```

```
y3<-sum(ybern3[])
```

```
y4<-sum(ybern4[])
```

```
k1<-y1+y2+y3+y4+alpha
```

```
k2<-u+v+w+x-y1-y2-y3-y4+beta
```

```
k3<-y1+y2+alpha1
```

```
k4<-y3+y4+beta1
```

```
k5<-w+x-y3-y4+alpha3
```

```
k6<-u+v-y1-y2+beta3
```

```
k7<-y1+y3+alpha2
```

```
k8<-y2+y4+beta2
```

```

k9<-v+x-y2-y4+alpha4
k10<-u+w-y1-y3+beta4
pi~dbeta(k1,k2)
eta1~dbeta(k3,k4)
teta1~dbeta(k5,k6)
eta2~dbeta(k7,k8)
teta2~dbeta(k9,k10)
eta1.estrela~dbeta(k4,k3)
teta1.estrela~dbeta(k6,k5)
eta2.estrela~dbeta(k8,k7)
teta2.estrela~dbeta(k10,k9)
gamma1<-(pi*eta1)/(pi*eta1+(1-pi)*(1-teta1))
delta1<-((1-pi)*teta1)/(pi*(1-eta1)+(1-pi)*teta1)
gamma2<-(pi*eta2)/(pi*eta2+(1-pi)*(1-teta2))
delta2<-((1-pi)*teta2)/(pi*(1-eta2)+(1-pi)*teta2)
}
Dados
list(u=38,v=87,w=2,x=35,alpha=1,beta=1,alpha1=21.96,beta1=5.49,alpha3=4.93,
beta3=2.38,alpha2=4.44,beta2=13.31,alpha4=71.25,beta4=3.75)
Valore Iniciais
list(pi=0.5,eta1=0.80,teta1=0.675,eta2=0.25,teta2=0.95)

```

G. Programas para o cálculo das estimativas via *Gibbs Sampling* do Critério de Must, Cole e Combinação dos dois Critérios.

*Software* utilizado: *WinBUGS* versão 1.4

#### a) Critério de Must para o Sexo Feminino

Modelo

```
{  
p1<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
p2<-(pi*(1-eta))/(pi*(1-eta)+(1-pi)*teta)  
for (i in 1:a) {ybern1[i]~dbern(p1)}  
for (i in 1:b) {ybern2[i]~dbern(p2)}  
y1<-sum(ybern1[])  
y2<-sum(ybern2[])  
k1<-y1+y2+alpha  
k2<-a+b-y1-y2+beta  
k3<-y1+alpha1  
k4<-y2+beta1  
k5<-b-y2+alpha2  
k6<-a-y1+beta2  
pi~dbeta(k1,k2)  
eta~dbeta(k3,k4)  
teta~dbeta(k5,k6)  
eta.estrela~dbeta(k4,k3)  
teta.estrela~dbeta(k6,k5)  
gamma<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
delta<-((1-pi)*teta)/(pi*(1-eta)+(1-pi)*teta)  
}
```

Dados

```
list(a=11,b=188,alpha=87.375,beta=611.625,alpha1=1,beta1=1,alpha2=1,beta2=1)
```

Valores Iniciais

```
list(pi=0.125,eta=0.5,teta=0.5)
```

## b) Critério de Cole para o sexo Feminino

Modelo

```
{  
p1<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
p2<-(pi*(1-eta))/(pi*(1-eta)+(1-pi)*teta)  
for (i in 1:a) {ybern1[i]~dbern(p1)}  
for (i in 1:b) {ybern2[i]~dbern(p2)}  
y1<-sum(ybern1[])  
y2<-sum(ybern2[])  
k1<-y1+y2+alpha  
k2<-a+b-y1-y2+beta  
k3<-y1+alpha1  
k4<-y2+beta1  
k5<-b-y2+alpha2  
k6<-a-y1+beta2  
pi~dbeta(k1,k2)  
eta~dbeta(k3,k4)  
teta~dbeta(k5,k6)  
eta~dbeta(k4,k3)  
teta~dbeta(k6,k5)  
gamma<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
delta<-((1-pi)*teta)/(pi*(1-eta)+(1-pi)*teta)  
}
```

Dados

```
list(a=9,b=190,alpha=87.375,beta=611.625,alpha1=1,beta1=1,alpha2=1,beta2=1)
```

Valores Iniciais

```
list(pi=0.125,eta=0.5,teta=0.5)
```

**c) Critério de Must para o Sexo Masculino**

Modelo

```
{  
p1<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
p2<-(pi*(1-eta))/(pi*(1-eta)+(1-pi)*teta)  
for (i in 1:a) {ybern1[i]~dbern(p1)}  
for (i in 1:b) {ybern2[i]~dbern(p2)}  
y1<-sum(ybern1[])  
y2<-sum(ybern2[])  
k1<-y1+y2+alpha  
k2<-a+b-y1-y2+beta  
k3<-y1+alpha1  
k4<-y2+beta1  
k5<-b-y2+alpha2  
k6<-a-y1+beta2  
pi~dbeta(k1,k2)  
eta~dbeta(k3,k4)  
teta~dbeta(k5,k6)  
eta~dbeta(k4,k3)  
teta~dbeta(k6,k5)  
gamma<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
delta<-((1-pi)*teta)/(pi*(1-eta)+(1-pi)*teta)  
}
```

Dados

```
list(a=28,b=191,alpha=87.375,beta=611.625,alpha1=1,beta1=1,alpha2=1,beta2=1)
```

Valores Iniciais

```
list(pi=0.125,eta=0.5,teta=0.5)
```

#### d) Critério de Cole para o Sexo Masculino

Modelo

```
{  
p1<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
p2<-(pi*(1-eta))/(pi*(1-eta)+(1-pi)*teta)  
for (i in 1:a) {ybern1[i]~dbern(p1)}  
for (i in 1:b) {ybern2[i]~dbern(p2)}  
y1<-sum(ybern1[])  
y2<-sum(ybern2[])  
k1<-y1+y2+alpha  
k2<-a+b-y1-y2+beta  
k3<-y1+alpha1  
k4<-y2+beta1  
k5<-b-y2+alpha2  
k6<-a-y1+beta2  
pi~dbeta(k1,k2)  
eta~dbeta(k3,k4)  
teta~dbeta(k5,k6)  
eta~dbeta(k4,k3)  
teta~dbeta(k6,k5)  
gamma<-(pi*eta)/(pi*eta+(1-pi)*(1-teta))  
delta<-((1-pi)*teta)/(pi*(1-eta)+(1-pi)*teta)  
}
```

Dados

```
list(a=20,b=199,alpha=87.375,beta=611.625,alpha1=1,beta1=1,alpha2=1,beta2=1)
```

Valores Iniciais

```
list(pi=0.125,eta=0.5,teta=0.5)
```

**e) Combinação do Critério de Must e Cole para o Sexo Feminino**

Modelo

{

$p1 <- (\pi * \eta_1 * \eta_2) / (\pi * \eta_1 * \eta_2 + (1 - \pi) * (1 - \tau_1) * (1 - \tau_2))$

$p2 <- (\pi * \eta_1 * (1 - \eta_2)) / (\pi * \eta_1 * (1 - \eta_2) + (1 - \pi) * (1 - \tau_1) * \tau_2)$

$p4 <- (\pi * (1 - \eta_1) * (1 - \eta_2)) / (\pi * (1 - \eta_1) * (1 - \eta_2) + (1 - \pi) * \tau_1 * \tau_2)$

for (i in 1:u) {ybern1[i] ~ dbern(p1)}

for (i in 1:v) {ybern2[i] ~ dbern(p2)}

for (i in 1:x) {ybern4[i] ~ dbern(p4)}

y1 <- sum(ybern1[])

y2 <- sum(ybern2[])

y4 <- sum(ybern4[])

k1 <- -y1 + y2 + y4 + alpha

k2 <- -u + v + x - y1 - y2 - y4 + beta

k3 <- -y1 + y2 + alpha1

k4 <- -y4 + beta1

k5 <- -x - y4 + alpha3

k6 <- -u + v - y1 - y2 + beta3

k7 <- -y1 + alpha2

k8 <- -y2 + y4 + beta2

k9 <- -v + x - y2 - y4 + alpha4

k10 <- -u - y1 + beta4

$\pi \sim \text{dbeta}(k1, k2)$

$\eta_1 \sim \text{dbeta}(k3, k4)$

$\tau_1 \sim \text{dbeta}(k5, k6)$

$\eta_2 \sim \text{dbeta}(k7, k8)$

$\tau_2 \sim \text{dbeta}(k9, k10)$

$\eta_1.\text{estrela} \sim \text{dbeta}(k4, k3)$

$\tau_1.\text{estrela} \sim \text{dbeta}(k6, k5)$

```

eta2.estrela ~ dbeta(k8,k7)
teta2.estrela ~ dbeta(k10,k9)
gamma1 <- (pi*eta1)/(pi*eta1+(1-pi)*(1-teta1))
delta1 <- ((1-pi)*teta1)/(pi*(1-eta1)+(1-pi)*teta1)
gamma2 <- (pi*eta2)/(pi*eta2+(1-pi)*(1-teta2))
delta2 <- ((1-pi)*teta2)/(pi*(1-eta2)+(1-pi)*teta2)
}

```

Dados

```

list(u=9,v=2,x=188,alpha=87.375,beta=611.625,alpha1=1,beta1=1,alpha3=1,beta3=1,
alpha2=1,beta2=1,alpha4=1,beta4=1)

```

Valores Iniciais

```

list(pi=0.125,eta1=0.5,teta1=0.5,eta2=0.5,teta2=0.5)

```

## f) Combinação do Critério de Must e Cole para o Sexo Masculino

Modelo

```

{
p1 <- (pi*eta1*eta2)/(pi*eta1*eta2+(1-pi)*(1-teta1)*(1-teta2))
p2 <- (pi*eta1*(1-eta2))/(pi*eta1*(1-eta2)+(1-pi)*(1-teta1)*teta2)
p4 <- (pi*(1-eta1)*(1-eta2))/( pi*(1-eta1)*(1-eta2)+(1-pi)*teta1*teta2)
for (i in 1:u) {ybern1[i] ~ dbern(p1)}
for (i in 1:v) {ybern2[i] ~ dbern(p2)}
for (i in 1:x) {ybern4[i] ~ dbern(p4)}
y1 <- sum(ybern1[])
y2 <- sum(ybern2[])
y4 <- sum(ybern4[])
k1 <- y1+y2+y4+alpha
k2 <- u+v+x-y1-y2-y4+beta
k3 <- y1+y2+alpha1

```

```

k4<-y4+beta1
k5<-x-y4+alpha3
k6<-u+v-y1-y2+beta3
k7<-y1+alpha2
k8<-y2+y4+beta2
k9<-v+x-y2-y4+alpha4
k10<-u-y1+beta4
pi~dbeta(k1,k2)
eta1~dbeta(k3,k4)
teta1~dbeta(k5,k6)
eta2~dbeta(k7,k8)
teta2~dbeta(k9,k10)
eta1.estrela~dbeta(k4,k3)
teta1.estrela~dbeta(k6,k5)
eta2.estrela~dbeta(k8,k7)
teta2.estrela~dbeta(k10,k9)
gamma1<-(pi*eta1)/(pi*eta1+(1-pi)*(1-teta1))
delta1<-((1-pi)*teta1)/(pi*(1-eta1)+(1-pi)*teta1)
gamma2<-(pi*eta2)/(pi*eta2+(1-pi)*(1-teta2))
delta2<-((1-pi)*teta2)/(pi*(1-eta2)+(1-pi)*teta2)
}
Dados
list(u=20,v=8,x=191,alpha=87.375,beta=611.625,alpha1=1,beta1=1,alpha3=1,beta3=1,
alpha2=1,beta2=1,alpha4=1,beta4=1)
Valores Iniciais
list(pi=0.125,eta1=0.5,teta1=0.5,eta2=0.5,teta2=0.5)

```

# Referências Bibliográficas

- [1] BLAIR, S. N.; HORTON, E.; LEON, A. S.; LEE, I-MIN; DRINKWATER, B. L.; DISHMAN, R. K.; MANEY, M.; KIENHOLZ, M. L.; (1996). Physical activity, nutrition, and chronic disease. *Medicine Science Sports Exercise*, **28**, n. 3, 335-349.
- [2] BROOKS, S.; ROBERTS, G.; (1995). Diagnosing Convergence of Markov Chain Monte Carlo Algorithms. *Thechnical Report, Dept. of Pure Maths and Mathematical Statistics*. University of Cambridge.
- [3] COLE, T. J.; BELLIZZI, M. C.; FLEGAL, K. M.; DIETZ, W. H.; (2000). Establishing a standard definition for child overweight and obesity worldwide: international survey. *British Medical Journal*, **320**, 1-6.
- [4] DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B.; (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal Royal Statistical Society, B.*, **39**, 1-38.
- [5] DIEBOLT, J.; ROBERT, C.P.; (1994). Estimation of Finite Mixture Distributions through Bayesian Sampling. *Journal Royal Statistical Society, B.*, **56**, n. 2, 363-375.
- [6] EGGER, G.; BOLTON, A.; NEILL, M. O.; FREEMAND; (1996). Effectiveness of an abnominal obesity reduction programme in mem: the GutBuster 'waist loss' programme. *International Journal of Obesity*, **20**, 227-231.

- [7] FLETCHER, R.; FLETCHER, S.; WAGNER, E.; (1996). *Epidemiologia clínica: Elementos essenciais*. Trad. Bruce B. Duncan, Maria Inês Schmidt. - 3 ed. - Porto Alegre: Artes Médicas.
- [8] GELFAND, A. E.; SMITH, A. F.M.; (1990). Sampling - based approaches to calculating marginal densities. *Journal of the American Association*, **85**, 398-409.
- [9] GELMAN, A. E.; RUBIN, D. R.; (1992). Inference from Iterative Simulation using Multiple Sequences. *Statistical Science*, **7**, 457-472.
- [10] GEMAN, S.; GEMAN, D.; (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721-740.
- [11] GEWEKE, J.; (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics*, **4**, ed. J.M.Bernardo, J.O.Berger, A.P.Dawid, A.F.M.Smith. Clarendon Press, Oxford, UK.
- [12] GUILLAUME, M.; (1999). Defining obesity in childhood: current practice. *The American Journal of Clinical Nutrition*, **70**(suppl.), 126S-130S.
- [13] HEIDELBERGER, P.; WELCH, P. D.; (1983). Simulation Run Length Control in the Presence of an Initial Transient. *Operations Research*, **31**, 1109-1144.
- [14] HITCHCOCK, D. B.; (2003). A History of the Metropolis-Hastings Algorithm. *The American Statistician*, **57**, n. 4, 254-257.
- [15] JOSEPH, L.; GYORKOS, T.; COUPAL, L.; (1995). Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. *American Journal of Epidemiology*, **141**, n.3, 263-272.
- [16] KALDOR, J.; CLAYTON, D.; (1985). Latent class analysis in chronic disease epidemiology. *Stat. Med.*, **4**, 327-335.

- [17] MADRUGA, M. R.; ESTEVES, L. G.; WECHSLER, S.; (2000). On the Bayesianity of Pereira-Stern Tests. Relatório Técnico do Departamento de Estatística, IME-USP, **10**, 1-9, (RT-MAE-2000-10).
- [18] MEDRONHO, A. R.; CARVALHO, D. M.; BLOCH, K. V., LUIZ, R. R.; WERNERCK, L. G.; (2002). Epidemiologia. São Paulo: Editora Atheneu.
- [19] MUST, A.; DALLAL, G. E.; DIETZ, W. H.; (1991). Reference data for obesity: 85<sup>th</sup> and 95<sup>th</sup> percentiles of body mass index (wt/ht<sup>2</sup>) and triceps skinfold thickness. *American Journal of Clinical Nutrition*, **53**, 839-846.
- [20] NEVES, D. P.; MELO, A. L.; GENARO, O.; LINARDI, P. M.; (2000). Parasitologia Humana. Ed. 10, São Paulo: Editora Atheneu.
- [21] PEREIRA, C. A. B.; STERN, J. M.; (1999). Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses. *Entropy*, **1**, 99-110.
- [22] PERRY, C. L.; STONE, E. J.; PARCEL, G. S.; ELLISON; NADER, P. R.; WEBBER, L. S.; LUEPKER, R. V.; (1990). School-based cardiovascular health promotion: the child and adolescent trial for cardiovascular health (CATCH). *Journal of School Health*, **60**, 406-413.
- [23] RAFTERY, A. E.; LEWIS, S.; (1992). How Many Iterations in the Gibbs Sampler? *In Bayesian Statistics*, **4**, ed. J.M.Bernardo, J.O.Berger, A.P.Dawid, A.F.M.Smith, 763-773. Oxford University Press.
- [24] SILVEIRA, V. D. R.; (1999). Inferência Bayesiana para pesquisa de mercado com erros de resposta utilizando modelos mistos. Tese de mestrado, Universidade de São Paulo, Instituto de Ciências Matemáticas.
- [25] SOTELO, Y. O. M.; (2003). Acurácia de Indicadores Antropométricos Transversais quando comparados com Critérios Longitudinais no Diagnóstico da Obesidade em Escolares-Projeto RRAMM. Tese de mestrado, Universidade Federal de São Paulo.

- [26] SPIEGELHALTER, D. J.; THOMAS, A.; BEST, N.; GILKS, W.; (1995). BUGS (Bayesian Inference Using Gibbs Sampling) Version 0.50. *MRC Biostatistics Unit*, Cambridge, UK.
- [27] STUNKARD, A. J.; (1990). The body mass index of twins who have been reared apart. *New England Journal of Medicine*, **322**, 1483-1487.
- [28] TANNER, M. A.; WONG, W. H.; (1987). The calculation of posterior densities by data augmentation. (With Discussion) *J. Am. Statist. Assoc.*, **82**, 528-550.
- [29] WALTER, S. D.; IRWIG, L.M.; (1988). Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. *J. Clin. Epidemiol*, **41**, 923-937.
- [30] ZLOCHEVSKY, E. R. M.; (1996). Obesidade na infância e adolescência. *Revista Paulista de Pediatria*, **14**, 124-133.