



Universidade de Évora  
Departamento de Matemática

MODELO COMBINADO  
CAPTURA-RECAPTURA E TRANSECTOS LINEARES:  
UMA ABORDAGEM BAYESIANA

João Filipe Gonçalves Monteiro

Orientador: Prof. Doutor Russell Gerardo Alpizar-Jara  
(Professor Associado)

Tese submetida para obtenção de grau de Doutor

em MATEMÁTICA

Évora  
2010

# Modelo combinado captura-recaptura e transectos lineares: Uma abordagem bayesiana

**Resumo:** Neste trabalho apresenta-se uma abordagem bayesiana para estimar a probabilidade de detectar um animal/objecto na distância zero, conhecida como  $g_0$ , utilizando o modelo combinado de captura-recaptura e transectos lineares (Alpizar-Jara e Pollock, 1999, Em *Marine Mammal Survey and Assessment Methods*, 99–114 pp.). Um estimador para o tamanho da população pode ser enviesado se a heterogeneidade não for considerada na modelação das probabilidades de captura, relativa às características inerentes dos indivíduos que são difíceis de medir ou não observáveis. Este tipo de problema tem sido tradicionalmente abordado mediante os modelos de captura-recaptura para populações fechadas, designados por  $\mathcal{M}_h$  e  $\mathcal{M}_{th}$ . Nesta tese formula-se um modelo generalizado combinado de captura-recaptura e transectos lineares para populações fechadas, que incorpora heterogeneidade nas probabilidades de detecção relativa às características inerentes dos indivíduos. A probabilidade de detectar um indivíduo em cima da linha do transecto percorrido, é estimada admitindo que é menor ou igual a 1. Assume-se que a probabilidade de avistar um animal depende de características individuais. A heterogeneidade observável nas probabilidades de captura dos indivíduos na população é modelada através da regressão logística utilizando covariáveis, tais como o sexo, a idade, ou o tamanho do grupo em que o animal se encontra. A heterogeneidade não observável é modelada através de um efeito aleatório, utilizando a inferência bayesiana. O parâmetro  $g_0$  é estimado como sendo uma média baseada na informação dos indivíduos como se estivessem na linha do transecto. O desempenho dos estimadores da probabilidade de um indivíduo ser observado na distância zero é analisado através de simulações, realizadas no programa R, e comparada com as situações em que apenas é modelada a heterogeneidade observável ou quando são modeladas ambas as heterogeneidades, observável e não observável. As distribuições *a posteriori* dos parâmetros que determinam a função de detecção são obtidas através do método de amostragem Gibbs através do método de Monte Carlo baseado em cadeias de Markov implementado no WINBUGS. Os resultados são ilustrados com dados reais da população de ungulados de montanha (*Rupicapra p. pyrenaica*) do Parque Nacional dos Perinéus (sul da França).

**Palavras chave:** função de detecção; heterogeneidade não-observável; inferência bayesiana; método de Monte Carlo baseado em cadeias de Markov; probabilidade de detecção na distância zero; selecção de modelos; WINBUGS.

# Combined model of capture-recapture and line transect: A Bayesian approach

**Abstract:** This work presents a Bayesian approach to estimate the probability of an animal/object being detected on the transect line, known as  $g_0$ , using the combined line transect and capture-recapture model (Alpizar-Jara and Pollock, 1999, In *Marine Mammal Survey and Assessment Methods*, 99–114 pp.). An estimator for population size is generally be biased under the presence of heterogeneity in capture probabilities, relative to the inherent characteristics of the individuals. That sort of heterogeneity is difficult to measure because it is not observable. This kind of problem has been traditionally approached using capture-recapture models for closed populations, designated by  $M_h$  and  $M_{th}$ . On this thesis it is formulated a generalized combined capture-recapture and line transect model used for closed populations that takes into account the heterogeneity in detection probability relative to the inherent characteristic of the individuals. The probability of sighting an animal (or an object) on the transect line, is estimated assuming that it is less or equal to 1. We assume that resighting probabilities depend on individuals' characteristics. Logistic regression is used to model observable heterogeneity in individual detection probabilities using covariates, such as, sex, age, and group size. Non-observable heterogeneity is modelled as a random effect.  $g_0$  is estimated as an average of individual based information as if each individual was on the center transect line. The performance of probability at distance zero estimators based on combined models is analysed through simulation, using software R, and compared when only observable heterogeneity is modelled, and when both observable and nonobservable heterogeneity are modelled. The posterior distributions of the key parameters of the detection function were obtained using Gibbs sampling through Markov chain Monte Carlo implemented in WINBUGS. The results are illustrated by an example using chamois (*Rupicapra p. pyrenaica*) population from Cauterets, Parc National des Pyrénées (southern France).

**Keywords:** Bayesian Inference; detection function; detection probability at distance zero; Markov chain Monte Carlo methods; model selection; nonobservable heterogeneity; WINBUGS.

*Aos meus pais,  
à Karine  
e aos meus amigos*

# Agradecimentos

Um trabalho de investigação como este tem a particularidade de ser, simultaneamente, solitário e irrealizável sem a colaboração de terceiros. Chegado a este momento, não posso deixar de me lembrar, com especial carinho, de todos aqueles que, das mais diferentes formas, me apoiaram.

- Ao meu orientador da tese e grande amigo, posso dizê-lo, o Professor Associado Doutor Russell Alpizar-Jara, do Departamento de Matemática da Universidade de Évora, que sempre se mostrou disponível e de quem recebi apoio e oportunas sugestões, não só durante a realização desta tese, mas também ao longo de todo o tempo em que trabalhei sob a sua orientação.
- À minha grande e dedicada amiga Soraya Gonçalves, uma palavra de simpatia, pela encorajadora confiança que sempre me transmitiu. Não posso deixar de realçar ainda as valiosas ideias e pistas sugeridas para a organização deste trabalho e cujo apoio na revisão desta tese se mostrou imprescindível.
- Com respeitosa amizade, recordo também o apoio, a força e o estímulo dos meus amigos e colegas de casa Evaldo Lima, Jorge Lopes e Romira Duarte, que me foram muito úteis durante o doutoramento. À Cláudia Marques, pelo apoio e encorajamento inicial para este doutoramento. Muito Obrigado!
- Exprimo os meus sinceros agradecimentos ao Prof. Jesús Maria Pérez da Universidade de Jaén, Espanha, pela ajuda prestada durante as investigações relacionadas com os dados de rebecos e na revisão de artigos publicados.
- Agradeço aos elementos do júri pelo cuidado na leitura e pelas sugestões feitas que muito melhoraram a apresentação desta tese.
- Estendo a minha sincera gratidão à Oriana Gonçalves, ao Hermes Baptista, à Anabela Afonso, ao Samora Lobo e ao Njalo Chipindu, que muito me apoiaram com a sua dedicada amizade.
- Especial agradecimento à minha muito querida esposa Karine, pelo amor, carinho, compreensão, paciência, estímulo, esforço e atenção. Minha companheira de pesquisa, minha companheira de vida. Não há gestos possíveis que possam expressar tamanha vontade de te ter ao meu lado, sempre.
- Finalmente, agradeço à minha família, em particular, aos meus pais Guilhermina e Miguel Monteiro, e à minha irmã Sofia Monteiro, mesmo à distância, nunca deixei de sentir o seu carinho e a sua atenção, especialmente nas horas mais difíceis. Um muito obrigado ao meu irmão Manuel Monteiro, por tudo.

Évora. Dezembro de 2010

*Filipe Monteiro*

## Agradecimentos institucionais

**FCT** Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR



Esta tese foi realizada no âmbito de uma bolsa de doutoramento (SFRH/BD/15346/2005), financiada pelo Gabinete de Relações Internacionais da Ciência e do Ensino Superior (GRICES) e pela Fundação para a Ciência e a Tecnologia (FCT).

- Agradeço ao Departamento de Matemática da Universidade de Évora pelo acolhimento e apoio logístico, e ao Centro de Investigação em Matemática e Aplicações (CIMA) da Universidade de Évora, pela ajuda financeira prestada nas deslocações aos congressos, onde apresentei parte do trabalho desenvolvido ao longo da elaboração desta tese.
- Gostaria de agradecer ao grupo de trabalho do Parque Nacional de Pirenéus, França, em particular ao Jean-Paul Campre, por ter disponibilizado os dados da amostragem dos rebecos, relativos aos anos 2002 e 2007.

# Notação e abreviaturas

Os termos que a seguir se apresentam foram elaborados de acordo com as traduções propostas no glossário estatístico inglês-português, realizado por comissões da Sociedade Portuguesa de Estatística e da Associação Brasileira de Estatística (SPE e ABE, 2007). As abreviaturas foram consideradas a partir das expressões em inglês, quando estas já são conhecidas e aceites. Nos restantes casos optou-se pela abreviatura da expressão em português.

## **Abreviaturas**

AIC – Critério de Informação de Akaike (*Akaike's Information Criteria*).

BIC – Critério de Informação Bayesiano (*Bayesian Information Criteria*).

CDS – Amostragem por Distâncias Convencional (*Conventional Distance Sampling*).

CR – Captura-Recaptura.

CV – Coeficiente de Variação.

D – Desviância (*Deviance*).

DAG – Grafos Dirigido Acíclico (*Directed Acyclic Graphs*).

DIC – Critério de Informação da Desviância (*Deviance Information Criteria*).

DS – Amostragem por Distâncias (*Distance Sampling*).

EMV – Estimadores de Máxima Verosimilhança.

ER – Enviesamento Relativo (em %).

HPD – Intervalo de maior Densidade *a Posteriori* (*Highest Posterior Density*).

---

H-T – Horvitz-Thompson.

IB – Inferência Bayesiana.

*i.i.d* – Independentes e Identicamente Distribuídos.

MC – Método de Monte Carlo.

MCAP – Modelo Combinado Alpizar-Pollock.

MCH – Modelo Combinado Heterogêneo (heterogeneidade observável).

MCH\* – Modelo Combinado Heterogêneo (heterogeneidade observável e heterogeneidade não observável).

MCHG\* – Modelo Combinado Heterogêneo Generalizado (heterogeneidade observável e heterogeneidade não observável).

MCMC – Método de Monte Carlo baseado em cadeias de Markov (*Markov Chain Monte Carlo*).

MCNH – Modelo Combinado Não Heterogêneo.

MV – Máxima Verosimilhança.

REQM – Raiz do Erro Quadrático Médio.

*sd* – Desvio Padrão (*Standard Deviation*).

*se<sub>b</sub>* – Desvio Padrão entre as médias das réplicas das cadeias (*Standard Error-Batch*).

## Notação

$c_{ij}$  – Função indicatriz que assume valor igual a 1, se o indivíduo  $i$  for capturado antes do momento  $j$ , e 0, no caso contrário.

$d_i$  – Distância horizontal entre o observador e o indivíduo  $i$ .

$f$  – Função de verosimilhança.

---

$g(y)$  – Função de detecção dado que o indivíduo está na distância  $y$ .

$g^*(y)$  – Função de detecção dado que o indivíduo está na distância  $y$ , na amostragem por distâncias convencional.

$h(\theta)$  – Distribuição *a priori* do parâmetro  $\theta$ .

$h(\theta|y)$  – Distribuição *a posteriori* do parâmetro  $\theta$  dada a informação por amostragem  $y$ .

$i$  – Índice referente aos indivíduos.

$j$  – Índice referente ao tempo.

$L$  – Comprimento do transecto linear.

$\mathcal{M}_h$  – Modelo de captura-recaptura, onde as probabilidade de captura incorporam heterogeneidade observável.

$\mathcal{M}_{h^*}$  – Modelo de captura-recaptura, onde as probabilidade de captura incorporam heterogeneidade observável e não observável.

$m_i$  – Ângulo entre a linha do transecto e o segmento de recta  $r_i$ .

$M_i$  – Modelo logístico definido para  $g_0$ ,  $i = 1, \dots, 7$ , tendo em conta as covariáveis  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$  e  $\delta_i$ . Todos os 7 modelos incluem o efeito aleatório individual  $\delta_i$ , além do efeito global  $\beta_0$ .

$\mathcal{M}_t$  – Modelo de captura-recaptura, onde as probabilidade de captura variam com o tempo.

$\mathcal{M}_{th}$  – Modelo de captura-recaptura, onde as probabilidade de captura variam com o tempo e incorporam heterogeneidade observável.

$\mathcal{M}_{th^*}$  – Modelo de captura-recaptura, onde as probabilidade de captura variam com o tempo e incorporam heterogeneidade observável e não observável.

$N$  – Tamanho da população.

---

$n(= n_{11} + n_{10} + n_{01})$  – Total de indivíduos marcados ou avistados.

$\mathbf{n} = (n_{10}, n_{01}, n_{11})$  – Vector com informação sobre os indivíduos marcados e avistados.

$n_{01}$  – Número de indivíduos avistados, mas que não foram marcados previamente.

$n_1$  – Número de indivíduos marcados.

$n_{10}$  – Número de indivíduos marcados, que não foram posteriormente avistados.

$n_{11}$  – Número de indivíduos avistados que foram marcados previamente.

$n_2$  – Número de indivíduos avistados através do protocolo de amostragem por transectos lineares.

$\hat{N}_{LP}$  – Estimador do tamanho da população do modelo de Lincoln-Petersen.

$\hat{N}_{MCAP}$  – Estimador do tamanho da população do modelo combinado de Alpizar-Jara e Pollock.

$\hat{N}_{MCH^*}$  – Estimador do tamanho da população do modelo combinado heterogéneo.

$\hat{N}_{MCHG^*}$  – Estimador do tamanho da população do combinado heterogéneo generalizado.

$\hat{N}_{TL}$  – Estimador do tamanho da população do modelo de transectos lineares.

$o$  – Índice referente aos observadores.

$p$  – Probabilidade de um indivíduo ter sido marcado ou avistado.

$\mathbf{p}$  – Matriz das probabilidades  $p_{ij}$ .

$p_1$  – Probabilidade de um indivíduo ter sido marcado.

$p_2$  – Probabilidade de um indivíduo ter sido avistado.

$p_{ij}$  – Probabilidade do indivíduo  $i$  ser capturado no momento de amostragem  $j$ .

$p_{ij}^*$  – Probabilidade de o indivíduo  $i$  ser capturado no momento  $j$  quando  $c_{ij} = 0$ .

---

$r_i$  – Distância entre o observador e o indivíduo  $i$ .  
 $\mathbf{x}$  – Matriz com informação sobre o histórico das capturas.  
 $x_{ij}$  – Função indicatriz que assume valor igual a 1, se o indivíduo  $i$  for capturado no momento  $j$ , e 0, no caso contrário.  
 $w$  – Distância máxima de observação de um indivíduo no transecto linear.  
 $y_i$  – Distância perpendicular, em relação à linha do transecto, do indivíduo  $i$ .  
 $\mathbf{y} = (y_1, \dots, y_{n_2})$  – Vector com informação sobre as distâncias perpendiculares.  
 $\mathbf{z}$  – Vector com informação sobre as características individuais.  
 $\beta_0$  – Explica o efeito global.  
 $\beta_{\text{idade}}$  – Explica o efeito idade.  
 $\beta_{\text{tempo}_j}$  – Explica a variação da probabilidade de captura entre os  $j - 1$  momentos de amostragem.  
 $\beta_{\text{tempo}} = (\beta_{\text{tempo}_1}, \dots, \beta_{\text{tempo}_t})$  – Vector dos parâmetros que explicam a variação entre os momentos de amostragem.  
 $\beta_{\text{peso}}$  – Explica o efeito peso.  
 $\beta_{\text{sexo}}$  – Explica o efeito sexo.  
 $\beta$  – Vector de parâmetros desconhecidos, relativo aos efeitos individuais observáveis.  
 $\delta_i$  – Efeito aleatório do indivíduo  $i$ .  
 $\delta = (\delta_1, \dots, \delta_n)$ .  
 $\theta$  – Vector dos parâmetros desconhecidos.  
 $\Theta$  – Espaço paramétrico de  $\theta$ .  
 $\gamma_{ij}$  – Probabilidade de o indivíduo  $i$  ser capturado no momento  $j$ , dado o histórico da captura e sabendo que foi capturado pelo menos uma vez.

# Índice

Resumo	i
Abstract	ii
Agradecimentos	iii
Notação e abreviaturas	vi
Índice	xi
Lista de tabelas	xiv
Lista de figuras	xviii
<b>1 Introdução</b>	<b>1</b>
1.1 Aspectos gerais . . . . .	2
1.2 Amostragem por distâncias . . . . .	3
1.3 Amostragem por captura-recaptura . . . . .	4
1.4 Modelo combinado captura-recaptura e transectos lineares . . . . .	7
1.5 O método bayesiano e o algoritmo MCMC . . . . .	8
1.5.1 Introdução . . . . .	8
1.5.2 Teorema de Bayes . . . . .	9
1.5.3 <i>Priori e posteriori</i> . . . . .	10
1.5.4 Modelos hierárquicos . . . . .	12
1.5.5 MCMC . . . . .	14
1.6 O modelo combinado e a inferência bayesiana . . . . .	15
1.7 Estrutura da tese . . . . .	17

<b>2</b>	<b>Modelando heterogeneidade não observável em probabilidades de captura-recaptura</b>	<b>20</b>
2.1	Introdução . . . . .	21
2.2	Modelo captura-recaptura tipo $\mathcal{M}_{th}$ . . . . .	25
2.3	Modelando a heterogeneidade não observável . . . . .	28
2.3.1	Função de verosimilhança . . . . .	29
2.3.2	Distribuições <i>a priori</i> . . . . .	31
2.3.3	Distribuições <i>a posteriori</i> . . . . .	33
2.4	Aplicação em capturas de ratos veadeiro . . . . .	37
2.4.1	Análises de diagnóstico . . . . .	38
2.4.2	Análise de sensibilidade . . . . .	41
2.4.3	Seleção de modelos . . . . .	47
2.4.4	Resultados e análise do modelo . . . . .	52
<b>3</b>	<b>Estimação bayesiana de <math>g_0</math> em amostragem por distâncias</b>	<b>56</b>
3.1	Introdução . . . . .	57
3.2	Modelo de transectos lineares . . . . .	58
3.2.1	Pressuposto crítico ( $g_0 = 1$ ) . . . . .	60
3.3	Combinação dos modelos de Lincoln-Petersen e transectos lineares . . . . .	61
3.3.1	Função de verosimilhança . . . . .	62
3.4	Estimação bayesiana de $g_0$ . . . . .	64
3.4.1	Distribuições <i>a priori</i> . . . . .	65
3.4.2	Distribuições <i>a posteriori</i> . . . . .	69
3.5	Estudo de simulação . . . . .	74
3.5.1	Análises de diagnóstico . . . . .	74
3.5.2	Resultados e discussão . . . . .	77
3.6	Experiência com latas de cervejas . . . . .	82
3.6.1	Análises de diagnóstico . . . . .	83
3.6.2	Análise de sensibilidade . . . . .	84
3.6.3	Resultados . . . . .	85
3.7	Aplicação nas estimação da densidade de ungulados de montanha . . . . .	88
3.7.1	Área de estudo . . . . .	88
3.7.2	Análises de diagnóstico . . . . .	89
3.7.3	Resultados . . . . .	90
<b>4</b>	<b>Estimação bayesiana da probabilidade na distância zero utilizando o modelo de dupla plataforma e heterogeneidade</b>	<b>94</b>
4.1	Introdução . . . . .	95
4.2	Modelando heterogeneidade observável e não observável em $g_0$ . . . . .	97

4.2.1	Função de verosimilhança . . . . .	98
4.2.2	Distribuição <i>a priori</i> . . . . .	100
4.2.3	Distribuições <i>a posteriori</i> . . . . .	101
4.3	Estimação da densidade de ungulados de montanha . . . . .	109
4.3.1	Análises de diagnóstico . . . . .	110
4.3.2	Análise de sensibilidade . . . . .	113
4.3.3	Seleção de modelos . . . . .	120
4.3.4	Resultados e análise do modelo . . . . .	122
4.4	Generalização do modelo de dupla plataforma e heterogeneidade . . . . .	126
4.4.1	Função de verosimilhança . . . . .	128
4.4.2	Distribuição <i>a priori</i> . . . . .	129
4.4.3	Distribuição <i>a posteriori</i> . . . . .	130
<b>5</b>	<b>Conclusão</b>	<b>132</b>
5.1	Notas sobre MCMC . . . . .	133
5.2	Principais desenvolvimentos e resultados . . . . .	135
5.2.1	Modelação da heterogeneidade não observável em probabilidades de captura-recaptura . . . . .	135
5.2.2	Modelo combinado sem heterogeneidade . . . . .	136
5.2.3	Modelo combinado que incorpora heterogeneidade observável e não observável em $g_0$ . . . . .	137
	<b>Bibliografia</b>	<b>139</b>
	<b>Apêndices</b>	<b>159</b>
<b>A</b>	<b>Código WINBUGS</b>	<b>159</b>
A.1	Modelo de captura-recaptura tipo $\mathcal{M}_{th}^*$ . . . . .	160
A.2	Modelo <i>MCNH</i> . . . . .	161
A.3	Modelo <i>MCH*</i> . . . . .	162
A.4	Seleção de modelos para o modelo <i>MCH*</i> . . . . .	163
A.5	Modelo <i>MCHG*</i> . . . . .	164
<b>B</b>	<b>Dados</b>	<b>165</b>
B.1	Dados de capturas de ratos . . . . .	166
B.2	Experiência de Otto e Pollock (1990) . . . . .	167
B.3	Dados de rebecos no ano 2002 . . . . .	168
B.4	Dados de rebecos no ano 2007 . . . . .	169

# Lista de Tabelas

2.1	Estimativa do período de aquecimento e da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50) para as cadeias de Markov dos parâmetros $\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}_1}, \beta_{\text{tempo}_2}, \beta_{\text{tempo}_3}, \beta_{\text{tempo}_4}, \beta_{\text{tempo}_5}, \tau_\delta, \delta_1, \delta_5, \delta_{10}, \delta_{15}, \delta_{20}, \delta_{25}, \delta_{30}$ e $\delta_{38}$ , quando o modelo utilizado é do tipo $\mathcal{M}_{th^*}$ . . . . .	41
2.2	Resultado da análise de sensibilidade para os parâmetros $\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}_1}$ e $\beta_{\text{tempo}_2}$ , tendo em conta as distribuições <i>a priori</i> nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição <i>a posteriori</i> é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD. . . . .	45
2.3	Resultado da análise de sensibilidade para os parâmetros $\beta_{\text{tempo}_3}, \beta_{\text{tempo}_4}, \beta_{\text{tempo}_5}, \tau_\delta, \delta_1$ e $\delta_{38}$ , tendo em conta as distribuições <i>a priori</i> nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição <i>a posteriori</i> é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD. . . . .	46
2.4	Comparação entre $p_\theta, p_D$ e os resultados obtidos para os diferentes critérios (AIC, BIC, D e DIC) para selecção dos modelos $\mathcal{M}_t, \mathcal{M}_h, \mathcal{M}_{th}$ e $\mathcal{M}_{th^*}$ , considerando o método de máxima verosimilhança (MV) e a inferência bayesiana (IB). . . . .	51

2.5 Estimativas, coeficiente de variação [CV%] e desvio padrão (*sd*) obtidos com os métodos de máxima verosimilhança (MV) e inferência bayesiana (IB), para os parâmetros  $N$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$  e  $\beta_{\text{tempo}_5}$ , considerando os modelos  $\mathcal{M}_t$ ,  $\mathcal{M}_h$ ,  $\mathcal{M}_{th}$  e  $\mathcal{M}_{th^*}$ . . . . . 52

3.1 Estimativa do período de aquecimento e da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50), para as cadeias de Markov dos parâmetros  $\lambda$ ,  $g_0$ ,  $\sigma^2$  e  $p_1$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários (1,2,3,4) considerados para a simulação de dados para o parâmetro  $g_0$ , respectivamente para  $g_0 = 0.50, 0.70, 0.90$  e  $1.00$ . 77

3.2 Percentagem de enviesamento relativo e coeficiente de variação (CV%), das estimativas *a posteriori*, dos parâmetros  $N_{CMAP}$ ,  $p_1$  e  $p_2(g_0, \sigma)$ , tendo em conta todos os cenários considerados para a simulação de dados. . . . . 79

3.3 Resumo da distribuição *a posteriori* do parâmetro  $g_0$  nos diferentes cenários considerados para a simulação de dados. Apresentam-se o primeiro quartil ( $Q_1$ ), a mediana (*me*), o terceiro quartil ( $Q_3$ ), o desvio padrão (*sd*), a percentagem do enviesamento relativo (%*ER*), a percentagem da raiz do erro quadrático médio (%*REQM*) e o coeficiente de variação (CV%). . . . . 80

3.4 Quadro comparativo entre as estimativas obtidas através do método de Monte Carlo, utilizando o estimador de máxima verosimilhança, com as estimativas da distribuição *a posteriori*, utilizando o método de Monte Carlo baseado em cadeias de Markov, para os parâmetros  $g_0$ ,  $N$ ,  $p_1$ ,  $p_2$  e  $\mu^*$ . Foram utilizados três modelos: Lincoln-Petersen, transectos lineares e combinado. Apresentam-se as respectivas estimativas pontuais, o intervalo de confiança ou credibilidade/HPD e o desvio padrão (*sd*). . . . . 86

3.5 Estimativa obtida da análise da densidade, do coeficiente de variação (CV%) e do intervalo de confiança (IC) para a população de rebecos em *Clot-Cayan*, Parque Nacional de Pirenéus, utilizando dados de amostragem realizada no mês de Outubro do ano 2002. Foram consideradas quatro funções de detecção com ajustamento de co-seno: Taxa de risco, Exponencial negativa, Uniforme e Semi- -Normal. A selecção de modelos é feita usando o Critério de Informação de Akaike (AIC). . . . . 90

- 3.6 Estimativa obtida da análise do tamanho da população ( $N$ ) e da estimativa para  $g_0$  e o respectivo coeficiente de variação (CV%) e intervalo de confiança (IC), utilizando o estimador de máxima verosimilhança do modelo combinado proposto por Alpizar-Jara e Pollock (1999), para a população de rebecos em *Clot-Cayan*, Parque Nacional de Pirenéus, utilizando dados de amostragem realizada no mês de Outubro do ano 2002. Foram consideradas quatro funções de detecção com ajustamento de co-seno: Taxa de risco, Exponencial negativa, Uniforme e Semi-Normal. . . . . 92
- 3.7 Estimativa, intervalo de credibilidade/HPD e desvio padrão ( $sd$ ) obtidos para os parâmetros  $g_0$ ,  $N$ ,  $p_1$  e  $p_2$ , utilizando o modelo combinado, para a população de rebecos em *Clot-Cayan*, Parque Nacional de Pirenéus, utilizando dados de amostragem do ano 2002. . . . . 93
- 4.1 Estimativa do período de aquecimento e da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50), para as cadeias de Markov dos parâmetros  $N$ ,  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\tau_\delta$ ,  $\delta_1$  e  $\delta_{72}$ , para o modelo *MCH\**. . . . . 113
- 4.2 Resultado da análise de sensibilidade para os parâmetros  $N$ ,  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ , tendo em conta as distribuições *a priori* nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição *a posteriori* é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD. . . . . 118
- 4.3 Resultado da análise de sensibilidade para os parâmetros  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$ ,  $\tau_\delta$ ,  $\delta_1$ ,  $\delta_{36}$  e  $\delta_{72}$ , tendo em conta as distribuições *a priori* nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição *a posteriori* é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD. . . . . 119

- 4.4 Estimativas e intervalo de credibilidade de  $N$  e estimativas para os parâmetros do modelo logístico para  $g_0$ , quando apenas a heterogeneidade observável foi modelada, e quando ambas, heterogeneidade observável e não observável, foram modeladas considerando diferente modelos,  $M_0 - M_7$ , Modelo Combinado Não Heterogéneo ( $MCNH$ ) e Modelo Combinado Heterogéneo ( $MCH$ ). Note-se que os Modelos 1-7 são Modelos Combinados Heterogéneo que incorporam factores observáveis e não observáveis ( $MCH^*$ ). . . . . 125
- 5.1 Resumo da bibliografia que tem sido utilizada para as abordagens clássica e bayesiana para os modelos:  $\mathcal{M}_0$ ,  $\mathcal{M}_t$ ,  $\mathcal{M}_h$ ,  $\mathcal{M}_{h^*}$ ,  $\mathcal{M}_{th}$ ,  $\mathcal{M}_{th^*}$ ,  $DS$ ,  $MC$ ,  $MCH^*$ ,  $MCHG^*$ . . . . . 134
- B.1 Dados de capturas de ratos (*Peromyscus maniculatus*) realizadas por V. Reid, na zona Este de Stuart Gulch em Colorado, EUA. Foram realizados 6 momentos de amostragem, de onde foram obtidas características individuais de 38 ratos, tais como o sexo (f=fêmea e m=macho), a idade (j=juvenil, sa=semi-adulto e a=adulto) e o peso (em gramas). . . . . 166
- B.2 Dados da experiência feita por Otto, em 1982 (Otto e Pollock, 1990), com latas castanhas de cerveja, num campo perto de Raleigh, nos Estados Unidos. Cada grupo individual de distâncias perpendiculares das latas em relação à linha central do transecto é apresentado juntamente com quantos dos 9 observadores que o avistaram. O número 0 significa que a lata de cerveja não foi avistada. . . . . 167
- B.3 Dados da amostragem, do ano 2002, das distâncias perpendiculares da população de rebecos (*Rupicapra p. pyrenaica*), em *Clot-Cayan*, Parque Nacional de Pirenéus. . . . . 168
- B.4 Dados da amostragem da população de rebecos (*Rupicapra p. pyrenaica*), em *Clot-Cayan*, do ano 2007, Parque Nacional de Pirenéus, França. Foram avistados 72 rebecos, dos quais foram registadas características individuais, tais como a idade (1=cria/juvenil e 2=adulto), o sexo (0=fêmea e 1=macho) e o tamanho do grupo do qual o animal pertencia. . . . . 169

# Lista de Figuras

1.1	Exemplo de um grafo dirigido acíclico (DAG) do modelo hierárquico Poisson-Gama para o tamanho da população ( $N$ ), $N \lambda \sim Poi(\lambda)$ e $\lambda \sim Gama(\tau, \eta)$ , onde $N = 1, 2, \dots$ , $\lambda > 0$ , $\tau, \eta \in \mathbb{R}$ . . . . .	13
1.2	Diagrama do modelo combinado heterogéneo generalizado ( $MCHG^*$ ) e dos submodelos subsequentes: amostragem por distâncias convencional ( $CDS$ ); modelos de captura-recaptura quando as probabilidades de captura: são constantes ( $\mathcal{M}_0$ ), variam com o tempo ( $\mathcal{M}_t$ ), dependem de características individuais ( $\mathcal{M}_h$ ), dependem de características individuais e de outros factores não observáveis ( $\mathcal{M}_{h^*}$ ), variam com o tempo e dependem de características individuais ( $\mathcal{M}_{th}$ ), variam com o tempo e dependem de características individuais e de outros factores não observáveis ( $\mathcal{M}_{th^*}$ ) e modelo combinado de captura-recaptura e transectos lineares, quando as probabilidades de avistar dependem de características individuais e de outros factores não observáveis ( $MCH^*$ ). . . . .	18
2.1	DAG do modelo de captura-recaptura quando as probabilidades de captura variam com o tempo e dependem de características individuais e de outros factores não observáveis ( $\mathcal{M}_{th^*}$ ) para estimar o tamanho da população. Os rectângulos maiores representam os ciclos em torno das probabilidades de captura $p_{ij}$ ( $i = 1, \dots, n$ e $j = 1, \dots, t$ ) e dos efeitos aleatórios individuais $\delta_i$ ( $i = 1, \dots, n$ ). . . . .	33
2.2	Traço das duas cadeias de Markov, com pontos iniciais diferentes, consideradas para os parâmetros $\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}_1}, \beta_{\text{tempo}_2}, \beta_{\text{tempo}_3}, \beta_{\text{tempo}_4}, \beta_{\text{tempo}_5}, \tau_\delta, \delta_1$ e $\delta_{38}$ , quando o modelo utilizado é do tipo $\mathcal{M}_{th^*}$ . . . . .	39

2.3 Teste de convergência Gelman e Rubin, utilizando duas cadeias de Markov com pontos iniciais diferentes, para os parâmetros  $\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}_1}, \beta_{\text{tempo}_2}, \beta_{\text{tempo}_3}, \beta_{\text{tempo}_4}, \beta_{\text{tempo}_5}, \tau_\delta, \delta_1$  e  $\delta_{38}$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ . - - - - - 97.5%, ——— mediana. . . . . 40

2.4 Gráfico das distribuições *a priori* escolhidas para análise de sensibilidade para os parâmetros:  $(\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}}, \delta)$  a) e  $\tau_\delta$  b). . . . . 43

2.5 Gráfico das distribuições *a posteriori* para os parâmetros  $\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}_1}, \beta_{\text{tempo}_2}, \beta_{\text{tempo}_3}, \beta_{\text{tempo}_4}, \beta_{\text{tempo}_5}, \tau_\delta, \delta_1$  e  $\delta_{38}$ , obtidas tendo em conta as distribuições *a priori* dos quatro cenários. 44

2.6 Função distribuição *a posteriori* do parâmetro  $N$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ . . . . . 53

2.7 Função distribuição *a posteriori* dos parâmetros  $\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}_1}, \beta_{\text{tempo}_2}, \beta_{\text{tempo}_3}, \beta_{\text{tempo}_4}, \beta_{\text{tempo}_5}, \tau_\delta, \delta_1$  e  $\delta_{38}$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ . . . . . 54

3.1 Transecto linear ao longo da faixa de largura  $2w$  e comprimento  $L$ , a). Os objectos estão representados por  $\bullet$  e os segmentos de recta ilustram as distâncias perpendiculares,  $y_i$ , aos objectos detectados. Ao longo do transecto medem-se as distâncias  $r_i$  e os ângulos  $m_i$  para calcular as distâncias perpendiculares  $y_i$ , b). Note-se que  $d_i$  é a distância horizontal entre o observador e o indivíduo  $i$ . Adaptado de Buckland *et al.* (1993, p. 5). . . . . 58

3.2 DAG do modelo combinado de captura-recaptura e transectos lineares (*MCNH*), para estimar a probabilidade de detectar um indivíduo na distância zero,  $g_0$ . . . . . 70

3.3 Traço das duas cadeias de Markov, com pontos iniciais diferentes, consideradas para o parâmetro  $g_0$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários considerados para a simulação de dados para  $g_0$ , i.e.  $g_0 = 0.50, 0.70, 0.90, 1.00$ . . . . . 75

3.4 Teste de convergência Gelman e Rubin, utilizando duas cadeias de Markov com pontos iniciais diferentes, para o parâmetro  $g_0$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários considerados para a simulação de dados para  $g_0$ , i.e.  $g_0 = 0.50, 0.70, 0.90, 1.00$ . - - - - - 97.5%, ——— mediana. . . . . 76

3.5 Distribuição *a posteriori* para  $g_0$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários considerados para a simulação de dados para  $g_0$ , i.e.  $g_0 = 0.50, 0.70, 0.90, 1.00$ . . . . . 81

3.6	Gráfico do teste de convergência de Gelman e Rubin (a), utilizando duas cadeias de Markov com pontos iniciais diferentes, e gráfico da estimativa da função de autocorrelação (b), utilizando uma cadeia de Markov, para o parâmetro $g_0$ . - - - - - 97.5%, ——— mediana. . . . .	84
3.7	Gráfico das distribuições <i>a priori</i> (a), escolhidas para análise de sensibilidade para o parâmetro $g_0$ , e as respectivas distribuições <i>a posteriori</i> (b). . . . .	85
3.8	Função distribuição <i>a posteriori</i> dos parâmetros $g_0$ , (a), e $N$ , (b). . . . .	87
3.9	Gráficos para o teste de convergência de Gelman e Rubin (a), utilizando duas cadeias de Markov com pontos iniciais diferentes, a estimativa da função de autocorrelação (b) e o traço (c), utilizando uma cadeia de Markov, para o parâmetro $g_0$ . - - - - - 97.5%, ——— mediana. . . . .	89
3.10	Distribuição <i>a posteriori</i> para $g_0$ . . . . .	93
4.1	DAG do modelo combinado de captura-recaptura e transectos lineares, quando as probabilidades de captura dependem de características individuais e de outros factores não observáveis ( $MCH^*$ ), para estimar a probabilidade de detectar um indivíduo na distância zero, $g_0$ . O retângulo maior representa o ciclo em torno dos $g_0$ individuais ( $g_{0_i}, i = 1, \dots, n_2$ ). . . . .	100
4.2	Transecto linear, ver linha traçada a branco, com 4.471 km de comprimento, realizado em “Clot-Cayan”, em Outubro de 2007, entre as 08:55 horas e 12:50 horas. . . . .	109
4.3	Traço das duas cadeias de Markov, com pontos iniciais diferentes, consideradas para os parâmetros $N, \lambda, p_1, \sigma^2, \beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}}, \tau_\delta, \delta_1, \delta_{36}$ e $\delta_{72}$ , para o modelo $MCH^*$ . . . . .	111
4.4	Teste de convergência Gelman e Rubin, utilizando duas cadeias de Markov com pontos iniciais diferentes, para os parâmetros $N, \lambda, p_1, \sigma^2, \beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}}, \tau_\delta, \delta_1, \delta_{36}$ e $\delta_{72}$ , para o modelo $MCH^*$ . - - - - - 97.5%, ——— mediana. . . . .	112
4.5	Gráfico das distribuições <i>a priori</i> escolhidas para a análise de sensibilidade para os parâmetros: $(\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}}, \boldsymbol{\delta})$ e $\tau_\delta$ , figuras a) e b) respectivamente. . . . .	114
4.6	Gráfico das distribuições <i>a posteriori</i> para o parâmetros $N$ , obtidas tendo em conta as distribuições <i>a priori</i> dos quatro cenários. Figuras a), b), c) e d), respectivamente. . . . .	116
4.7	Gráfico das distribuições <i>a posteriori</i> para os parâmetros $\lambda, p_1, \sigma^2, \beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}}, \tau_\delta, \delta_1, \delta_{36}$ e $\delta_{72}$ , obtidas tendo em conta as distribuições <i>a priori</i> dos quatro cenários. . . . .	117

- 
- 4.8 Distribuição *a posteriori* para selecção de modelo. Os números e as barras designam os modelos e o seu respectivo valor de probabilidade *a posteriori*. . . . . 123
- 4.9 Distribuição *a posteriori* do parâmetro  $g_0$  quando se considera o modelo não heterogéneo a) e o modelo heterogéneo que incorpora factores observáveis e não observáveis b). . . . . 126
- 4.10 DAG do modelo combinado generalizado de captura-recaptura e transectos lineares, quando as probabilidades de captura dependem de características individuais e de outros factores não observáveis ( $MCHG^*$ ), para estimar a probabilidade de detectar um indivíduo na distância zero pelo observador  $o$ ,  $g_0^o$ . Os rectângulos maiores representam os ciclos em torno dos  $g_0^o$  individuais ( $g_{0,i}^o, i = 1, \dots, n_2^o, o = 1, \dots, O$ ). . . 131

# Capítulo 1

## Introdução

## 1.1 Aspectos gerais

As estimativas do tamanho das populações naturais são necessárias para a sua monitorização e para a avaliação do estado dos ecossistemas. Neste sentido, formulam-se modelos matemáticos e desenvolvem-se métodos que tentam estimar a abundância e densidades populacionais (Cormack, 1988; Schwarz e Seber, 1999; Seber, 1986, 1992; Seber e Schwarz, 2002). O tipo de modelo utilizado para estimação depende da natureza da população em estudo, que pode ser caracterizada como população “aberta” ou “fechada”. Em contraste com uma população aberta, numa população fechada assume-se que não existem nascimentos ou imigrações e nem mortes ou emigrações de indivíduos. Durante os últimos dois séculos, métodos tais como captura-recaptura<sup>1</sup> (CR), amostragem por distâncias<sup>2</sup> (DS), e outros métodos *ad hoc* têm sido desenvolvidos para estimar populações animais (e.g., Borchers *et al.*, 2002; Buckland *et al.*, 2000; Pollock, 1991, 2000). Os modelos mais utilizados na estimação do tamanho de populações fechadas de animais baseiam-se em amostragem por captura-recaptura (Williams *et al.*, 2002) e amostragem por distâncias (Buckland *et al.*, 2001, 2004).

---

<sup>1</sup>O primeiro desenvolvimento do método de captura-recaptura é atribuído ao matemático francês Pierre Laplace (1749-1827), que também é considerado por muitos o pai da probabilidade. Em 1802, utilizou o método de captura-recaptura para estimar o tamanho da população da França (Cochran, 1978; Stigler, 1986). Contudo, muitas vezes este método é também associado ao inglês John Graunt (1620-1674), que utilizou um método similar para estimar o efeito de pragas, e também a população da Inglaterra no início de 1600 (Hald, 1990).

<sup>2</sup>A amostragem por distâncias foi utilizada, pela primeira vez, em 1906, por Alfred Otto Gross (1883-1970) para amostrar pássaros no Illinois, nos Estados Unidos (Forbes, 1907; Forbes e Gross, 1921). Segundo Leopold (1933) e Gates (1979), entre 1929 e 1932, Ralph Terence King fez a primeira tentativa, num trabalho não publicado, de utilização da técnica de amostragem por distâncias para estimar abundância de galos silvestre (*Bonasa umbellus*), na floresta Cloquet, nos Estados Unidos. Desde então, e sobretudo depois da década de 80, muitos autores têm contribuído para o desenvolvimento deste método (e.g., Buckland *et al.*, 1993, 2001, 2004; Burnham *et al.*, 1980).

## 1.2 Amostragem por distâncias

A amostragem por distâncias é uma técnica muito utilizada para estimar a abundância de animais. Tem sido proposta para estimar a densidade de algumas populações de animais, como por exemplo, urso castanhos (*Ursus arctos middendorffi*), na ilha de Kodiak, no Alaska (Quang e Becker, 1999), toninhas do cais (*Phocoena phocoena*), no norte da Califórnia (Carretta *et al.*, 2001), foca caranguejeira (*Lobodon carcinophagus*), no Antártico (Borchers *et al.*, 2006), e várias espécies de aves (Kenneth, 2001). Esta técnica baseia-se em avistar os animais e na medição das distâncias perpendiculares,  $y$ , da linha (modelo de transecto linear) ou distância radial a partir de um ponto (modelo de transecto pontual) até aos indivíduos observados. Depois disso, é ajustada uma função, designada por função de detecção,  $g(y)$ . Esta função representa a probabilidade de detectar um animal dado que está à distância  $y$  do transecto percorrido (Buckland *et al.*, 2001).

Na amostragem por distâncias convencional (CDS), um pressuposto crítico é que todos os animais em cima da linha do transecto são detectados com probabilidade 1, i.e.,  $g_0 = g(0) = 1$ .

Este pressuposto é muitas vezes violado, por exemplo nas amostragens aéreas e marinhas, onde existem duas fontes de enviesamento: percepção e disponibilidade. O enviesamento devido a percepção pode ocorrer, porque o observador tem pouca capacidade de detectar os animais ou devido às más condições ambientais. Por exemplo, o vento, a precipitação e o grau de luminosidade podem afectar a percepção do observador. Também o observador pode não ter muita experiência, e pode ter dificuldades de visão ou audição. O enviesamento devido a disponibilidade pode

ocorrer porque o animal pode estar escondido debaixo de uma árvore ou submergido num lago.

O estimador para a densidade da população é enviesado quando se assume que a probabilidade de detectar um indivíduo na linha do transecto é igual a 1, i.e.  $g_0 = 1$ , mas na verdade tem-se que é inferior a 1,  $g_0 < 1$ . Neste caso, o efeito da heterogeneidade nas probabilidades de detecção e o enviesamento nas estimativas podem ser severos. A heterogeneidade nas probabilidades de detecção pode ocorrer devido à distância e muitas outras variáveis, como por exemplo o tamanho ou o comportamento do animal, a estação do ano, etc. Este problema (da incerteza em torno de  $g_0$ ) foi levantado pela primeira vez por Pollock e Kendall (1987) e posteriormente abordado por outros autores (e.g., Alpizar-Jara e Pollock, 1996, 1999; Borchers, 1996; Borchers *et al.*, 1998a,b; Buckland e Turnock, 1992; Laake e Borchers, 2004; Manly *et al.*, 1996; Okamura *et al.*, 2003; Palka, 1995; Quang e Becker, 1997), que recomendam que o método de captura-recaptura deve ser utilizado para diminuir o enviesamento causado pela percepção e que permite estimar  $g_0$ .

### 1.3 Amostragem por captura-recaptura

A amostragem por captura-recaptura consiste em captura-marcação-recaptura de indivíduos numa população que se assume como fechada (Borchers *et al.*, 2002) ou aberta (Pollock e Alpizar-Jara, 2005). Este método é frequentemente utilizado para estimar parâmetros demográficos, tais como abundância animal, taxas de sobrevivência, etc. (e.g., Borchers *et al.*, 2002; Seber, 1982; Williams *et al.*, 2002). Entretanto, este modelo tem sido usado em diferentes áreas, como por exemplo,

aplicações na epidemiologia (Hook e Regal, 1995; IWGDMF, 1995a,b), pesquisas na área de VIH/SIDA (Bernillon *et al.*, 2000) e também para avaliar a fiabilidade de programas de computação (Basu e Ebrahimi, 2001).

O modelo mais simples de captura-recaptura para populações fechadas é conhecido como o modelo de Lincoln-Petersen<sup>3</sup> (Lincoln, 1930; Petersen, 1896), que envolve apenas dois momentos de amostragem. A partir dos anos 50, outros autores publicaram vários trabalhos para estimar parâmetros populacionais por captura-marcação-recaptura tais como Burnham (1972); Chapman (1951); Cowan e Malec (1986); Darroch (1958, 1959); Jolly (1965, 1982); Parker (1963); Pollock (1974, 1991); Pollock e Otto (1983); Pollock *et al.* (1984); Seber (1965, 1982, 1986, 1992); Wang (2002).

Sendo  $N$  o número total de indivíduos que se pretende estimar, numa primeira amostra, capturam-se  $n_1$  indivíduos. Marcam-se esses indivíduos capturados que posteriormente são libertados para o seu habitat natural, juntando-se aos restantes indivíduos não marcados. Numa segunda amostra, capturam-se  $n_2$  indivíduos, podendo alguns destes,  $n_{11}$ , já terem sido capturados na primeira amostra, obtendo-se assim a recaptura.

Espera-se que a proporção de indivíduos marcados na segunda amostra (recaptura) seja aproximadamente igual à proporção de indivíduos marcados no total da população:

$$\frac{n_{11}}{n_2} \approx \frac{n_1}{N} \Rightarrow \hat{N}_{LP} = \frac{n_1 n_2}{n_{11}}, n_{11} > 0,$$

---

<sup>3</sup>Este modelo foi utilizado pelo dinamarquês Carl George Johannes Petersen (1860-1928), em 1896, para estudar o fluxo migratório de peixes no mar Báltico e, em 1930, pelo americano Frederick Charles Lincoln (1892-1960) para estimar o tamanho de populações de patos selvagens na América do Norte.

obtendo-se assim intuitivamente um estimador do número de indivíduos na população.

O modelo de Lincoln-Petersen assume igual probabilidade de captura para os indivíduos em cada amostra; as marcas não se perdem e na amostragem os indivíduos são contados inequivocamente (White *et al.*, 1982).

O primeiro pressuposto acima mencionado nem sempre é válido, pois a probabilidade de captura varia: entre os diferentes momentos de amostragem, devido a factores externos, tais como condições ambientais (tempo - t); entre os indivíduos, devido às suas características inerentes na probabilidade de captura, tais como sexo, idade, tamanho ou grupo social dominante (heterogeneidade - h); e depois da primeira captura, porque os indivíduos podem reagir de forma diferente nas capturas subsequentes com comportamento do tipo *trap happy* ou *trap shy* (comportamento - b)<sup>4</sup>.

A combinação desses três factores dá um conjunto de oito possíveis modelos:  $\mathcal{M}_0$ ,  $\mathcal{M}_t$ ,  $\mathcal{M}_b$ ,  $\mathcal{M}_h$ ,  $\mathcal{M}_{tb}$ ,  $\mathcal{M}_{th}$ ,  $\mathcal{M}_{bh}$  e  $\mathcal{M}_{tbh}$  (Otis *et al.*, 1978). O modelo mais simples é o  $\mathcal{M}_0$ , que considera igual probabilidade de captura para todos os indivíduos na população, e para todos os momentos de amostragem da captura. O primeiro desenvolvimento do modelo  $\mathcal{M}_h$  foi realizado por Burnham e Overton (1978), utilizando o método de *jackknife*. Mais tarde, Cormack (1989) tentou estimar os parâmetros do modelo heterogéneo ( $\mathcal{M}_h$ ) com a metodologia loglinear, e Norris e Pollock (1996) e Pledger (2000) utilizaram o método de máxima verosimilhança e modelos de misturas. Os estimadores do modelo tipo  $\mathcal{M}_t$  foram desenvolvidos por Chao (1989) e Schnabel (1938) e, mais tarde, Chao *et al.* (2000) propuseram os estimadores para o modelo do tipo  $\mathcal{M}_{th}$ . O modelo que varia com o tempo e considera a heterogenei-

---

<sup>4</sup>A sigla *b* advém da palavra inglesa *behaviour*.

dade ( $\mathcal{M}_{th}$ ) foi desenvolvido por Alho (1990) e Huggins (1989), utilizando técnicas de regressão logística. Outros autores têm utilizado inferência bayesiana (IB) para estimar os parâmetros dos modelos do tipo:  $\mathcal{M}_t$  (Castledine, 1981; George e Robert, 1992; Smith, 1988, 1991),  $\mathcal{M}_h$  (Tardella, 2002),  $\mathcal{M}_{tb}$  (Lee e Chen, 1998) e  $\mathcal{M}_{bh}$  (Ghosh e Norris, 2005).

O modelo de captura-recaptura não reduz o enviesamento devido a disponibilidade, mas pode diminuir o enviesamento causado pela percepção. Contudo, se a heterogeneidade não for modelada adequadamente, o modelo de captura-recaptura pode introduzir ainda mais enviesamento que o modelo de transecto linear (Laake e Borchers, 2004). A heterogeneidade observável nas probabilidades de detecção tem sido modelada utilizando modelos de dupla plataforma, que combinam os modelos de transecto linear e captura-recaptura, incorporando outras covariáveis para além da distância (Borchers *et al.*, 1998a,b). Por outro lado, o modelo proposto por Borchers *et al.* (1998a,b) não considera o problema da heterogeneidade não modelada, relativo às características inerentes aos indivíduos, que são difíceis de medir ou observar. Borchers *et al.* (2006) utilizaram o modelo de dupla plataforma para modelar heterogeneidade não observável.

## 1.4 Modelo combinado captura-recaptura e transectos lineares

O modelo combinado tem sido estudado por vários autores para estimar a probabilidade de detectar um animal na linha do transecto (e.g., Alpizar-Jara e Pollock, 1996, 1999; Borchers *et al.*, 1998a,b; Buckland e Turnock, 1992; Chen, 2000; Manly *et al.*, 1996; Okamura *et al.*, 2003; Palka, 1995; Quang e Becker, 1997; Schweder

*et al.*, 1999). Um modelo mais geral, que considera detecção de animais por mais do que um observador, foi considerado por Alpizar-Jara e Pollock (1996). Estes autores consideraram um modelo que combina a informação das distâncias obtidas através da amostragem com as detecções dos animais pelos observadores como analogia às múltiplas capturas no modelo de captura-recaptura. Este modelo é considerado mais realístico, pois toma em conta que as probabilidades de detecção variam de observador para observador e permite a violação do pressuposto chave na CDS.

O modelo combinado pode envolver muitos parâmetros, o que dificulta a análise sob o ponto de vista frequentista e analítico. Neste sentido, parece natural apresentar uma abordagem bayesiana para estimar os muitos parâmetros desse modelo.

## 1.5 O método bayesiano e o algoritmo MCMC

### 1.5.1 Introdução

A inferência bayesiana<sup>5</sup> permite incorporar informação *a priori* nas variáveis não observáveis, e a distribuição *a posteriori* dá uma interpretação probabilística das variáveis chave, o que faz com que esta abordagem seja eficiente na resolução de problemas complexos.

Os métodos bayesianos podem ser interpretados como uma extensão da estatística clássica. Enquanto que, na inferência clássica, o parâmetro  $\theta$  (ou o vector parâmetro  $\boldsymbol{\theta}$ ) assume-se desconhecido, mas fixo, na inferência bayesiana este passa

---

<sup>5</sup>A filosofia Bayesiana foi desenvolvida pelo reverendo inglês Thomas Bayes (1702-1761), no final do século XVIII. Ele foi o primeiro matemático a usar probabilidade indutivamente e a estabelecer as bases matemáticas para inferência probabilística. As descobertas em probabilidade foram publicadas postumamente, no artigo intitulado *Essay Towards Solving a Problem in the Doctrine of Chances* (Bayes, 1763), no *Philosophical Transactions of the Royal Society* de Londres.

a ser uma variável aleatória não observável.

Para os bayesianos, qualquer informação *a priori* externa em relação à amostra ou experiência é muito importante para ser ignorada e deve ser quantificada em termos de probabilidade, que é expressa por uma distribuição de probabilidades,  $h(\theta)$ , designada por distribuição *a priori*.

Geralmente, a distribuição de probabilidade *a priori* é interpretada como uma medida do grau de credibilidade que uma dada pessoa, que procede à análise, atribui a um determinado parâmetro  $\theta$ . No caso contínuo, tem-se a função de densidade *a priori*, e  $h(\theta)d\theta$  entende-se como sendo o grau de credibilidade que se atribui ao intervalo  $(\theta, \theta + d\theta)$ ,  $\theta \in \mathbb{R}$ .

Para além da informação anterior à experiência, considera-se ainda a informação por amostragem,  $f(y|\theta)$ , obtida através de experiências realizadas em condições idênticas. É de notar que na inferência clássica, toma-se em conta apenas a informação por amostragem, desprezando assim a informação *a priori*.

### 1.5.2 Teorema de Bayes

O teorema de Bayes, a seguir enunciado, possibilita a actualização das credibilidades iniciais tendo em conta a informação por amostragem.

**Teorema 1.1** (Bayes). *Sejam  $A_1, A_2, \dots$  uma partição do conjunto  $A$ , onde  $A$  é o universo e  $B$  é um conjunto arbitrário. Então para qualquer  $i = 1, 2, \dots$  vem que:*

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\infty} P(B|A_j)P(A_j)}, \quad (1.1)$$

*dado que  $P(B|A_i), P(A_i)$  existem e  $\sum_{j=1}^{\infty} P(B|A_j)P(A_j) \neq 0$ .*

Seja  $y$  uma realização particular da variável aleatória  $Y$ ,  $f(y|\theta)$  um elemento da família  $\mathcal{F}$  tal que:

$$\mathcal{F} = \{f(y|\theta) : \theta \in \Theta\},$$

onde o espaço amostral  $\mathcal{Y}$ ,  $y \in \mathcal{Y}$ , é o conjunto de amostras possíveis.

Tendo em conta a distribuição *a priori*  $h(\theta)$ , dado o Teorema de Bayes, (1.1), resulta que a distribuição *a posteriori* para a variável contínua  $\theta$ ,  $h(\theta|y)$ , é expressa por:

$$h(\theta|y) = \frac{f(y|\theta)h(\theta)}{\int_{\Theta} f(y|\theta)h(\theta)d\theta} \propto f(y|\theta)h(\theta), \quad (1.2)$$

onde  $\propto$  exprime a proporcionalidade. É de referir que a relação de proporcionalidade expressa em (1.2) faz sentido, uma vez que o denominador,  $\int_{\Theta} f(y|\theta)h(\theta)d\theta$ , não depende de  $\theta$ .

### 1.5.3 *Priori e posteriori*

No contexto bayesiano, a distribuição *a posteriori* é o elemento fulcral, pois constitui a base da inferência estatística. Depois da escolha da distribuição *a priori* e da actualização por amostragem, obtém-se a distribuição *a posteriori* através da equação (1.2).

Quando a informação *a priori* é inexistente, ou então, o conhecimento efectivo sobre  $\theta$  é excessivamente vago, a distribuição *a priori* para o parâmetro será minimamente informativa e, neste caso, é denominada por distribuição *a priori* não informativa.

Uma distribuição *a priori* não informativa foi proposta por Bayes (1763), que considerou a distribuição Uniforme como sendo a distribuição *a priori* num cenário binomial, e Laplace (1812), que usou uma distribuição *a priori* Uniforme imprópria para o caso da Normal. A ideia é que quando há pouca informação *a priori* deve-se adoptar a equiprobabilidade para os acontecimentos.

Se  $\Theta$  for finito, i.e.  $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$ , não há nenhum problema com a distribuição *a priori* Uniforme, e resulta que:

$$h(\theta) = \frac{1}{k}, \quad \theta = \theta_1, \theta_2, \dots, \theta_k.$$

Contudo, se  $\Theta$  for infinito numerável, deixa de existir uma distribuição Uniforme em  $\Theta$ . Neste caso,  $h(\theta)$  é designada por distribuição *a priori* imprópria. Se  $\Theta$  for infinito não numerável e ilimitado, também não existe nenhuma distribuição Uniforme em  $\Theta$ , uma vez que  $\int_{\theta} h(\theta) d\theta$  é infinito, não satisfazendo o axioma de probabilidade total unitária.

Entretanto, nem sempre a distribuição Uniforme é invariante a transformações injectivas na representação formal da “ignorância”. Assim, para resolver este problema, Jeffreys (1946, 1961) propôs para a distribuição *a priori* o uso da medida de informação de Fisher para  $\theta \in \mathbb{R}$ ,  $I(\theta)$ :

$$h(\theta) \propto \sqrt{|I(\theta)|}, \tag{1.3}$$

onde:

$$I(\theta) = E \left[ - \frac{\partial^2 \log f(Y|\theta)}{\partial \theta^2} \Big|_{\theta} \right], \quad \theta \in \mathbb{R} \text{ (Lehmann e Casella, 1998).}$$

Jeffreys sugere que a distribuição *a priori* mais plausível para  $\theta$  é aquela que apresenta maior informação de Fisher, predominando neste caso a informação por amostragem face à informação *a priori*, daí a denominação de distribuição *a priori* não informativa.

#### 1.5.4 Modelos hierárquicos

Nos modelos com muitos parâmetros é comum a adoção de modelos hierárquicos. Assim, a incerteza nos hiperparâmetros<sup>6</sup> é modelada assumindo vários níveis para a distribuição *a priori*. Os modelos hierárquicos têm sido utilizados nos métodos de captura-recaptura, como por exemplo, os de Johnson e Hoeting (2003) e Link e Barker (2005), uma vez que os modelos podem ser expressos de maneira natural pelos ecologistas. Também permitem que as distribuições *a posteriori* complexas sejam subdivididas em distribuições condicionais, que são mais fáceis de trabalhar (Royle e Dorazio, 2008).

Geralmente, os modelos bayesianos hierárquicos apresentam apenas dois níveis. A distribuição *a priori* do primeiro nível é quase sempre uma distribuição conjugada<sup>7</sup> do modelo amostral, quando tal é possível, com o intuito de facilitar a implementação

---

<sup>6</sup>São parâmetros que, por definição, dependem de outros parâmetros diferentes, que, por sua vez, podem depender ou não de outros.

<sup>7</sup>Uma distribuição *a priori* é dita conjugada com a função de verossimilhança se a distribuição *a posteriori* resultante é da mesma família de distribuições da *a priori*.

computacional. Entretanto, a distribuição do último nível é muitas vezes de caráter não informativo, uma vez que se torna muito difícil descrever os hiperparâmetros quando o nível é superior a dois.

Os modelos hierárquicos são representados por grafos dirigidos acíclicos de fácil percepção, designados por DAG (*Directed Acyclic Graphs*), que descrevem os parâmetros através dos nós e setas que explicam as probabilidades condicionadas. Por exemplo, o modelo hierárquico Poisson-Gama para o tamanho da população tem dois nós  $N$  e  $\lambda$ , onde a seta que faz a ligação entre eles representa a relação condicional  $N|\lambda$  como sendo a distribuição Poisson com média  $\lambda$ ,  $N|\lambda \sim Poi(\lambda)$  e  $\lambda \sim Gama(\tau, \eta)$ , onde  $N = 1, 2, \dots$ ,  $\lambda > 0$ ,  $\tau, \eta \in \mathbb{R}$ .

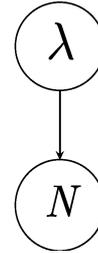


Figura 1.1: Exemplo de um grafo dirigido acíclico (DAG) do modelo hierárquico Poisson-Gama para o tamanho da população ( $N$ ),  $N|\lambda \sim Poi(\lambda)$  e  $\lambda \sim Gama(\tau, \eta)$ , onde  $N = 1, 2, \dots$ ,  $\lambda > 0$ ,  $\tau, \eta \in \mathbb{R}$ .

Na Figura 1.1 apresenta-se a relação estocástica  $N|\lambda$  e diz-se que  $\lambda$  é “parente” de  $N$ . Esta relação é também “acíclica”, no sentido de que não há outro caminho possível através de um nó diferente. O procedimento da inferência bayesiana reside então em especificar as distribuições *a priori* para os parâmetros que não têm “parentes”, neste caso particular seria o parâmetro  $\lambda$ .

### 1.5.5 MCMC

As primeiras tentativas de aplicação da inferência bayesiana para estimar parâmetros de populações fechadas têm sido desafiadas por causa da adoção de distribuições *a priori* pouco realísticas e das dificuldades computacionais encontradas no cálculo da distribuição *a posteriori*. Entretanto, grande parte das aplicações reais utilizando os modelos de captura-recaptura, transecto lineares ou o modelo combinado têm distribuições *a posteriori* que implicam trabalhar com integrais múltiplos com dimensão por vezes muito elevadas, só se tornando assim possível obter inferências de forma directa para um conjunto restrito de parâmetros. Não obstante, a distribuição *a posteriori* completa (ou as correspondentes distribuições marginais de cada parâmetro) proporciona estimativas pontuais e intervalos de credibilidade para os parâmetros.

O processo da inferência é feito através da distribuição *a posteriori*, traduzindo naquilo que é a fonte da complexidade computacional da abordagem bayesiana. As estimativas do vector de parâmetros,  $\theta$ , requerem que se condicionem os outros parâmetros. O integral de dimensão elevada associado à distribuição *a posteriori* é estimado através da construção da cadeia de Markov, cuja distribuição estacionária é igual à distribuição *a posteriori* de interesse. Todo esse seguimento é realizado mediante a utilização dos métodos das cadeias de Markov baseado em Monte Carlo (MCMC<sup>8</sup>, Gilks *et al.*, 1996), que utiliza o amostrador de Gibbs (Gelfand e Smith, 1990) ou o algoritmo Metropolis-Hastings (Hastings, 1970; Metropolis *et al.*, 1953) no cálculo de integrais de difícil resolução analítica.

É de notar que os avanços computacionais e o desenvolvimento de programas como o WINBUGS (*Bayesian inference Using Gibbs Sampling*, Spiegelhalter *et al.*,

---

<sup>8</sup>A abreviatura MCMC advém de Markov Chain Monte Carlo.

2004), de distribuição gratuita (o programa pode ser obtido em [www.mrc-bsu.cam.ac.uk/bugs](http://www.mrc-bsu.cam.ac.uk/bugs)) e de fácil implementação, que utiliza os métodos MCMC para resolver integrais complexos, fizeram com que a inferência bayesiana se tivesse tornado numa ferramenta muito útil para estimar a abundância animal (King e Brooks, 2002). O WINBUGS pode ser utilizado no programa R (R Development Core Team, 2006), também de acesso livre, através do pacote R2WINBUGS (Sturtz *et al.*, 2005).

No programa WINBUGS existe uma função denominada por *Doodle*, que permite representar um modelo em forma de DAG. Nesta representação gráfica, as elipses são quantidades não observáveis, os rectângulos pequenos, as constantes e os rectângulos maiores, os ciclos. As setas simples representam dependências estocásticas e as setas duplas representam funções lógicas.

## 1.6 O modelo combinado e a inferência bayesiana

Apesar de tudo, ainda existem alguns problemas a serem resolvidos, relacionados com a aplicação do método bayesiano para estimar os parâmetros em modelos para populações fechadas, tais como, a não identifiabilidade e a não estimabilidade de parâmetros (Seber e Schwarz, 2002).

No contexto dos modelos de captura-recaptura, apenas há cerca de poucos anos é que a abordagem bayesiana foi parcialmente proposta para resolver os problemas de identifiabilidade e de estimabilidade (Holzmann *et al.*, 2006; Huggins, 2002; Lee e Chen, 1998; Link, 2003; Pledger, 2000; Tardella, 2002; Wang, 2002). A metodologia bayesiana também foi proposta por Karunamuni e Quinn (1995), de forma independente, para estimar os parâmetros do modelo de transecto linear. Recente-

mente Eguchi e Gerrodette (2009) estenderam esse modelo, incluindo uma função de máxima verosimilhança binomial para o número de objectos detectados.

Quando se utiliza o MCMC para fazer sumário da distribuição *a posteriori* ou calcular qualquer que seja a quantidade de interesse, há que ter em conta duas questões no que toca à simulação das cadeias de Markov. Primeiro, há que saber se a cadeia de Markov atingiu a estacionaridade, ou a distribuição *a posteriori* desejada. Em segundo lugar, há que determinar o número de iterações necessárias para que a cadeia de Markov atinga a estacionaridade.

Note-se que muitas das ferramentas de diagnósticos são utilizadas para verificar uma condição de convergência necessária mas não suficiente. Não existem testes conclusivos que possam dizer quando a cadeia de Markov convergiu para a distribuição estacionária. Sendo assim, é recomendável a utilização de mais do que um teste, e a convergência deve ser verificada para todos os parâmetros, antes de se proceder a qualquer inferência.

Nesta tese, para análise de diagnóstico de convergência foram usadas duas cadeias com valores iniciais diferentes, um gerado pelo WINBUGS e o outro como sendo o valor esperado da distribuição *a priori* correspondente a cada um dos parâmetros. A convergência das cadeias foi verificada, utilizando entre outros, o método de convergência de Gelman e Rubin (1992), que estão disponíveis no próprio WINBUGS ou podem ser feitas usando as rotinas de programa R, nomeadamente CODA (*Convergence Diagnosis and Output Analysis*, Plummer *et al.*, 2006) ou BOA (*Bayesian Output Analysis*, Smith, 2005), que se baseia em análise de múltiplas cadeias simuladas através de MCMC por comparação das variâncias dentro de cada cadeia e entre as cadeias. Uma grande discrepância entre as duas variâncias indica que não

houve convergência e, possivelmente, estaríamos na presença de distribuições *a posteriori* com múltiplas modas (cadeias diferentes convergem para diferentes modas) ou então há necessidade de cadeias com mais iterações (o período de aquecimento não foi atingido).

A inferência bayesiana é realizada usando o programa WINBUGS e os intervalos HPD (*highest posterior density*) de 95% para os parâmetros de interesse são obtidos com recurso ao programa BOA implementado no programa R. As análises utilizando o método de máxima verosimilhança para o modelo de transectos lineares foram realizadas programa DISTANCE (Thomas *et al.*, 2006).

Para analisar os efeitos da especificação das distribuições *a priori* sobre os resultados, conduziu-se uma análise de sensibilidade, onde estimativas bayesianas foram obtidas tendo em conta diferentes possibilidades para a distribuição *a priori* do parâmetro considerado chave em cada modelo analisado nos respectivos capítulos.

## 1.7 Estrutura da tese

O modelo combinado ainda não foi abordado sob o ponto de vista da inferência bayesiana, e esta tese pretende dar um contributo nesse sentido, desenvolvendo um modelo combinado generalizado de captura-recaptura e transectos lineares, que toma em conta a heterogeneidade nas probabilidades de detecção relativa às características inerentes nos indivíduos (*MCHG\**), que são difíceis de medir ou não observáveis. O desempenho do modelo *MCHG\** e dos submodelos (Figura 1.2), são analisados e comparados mediante simulações e também aplicados utilizando dados reais.

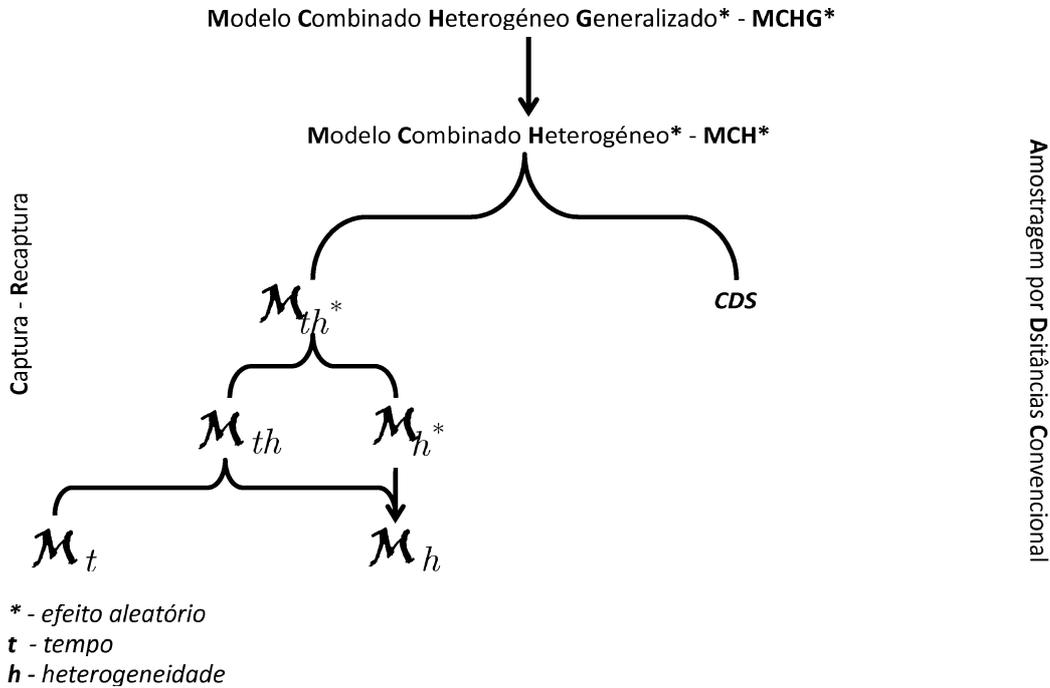


Figura 1.2: Diagrama do modelo combinado heterogêneo generalizado ( $MCHG^*$ ) e dos submodelos subsequentes: amostragem por distâncias convencional ( $CDS$ ); modelos de captura-recaptura quando as probabilidades de captura: são constantes ( $\mathcal{M}_0$ ), variam com o tempo ( $\mathcal{M}_t$ ), dependem de características individuais ( $\mathcal{M}_h$ ), dependem de características individuais e de outros factores não observáveis ( $\mathcal{M}_{h^*}$ ), variam com o tempo e dependem de características individuais ( $\mathcal{M}_{th}$ ), variam com o tempo e dependem de características individuais e de outros factores não observáveis ( $\mathcal{M}_{th^*}$ ) e modelo combinado de captura-recaptura e transectos lineares, quando as probabilidades de avistar dependem de características individuais e de outros factores não observáveis ( $MCH^*$ ).

Esta tese está organizada da seguinte maneira: cada capítulo pode ser visto de forma independente, mas o Capítulo 4 é também uma integração dos modelos apresentados no Capítulo 2 e 3. Os modelos de captura-recaptura que retratam o problema de heterogeneidade observável, utilizando covariáveis e a heterogeneidade não observável, são descritos no Capítulo 2. Comparam-se os estimadores dos

modelos  $\mathcal{M}_t$  e as variantes dos modelos que incorporam heterogeneidade, quando se modela apenas heterogeneidade observável ( $\mathcal{M}_h$  e  $\mathcal{M}_{th}$ ) e quando é modelada a heterogeneidade não observável,  $\mathcal{M}_{h^*}$  e  $\mathcal{M}_{th^*}$ , respectivamente. No Capítulo 3, apresenta-se uma abordagem bayesiana do modelo combinado, que considera apenas dois momentos de amostragem e não toma em conta a heterogeneidade nas probabilidades de captura. No Capítulo 4, destaca-se o modelo combinado que incorpora heterogeneidade (observável e não observável) na probabilidade de detecção na linha do transecto. Este modelo é generalizado para quando existem múltiplos observadores. Finalmente, no Capítulo 5 apresentam-se as conclusões.

Nos Apêndices A.1, A.2, A.3, A.4 e A.5 podem ser consultados os respectivos códigos de WINBUGS dos modelos apresentados nesta tese, que também podem ser acedidos no sítio <http://home.uevora.pt/~jfgm/codigos.htm>.

## Capítulo 2

# Modelando heterogeneidade não observável em probabilidades de captura-recaptura

Parte deste capítulo encontra-se sob a forma de publicação:

Monteiro, J. F. G. e Alpizar-Jara, R. (2008). Modelando heterogeneidade não observável em probabilidades de captura-recaptura. Em *Estatística da Teoria à Prática - Actas do XV Congresso Anual da SPE*. (Hill, M. M., Ferreira, M. A., Dias, J. G., Salgueiro, M. F., Carvalho, H., Vicente, P. e Braumann, C. eds). Sociedade Portuguesa de Estatística, Lisboa, 365-374 pp.

## 2.1 Introdução

Neste capítulo apresenta-se um modelo de captura-recaptura para estimar o tamanho de populações fechadas. Geralmente, considera-se que o tempo de amostragem dos animais é relativamente curto, de modo a assegurar que o pressuposto de oclusão demográfica seja satisfeito. Pollock (1974) foi o primeiro a reflectir sobre o pressuposto de as probabilidades de capturas serem iguais nos modelos de captura-recaptura, apresentando um conjunto de modelos, que são analisados em detalhe por Otis *et al.* (1978) e White *et al.* (1982).

Um estimador para o tamanho da população pode ser enviesado se na modelação não for considerada a heterogeneidade nas probabilidades de captura, relativa às características inerentes dos indivíduos, que são tipicamente difíceis de medir ou não observáveis. O problema da modelagem da heterogeneidade não observável nas probabilidades de captura tem sido para muitos estatísticos um desafio sob o ponto de vista teórico. A primeira abordagem *ad hoc* baseou-se no método de *jackknife* para corrigir o enviesamento, desenvolvendo-se um estimador para o tamanho da população em função das frequências de capturas (Burnham e Overton, 1978, 1979; Pollock e Otto, 1983, para modelos do tipo  $\mathcal{M}_h$  e  $\mathcal{M}_{bh}$ , respectivamente). Outros métodos não paramétricos, tais como *bootstrap* e estimadores de máxima verosimilhança não paramétrico, foram utilizados respectivamente por Smith e van Belle (1984) e Norris e Pollock (1995, 1996).

Pollock *et al.* (1984) consideraram nas análises para modelar as probabilidades de captura um modelo logístico que incluía covariáveis individuais, como por exemplo o tamanho do corpo ou o peso do animal. Contudo, nos modelos logísticos para

as probabilidades de captura, utilizando variáveis individuais, apesar de terem o propósito de considerar a heterogeneidade, alguma heterogeneidade inerente pode ficar por explicar devido a outras variáveis não observáveis (Alho, 1990; Borchers *et al.*, 1998b; Chao, 2001; Huggins, 1989, 1991; Pledger, 2000; Pollock *et al.*, 1984).

Chao (1989) apresentou vários estimadores para modelos do tipo  $\mathcal{M}_t$ , que mais tarde foram generalizados para o modelo do tipo  $\mathcal{M}_{th}$  (Chao *et al.*, 2000). Os autores Chao *et al.* (1992), Lee e Chao (1994) e Chao (2001) sugeriram modelos, do tipo  $\mathcal{M}_{bh}$ , que utilizam o método que se baseia em cobertura amostral.

Os modelos log-lineares para tabelas de contingência incompletas têm sido aplicados para estimar os parâmetros do modelo que considera heterogeneidade nas probabilidades de captura, ver por exemplo (Agresti, 1994; Bishop *et al.*, 1975; Cormack, 1989; Coull e Agresti, 1999; Fienberg, 1972; Rivest e Daigle, 2004). Outros autores sugeriram o uso de martingalas (Becker, 1984; Lloyd, 1994; Lloyd e Yip, 1991; Yip, 1989, 1991; Yip e Fong, 1993).

Pledger (2000), baseando-se nos desenvolvimentos feitos por Norris e Pollock (1995, 1996), utilizou distribuições de misturas finitas e introduziu heterogeneidade nas probabilidades de captura através do método de máxima verossimilhança.

Outras alternativas têm sido desenvolvidas, para casos onde a função de verossimilhança pode ser calculada. Como por exemplo, o modelo de Rasch (Coull e Agresti, 1999; Bartolucci e Forcina, 2001; Fienberg *et al.*, 1999),  $\mathcal{M}_{th}$ , o uso de covariáveis para incorporar heterogeneidade (Huggins, 1991; Stanghellini e van der Heijden, 2004),  $\mathcal{M}_{tth}$  e  $\mathcal{M}_{th}$ , os modelos de mistura utilizando distribuição Beta-Binomial (Dorazio e Royle, 2003),  $\mathcal{M}_{th}$ , os modelos de mistura utilizando distribuições Binomial e Beta-Binomial (Morgan e Ridout, 2007),  $\mathcal{M}_{th}$ .

De forma independente, os autores Alho (1990) e Huggins (1989, 1991) propuseram um modelo logístico, utilizando funções de verosimilhança condicionada aos históricos das capturas dos animais observados para modelar as probabilidades individuais de captura, em termos de características observáveis utilizando covariáveis em modelos de captura-recaptura do tipo  $\mathcal{M}_{th}$  para populações fechadas. Eles utilizaram um estimador generalizado do tipo Horvitz-Thompson<sup>9</sup> (H-T), para estimar o tamanho da população. Essa abordagem tem a vantagem de ser simples, uma vez que não há necessidade de especificar distribuições para as covariáveis, e a desvantagem de não utilizar a função de verosimilhança completa para estimar o tamanho da população. Huggins (1991) apresentou 8 diferentes modelos, desde o caso mais simples do tipo  $\mathcal{M}_0$  até um caso particular de  $\mathcal{M}_{tbh}$ .

Os estudos utilizando a inferência bayesiana têm sido mais para modelos que permitem apenas a variação de probabilidades entre os momentos de amostragens,  $\mathcal{M}_t$  (Castledine, 1981; George e Robert, 1992) ou quando abordam o problema da heterogeneidade não observável,  $\mathcal{M}_{h^*}$  (Basu e Ebrahimi, 2001). Ghosh e Norris (2005) desenvolveram um modelo utilizando uma abordagem bayesiana, onde as probabilidades de captura não variam com o tempo, mas variam de indivíduo para indivíduo, e dependem da reação dos mesmos à armadilha,  $\mathcal{M}_{bh}$ . Modelos do tipo  $\mathcal{M}_{th}$  com abordagem bayesiana foram analisados por Durban e Elston (2005), Fienberg *et al.* (1999) e Royle *et al.* (2007). Note-se que Link (2003) considerou o problema da não identificabilidade quando há heterogeneidade relacionada com a estimação do tamanho da população.

---

<sup>9</sup>O estimador do tipo Horvitz-Thompson (1952) trata-se de um estimador não enviesado da população total de uma população finita, aplicável de uma forma geral, nos casos onde os indivíduos são amostrados considerando que as probabilidades de inclusão são diferentes.

Outros autores também propuseram abordagem bayesiana para modelos de captura-recaptura, entre os quais Ananda (1997), Castledine (1981), George e Robert (1992), Hunter e Griffiths (1978), King e Brooks (2001), Leite *et al.* (2000), Madigan e York (1997), Mingoti (2000), Smith (1988, 1991), Wang (2002) e Yoshida *et al.* (1999).

Grande parte dos modelos clássicos para estimar o tamanho da população, acima referidos, podem ser analisados usando o programa de acesso gratuito MARK (White e Burnham, 1999), que permite reproduzir os modelos que são tratados no programa CAPTURE (Rexstad e Burnham, 1991), no qual é possível estimar os parâmetros associados aos modelos clássicos para populações fechadas, tendo em conta os diferentes modelos propostos por Otis *et al.* (1978). O desenvolvimento de novas metodologias, os avanços recentes a nível computacional e a disponibilidade de programas flexíveis, como por exemplo o programa WINBUGS, proporcionaram nas últimas décadas o aumento da aplicação dos métodos bayesianos na área da ecologia de populações (Clark, 2005; Ellison, 2004; Gimenez *et al.*, 2008; McCarthy, 2007).

Neste capítulo considera-se um modelo de captura-recaptura do tipo  $\mathcal{M}_{th}$  para populações fechadas onde as probabilidades de captura variam entre os diferentes momentos de amostragem e dependem das características inerentes a cada indivíduo, que podem ser observáveis, como por exemplo o sexo, a idade e o peso do animal, como as que são difíceis de observar ou de medir. Neste trabalho, utilizou-se uma abordagem híbrida, onde a heterogeneidade observável nas probabilidades de captura dos indivíduos na população é modelada utilizando a técnica de regressão logística, cuja abordagem foi proposta por Alho (1990) e Huggins (1989, 1991). En-

quanto que a heterogeneidade não observável será modelada utilizando a inferência bayesiana, introduzindo um efeito aleatório referente a cada indivíduo, ao qual se atribui uma distribuição *a priori*. Os resultados são ilustrados com um exemplo, utilizando capturas de ratos (*Peromyscus maniculatus*) realizadas por V. Reid, onde se comparam os estimadores do modelo de Huggins (1989, 1991), através de análises realizadas utilizando a abordagem clássica e a bayesiana.

## 2.2 Modelo captura-recaptura tipo $\mathcal{M}_{th}$

Os dados da captura são geralmente apresentados em forma de histórico da captura para cada animal observado individualmente na área de estudo (Williams *et al.*, 2002). O histórico da captura de um indivíduo é uma sequência simples de 0 e 1.

Seja  $i = 1, \dots, N$  o número de indivíduo na população que podem ser capturados nos  $j$  ( $j = 1, \dots, t$ ) momentos de amostragem com probabilidade,  $p_{ij}$ , e seja

$$x_{ij} = \begin{cases} 1 & \text{se indivíduo } i \text{ foi capturado no momento } j \\ 0 & \text{c.c.} \end{cases} .$$

A matriz  $\mathbf{x}$  contém a informação sobre o histórico das capturas e  $\mathbf{p}$  é a matriz dos  $p_{ij}$  para  $i = 1, \dots, N$  e  $j = 1, \dots, t$ . Assumindo que os indivíduos se comportam independentemente uns dos outros e que as capturas são realizadas de modo independente, a função de verosimilhança completa para o modelo de captura-recaptura (Otis *et al.*, 1978) é dada por:

$$f(\mathbf{x}|N, \mathbf{p}) \propto \prod_{i=1}^N \prod_{j=1}^t p_{ij}^{x_{ij}} (1 - p_{ij})^{(1-x_{ij})}. \quad (2.1)$$

Considerando-se  $i = 1, \dots, n$  o número de indivíduos capturados e  $i = n + 1, \dots, N$  os indivíduos que não foram capturados na população no momento  $j$ , ao igual que Huggins (1989) define-se  $\gamma_{ij}$  como sendo a probabilidade de o indivíduo  $i$  ser capturado no momento  $j$  dado o histórico da captura e sabendo que foi capturado pelo menos uma vez, por:

$$\gamma_{ij} = \frac{p_{ij}}{1 - (1 - c_{ij}) \prod_{l=j}^t (1 - p_{il}^*)}, \quad (2.2)$$

onde  $p_{ij}^*$  é igual à probabilidade do indivíduo  $i$  ser capturado no momento  $j$  quando  $c_{ij} = 0$  e

$$c_{ij} = \begin{cases} 1 & \text{se indivíduo } i \text{ foi capturado antes do momento } j \\ 0 & \text{c.c.} \end{cases}.$$

Condicionando a função de verosimilhança da equação (2.1) aos indivíduos que foram capturados pelo menos uma vez e tendo em conta (2.2), a função de verosimilhança condicional é expressa por:

$$f(\mathbf{x}|\mathbf{p}) \propto \prod_{i=1}^n \prod_{j=1}^t \gamma_{ij}^{x_{ij}} (1 - \gamma_{ij})^{(1-x_{ij})} = \prod_{i=1}^n \prod_{j=1}^t \frac{p_{ij}^{x_{ij}} (1 - p_{ij})^{(1-x_{ij})}}{1 - \prod_{l=1}^t (1 - p_{il}^*)}. \quad (2.3)$$

Apresenta-se a seguir o modelo  $\mathcal{M}_{th}$  como sendo um caso particular do modelo  $\mathcal{M}_{thh}$  proposto por Huggins (1991) utilizando uma transformação logística linear

para modelar os  $p_{ij}$ :

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j}, \quad (2.4)$$

com  $i = 1, \dots, n$ ,  $j = 1, \dots, t$  e note-se que  $\beta_{\text{tempo}_t} = 0$ .

Na equação (2.4),  $\beta_{\text{sexo}}$  e  $\text{sexo}_i$  significam, respectivamente, o efeito sexo e o género do animal. O efeito idade e a idade de cada indivíduo são representados por  $\beta_{\text{idade}}$  e  $\text{idade}_i$ ; o efeito por unidade de peso e peso do indivíduo  $i$  são expressos por  $\beta_{\text{peso}}$  e  $\text{peso}_i$ . Portanto,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$  e  $\beta_{\text{peso}}$  explicam a variação de  $p_{ij}$  entre os indivíduos. O efeito da variação do tempo,  $\beta_{\text{tempo}_j}$ , explica a variação da probabilidade de captura entre os momentos de amostragem.

O estimador para o tamanho da população proposto por Huggins (1989, 1991) é do tipo Horvitz-Thompson, de tal forma que:

$$\hat{N} = \sum_{i=1}^n \frac{1}{\hat{p}_i}, \hat{p}_i > 0,$$

onde

$$\hat{p}_i = 1 - \prod_{j=1}^t (1 - \hat{p}_{ij}^*)$$

é a estimativa da probabilidade do animal  $i$  ser capturado 1 ou mais vezes, ou seja 1 menos a probabilidade de nunca vir a ser capturado.

## 2.3 Modelando a heterogeneidade não observável

Na prática, é comum em muitas aplicações seleccionar múltiplas medições sobre um indivíduo. Os modelos lineares mistos (Laird e Ware, 1982; Longford, 1993) têm por objectivo considerar dependências entre os indivíduos e nas múltiplas medições, introduzindo um ou mais variáveis latentes específicas (i.e, efeitos aleatórios) no modelo de regressão.

O uso de métodos bayesianos é, sobretudo, sugerido quando há necessidade de incorporar covariáveis não observáveis. Os métodos bayesianos automaticamente integram variáveis aleatórias não observadas utilizando integração numérica ou métodos de amostragens MCMC (Pollock, 2002).

A heterogeneidade está sempre presente nos animais, pois alguns são mais fáceis de serem capturados que outros. O modelo  $\mathcal{M}_{th}$  (2.4) apenas leva em conta a heterogeneidade que seja observável ou mensurável, ficando por explicar e incorporar factores desconhecidos ou não mensuráveis, que afectam as probabilidades individuais de capturas. Ora, se a heterogeneidade que não é mensurável for ignorada, os animais capturados poderão ser aqueles com maior tendência para serem capturados, e, neste caso, a probabilidade de captura individual será sobrestimada. Portanto, o tamanho da população será subestimado, uma vez que as probabilidades individuais dos animais capturados correspondem ao denominador nos estimadores do tipo H-T.

A heterogeneidade não observável inerente ao indivíduo  $i$  é aqui explicada por um efeito aleatório  $\delta_i$ , a qual se atribuirá uma distribuição *a priori*. Tendo em conta (2.4), o modelo que incorpora heterogeneidade não observável, que se designa por

$\mathcal{M}_{th^*}$ , define-se como sendo:

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i. \quad (2.5)$$

O modelo proposto em (2.5) coincide com uma reparametrização do modelo  $\mathcal{M}_0$  de Otis *et al.* (1978), quando se tem apenas  $\beta_0$  no segundo membro da equação;  $\mathcal{M}_t$  para  $\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_{\text{tempo}_j}$ ; um caso particular de  $\mathcal{M}_h$  se  $\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i$ . Todos estes submodelos, nomeadamente  $\mathcal{M}_t$ ,  $\mathcal{M}_h$  e  $\mathcal{M}_{th}$ , e o modelo  $\mathcal{M}_{th^*}$  serão analisados através de uma abordagem bayesiana, que permite incorporar a informação de estudos similares anteriores, obter regiões de probabilidade para os parâmetros e incorporar decisões com significado ecológico.

### 2.3.1 Função de verosimilhança

O uso do efeito aleatório individual levanta a questão do cálculo da função de máxima verosimilhança, que é obtida através da integração em função do efeito aleatório. Este é, de facto, um problema envolvendo integrais múltiplos complexos que só poderiam ser resolvidos mediante a utilização de aproximações da distribuição *a posteriori* com base nos métodos de Laplace (Chavez-Demoulin, 1999), de argumentos assintóticos (Gould e Nichols, 1998; Burnham e White, 2002), ou através de integração numérica (e.g., Skaug e Fournier, 2006; Wintrebert *et al.*, 2005). Em contrapartida, a abordagem bayesiana proporciona uma solução para este tipo de problema (Brooks *et al.*, 2000, 2002) e o programa WINBUGS tem demonstrado ser uma ferramenta complementar aos programas *clássicos*, tais como o MARK ou CAPTURE.

Para  $i = 1, \dots, n$  e  $j = 1, \dots, t$ , de (2.5) tem-se que:

$$p_{ij} = \frac{e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}} \quad (2.6)$$

$$1 - p_{ij} = \frac{1}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}} \quad (2.7)$$

e de (2.6) e (2.7) tem-se que:

$$p_{ij}^{x_{ij}} (1 - p_{ij})^{(1-x_{ij})} = \frac{e^{(\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i) x_{ij}}}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}} \quad (2.8)$$

Seja  $\mathbf{z} = (\text{sexo}_1, \dots, \text{sexo}_n, \text{idade}_1, \dots, \text{idade}_n, \text{peso}_1, \dots, \text{peso}_n)$  o vector que contém a informação sobre as características individuais, nomeadamente sexo, idade e peso.

Sejam  $\boldsymbol{\beta}_{\text{tempo}} = (\beta_{\text{tempo}_1}, \dots, \beta_{\text{tempo}_t})$ ,  $\boldsymbol{\beta} = (\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \boldsymbol{\beta}_{\text{tempo}})$  e  $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)$ .

A função de verosimilhança dada em (2.3) pode ser reescrita, utilizando (2.8) da seguinte maneira:

$$f(\mathbf{x}, \mathbf{z} | \boldsymbol{\beta}, \boldsymbol{\delta}) = \prod_{i=1}^n \prod_{j=1}^t \left[ \frac{e^{(\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i) x_{ij}}}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}} \right. \\ \left. \times \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_l} + \delta_i}}} \right]. \quad (2.9)$$

A função de verosimilhança do modelo de captura-recaptura do tipo  $\mathcal{M}_{th^*}$ , (2.9), tem um total de  $4 + t + n$  parâmetros a serem estimados. Tendo em conta que a estatística suficiente mínima é obtida a partir de  $(\mathbf{x}, \mathbf{z})$ , cuja dimensão é  $n(3 + t)$ ,

todos os parâmetros do modelo podem ser estimados.

### 2.3.2 Distribuições *a priori*

Na função de verosimilhança (2.3), quando se considera o modelo proposto em (2.5), os parâmetros desconhecidos são  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}}$ , e  $\delta$ . Assumindo-se que todos os parâmetros são independentes e Normalmente distribuídos, considerou-se uma distribuição *a priori* Gaussiana univariável com média 0 e um determinado valor para a precisão, de tal modo que:

$$\beta_0 \sim N(\mu_0, \tau_0) \quad (2.10)$$

$$\beta_{\text{sexo}} \sim N(\mu_{\text{sexo}}, \tau_{\text{sexo}}) \quad (2.11)$$

$$\beta_{\text{idade}} \sim N(\mu_{\text{idade}}, \tau_{\text{idade}}) \quad (2.12)$$

$$\beta_{\text{peso}} \sim N(\mu_{\text{peso}}, \tau_{\text{peso}}) \quad (2.13)$$

$$\beta_{\text{tempo}_j} \stackrel{i.i.d}{\sim} N(\mu_{\text{tempo}}, \tau_{\text{tempo}}), \quad (2.14)$$

para  $j = 1, \dots, t$ , onde  $\mu_0$ ,  $\mu_{\text{sexo}}$ ,  $\mu_{\text{idade}}$ ,  $\mu_{\text{peso}}$  e  $\mu_{\text{tempo}}$  são reais e  $\tau_0$ ,  $\tau_{\text{sexo}}$ ,  $\tau_{\text{idade}}$ ,  $\tau_{\text{peso}}$  e  $\tau_{\text{tempo}}$  são reais positivos.

Os valores de  $\mu_0$ ,  $\mu_{\text{sexo}}$ ,  $\mu_{\text{idade}}$ ,  $\mu_{\text{peso}}$  e  $\mu_{\text{tempo}}$  foram escolhidos como sendo iguais a 0. De modo a garantir pouca precisão para as respectivas distribuições, os valores de  $\tau_{\text{sexo}}$ ,  $\tau_{\text{idade}}$ ,  $\tau_{\text{peso}}$  e  $\tau_{\text{tempo}}$  são pequenos e iguais a 0.1, e  $\tau_0 = 1$ .

Um valor grande para a precisão de cada uma das distribuições *a priori* representa para cada um dos coeficientes um valor perto de 0. Assumir-se que a média para a distribuição *a priori* de cada coeficiente é igual a 0 não implica, no entanto,

que o valor estimado *a posteriori* para cada um dos coeficientes será exactamente 0 (Genkin *et al.*, 2007).

O tratamento hierárquico dos efeitos aleatórios é muito comum em modelos lineares mistos, onde as distribuições são tipicamente Gaussianas (Clark *et al.*, 2004).

Atendendo que não há informação sobre os factores não observáveis, e que isso afecta a probabilidade de captura individual, define-se um modelo hierárquico Normal-Gama para o efeito aleatório individual, de tal modo que  $\delta_i$  depende do hiperparâmetro  $\tau_\delta$ , para o qual se assume a distribuição Gama:

$$\delta_i | \tau_\delta \stackrel{i.i.d}{\sim} N(\mu_\delta, \tau_\delta), \quad i = 1, \dots, n \quad (2.15)$$

$$\tau_\delta \sim \text{Gama}(\xi, \nu) \quad e \quad \xi, \nu \in \mathbb{R}^+. \quad (2.16)$$

A distribuição Gama para a precisão é equivalente a uma distribuição Gama Invertida para a variância. Neste caso, este é a especificação típica para uma distribuição *a priori* vaga (Gelman, 2006; Lambert *et al.*, 2005; van Dongen, 2006).

Note-se que  $\mu_\delta$ , que se assume como sendo igual a 0, representa o valor esperado do efeito aleatório da probabilidade de captura na escala logística. Uma vez que a informação sobre  $\mu_\delta$  é escassa, tem-se uma distribuição *a priori* vaga com uma precisão pequena. Assim, os valores  $\xi$  e  $\nu$  são ambos iguais a 10.0, o que corresponde a um valor esperado para  $\tau_\delta$  igual a 1, mas com uma precisão de 10.0.

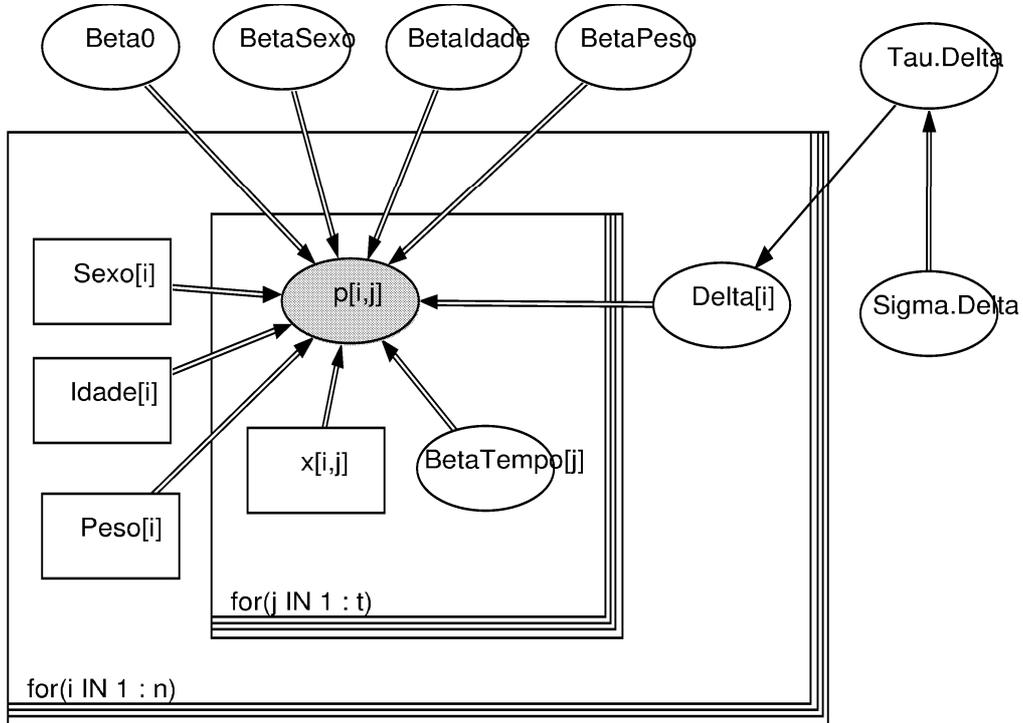


Figura 2.1: DAG do modelo de captura-recaptura quando as probabilidades de captura variam com o tempo e dependem de características individuais e de outros factores não observáveis ( $\mathcal{M}_{th^*}$ ) para estimar o tamanho da população. Os retângulos maiores representam os ciclos em torno das probabilidades de captura  $p_{ij}$  ( $i = 1, \dots, n$  e  $j = 1, \dots, t$ ) e dos efeitos aleatórios individuais  $\delta_i$  ( $i = 1, \dots, n$ ).

### 2.3.3 Distribuições *a posteriori*

Seja  $\theta = (\beta, \delta, \tau_\delta)$  o vector dos parâmetros desconhecidos, a distribuição *a posteriori* completa de  $\theta$  dado  $\mathbf{x}$  e  $\mathbf{z}$  (ver DAG do modelo  $\mathcal{M}_{th^*}$  na Figura 2.1), sob o pressuposto de independência entre as componentes do vector  $\theta$ , é proporcional ao produto entre a função de verosimilhança (2.9) e as distribuições *a priori* (2.10) a (2.16):

$$\begin{aligned}
 h(\boldsymbol{\theta}|\mathbf{x}, \mathbf{z}) &\propto f(\mathbf{x}, \mathbf{z}|\boldsymbol{\theta})h(\beta_0)h(\beta_{\text{sexo}})h(\beta_{\text{idade}})h(\beta_{\text{peso}})h(\beta_{\text{tempo}_1}) \dots h(\beta_{\text{tempo}_t}) \\
 &\times h(\delta_1) \dots h(\delta_n)h(\tau_\delta) \\
 &= \prod_{i=1}^n \prod_{j=1}^t \left[ \frac{e^{(\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i) x_{ij}}}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}} \right. \\
 &\times \left. \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_l} + \delta_i}}} \right] \\
 &\times \frac{\tau_\delta^{\frac{n}{2}}}{(20\pi)^{\frac{t+3}{2}} (2\pi)^{\frac{n+1}{2}}} e^{-\frac{10\beta_0^2 + \beta_{\text{sexo}}^2 + \beta_{\text{idade}}^2 + \beta_{\text{peso}}^2 + \sum_{j=1}^t \beta_{\text{tempo}_j}^2 + 10\tau_\delta \sum_{i=1}^n \delta_i^2}{20}} \\
 &\times \frac{10^{10} \tau_\delta^9 e^{-10\tau_\delta}}{\Gamma(10)}.
 \end{aligned}$$

**Proposição 2.1.** *A distribuição a posteriori condicional completa de  $(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$  dada a informação sobre o histórico das capturas, o número de indivíduos e as informações individuais,  $(\mathbf{x}, \mathbf{z})$ , é dada por:*

a) Para  $i = 1, \dots, n$ ,  $h(\delta_i|\mathbf{x}, \mathbf{z}, \boldsymbol{\beta}, \delta_1, \dots, \delta_{i-1}, \delta_{i+1}, \dots, \delta_n, \tau_\delta)$

$$\begin{aligned}
 &\propto \prod_{j=1}^t \left[ \frac{e^{(\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i) x_{ij}}}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}} \right. \\
 &\times \left. \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_l} + \delta_i}}} \right] e^{-\frac{\tau_\delta \delta_i^2}{2}};
 \end{aligned}$$

b)

$$\tau_\delta \sim \text{Gama} \left( \xi + \frac{1}{2}, v - \sum_{i=1}^n \frac{\delta_i^2}{2} \right);$$

 c)  $h(\beta_0 | \mathbf{x}, \mathbf{z}, \beta_{sexo}, \beta_{idade}, \beta_{peso}, \boldsymbol{\beta}_{tempo}, \boldsymbol{\delta}, \tau_\delta)$ 

$$\begin{aligned} &\propto \prod_{i=1}^n \prod_{j=1}^t \left[ \frac{e^{\left( \beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i \right) x_{ij}}}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i}} \right] \\ &\times \left[ \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_l} + \delta_i}}} \right] e^{-\frac{\beta_0^2}{2}}; \end{aligned}$$

 d)  $h(\beta_{sexo} | \mathbf{x}, \mathbf{z}, \beta_0, \beta_{idade}, \beta_{peso}, \boldsymbol{\beta}_{tempo}, \boldsymbol{\delta}, \tau_\delta)$ 

$$\begin{aligned} &\propto \prod_{i=1}^n \prod_{j=1}^t \left[ \frac{e^{\left( \beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i \right) x_{ij}}}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i}} \right] \\ &\times \left[ \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_l} + \delta_i}}} \right] e^{-\frac{\beta_{sexo}^2}{20}}; \end{aligned}$$

 e)  $h(\beta_{idade} | \mathbf{x}, \mathbf{z}, \beta_0, \beta_{sexo}, \beta_{peso}, \boldsymbol{\beta}_{tempo}, \boldsymbol{\delta}, \tau_\delta)$ 

$$\begin{aligned} &\propto \prod_{i=1}^n \prod_{j=1}^t \left[ \frac{e^{\left( \beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i \right) x_{ij}}}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i}} \right] \\ &\times \left[ \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_l} + \delta_i}}} \right] e^{-\frac{\beta_{idade}^2}{20}}; \end{aligned}$$

f)  $h(\beta_{peso} | \mathbf{x}, \mathbf{z}, \beta_0, \beta_{sexo}, \beta_{idade}, \boldsymbol{\beta}_{tempo}, \boldsymbol{\delta}, \tau_\delta)$

$$\begin{aligned} &\propto \prod_{i=1}^n \prod_{j=1}^t \left[ \frac{e^{(\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i) x_{ij}}}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i}} \right. \\ &\times \left. \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_l} + \delta_i}}} \right] e^{-\frac{\beta_{peso}^2}{20}}; \end{aligned}$$

g) Para  $j = 1, \dots, t$ ,

$$h(\beta_{tempo_j} | \mathbf{x}, \mathbf{z}, \beta_0, \beta_{sexo}, \beta_{idade}, \beta_{peso}, \beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_t, \boldsymbol{\delta}, \tau_\delta)$$

$$\begin{aligned} &\propto \prod_{i=1}^n \left[ \frac{e^{(\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i) x_{ij}}}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_j} + \delta_i}} \right. \\ &\times \left. \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{sexo} \times sexo_i + \beta_{idade} \times idade_i + \beta_{peso} \times peso_i + \beta_{tempo_l} + \delta_i}}} \right] e^{-\frac{\beta_{tempo_j}^2}{20}}; \end{aligned}$$

*Demonstração.* Demonstra-se apenas a alínea b), uma vez que as outras alíneas surgem directamente do produto entre as respectivas distribuições *a priori* e a função de verosimilhança dada em (2.9).

b) Tendo em conta a função de verosimilhança dada em (2.9) e as distribuições *a priori* (2.15) e (2.16) tem-se que:

$$\begin{aligned}
 & h(\tau_\delta | \mathbf{x}, \mathbf{z}, n, \boldsymbol{\beta}, \boldsymbol{\delta}) \propto f(\mathbf{x}, n, \mathbf{z} | \boldsymbol{\theta}) h(\delta_1) \dots h(\delta_n) h(\tau_\delta) \\
 & = \prod_{i=1}^n \prod_{j=1}^t \left[ \frac{e^{(\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i) x_{ij}}}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_j} + \delta_i}} \right. \\
 & \times \left. \frac{1}{1 - \prod_{l=1}^t \frac{1}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{peso}} \times \text{peso}_i + \beta_{\text{tempo}_l} + \delta_i}}} \right] \\
 & \times \sqrt{\frac{\tau_\delta}{2\pi}} e^{-\sum_{i=1}^n \frac{\tau_\delta \delta_i^2}{2}} \frac{\xi^v \tau_\delta^{\xi-1} e^{v\tau_\delta}}{\Gamma(\xi)} \\
 & \propto \frac{\xi^v \tau_\delta^{\xi-\frac{1}{2}} e^{\tau_\delta \left(v - \sum_{i=1}^n \frac{\delta_i^2}{2}\right)}}{\Gamma(\xi)} \propto \frac{\left(\xi + \frac{1}{2}\right)^{\left(v - \sum_{i=1}^n \frac{\delta_i^2}{2}\right)} \tau_\delta^{\xi-\frac{1}{2}} e^{\tau_\delta \left(v - \sum_{i=1}^n \frac{\delta_i^2}{2}\right)}}{\Gamma\left(\xi + \frac{1}{2}\right)}.
 \end{aligned}$$

□

## 2.4 Aplicação em capturas de ratos veadeiro

Utilizaram-se como exemplo os dados de capturas de ratos veadeiro (*Peromyscus maniculatus*) realizadas por V. Reid na zona Este de Stuart Gulch, em Colorado nos EUA (Rexstad e Burnham, 1991). As probabilidades das capturas dependem de covariáveis individuais, tais como o sexo, a idade (categorizada em juvenis, semi-adultos e adultos) e o peso (em gramas), e do histórico das capturas. Foram realizados seis momentos de capturas, onde foram capturados 38 ratos. Assim como em Huggins (1991), também se agruparam os três semi-adultos com os adultos. Os dados também analisados por Huggins (1991) são apresentados no Apêndice B.3, na Tabela B.1.

### 2.4.1 Análises de diagnóstico

Na Figura 2.2 verifica-se que as duas cadeias relativamente aos parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$ ,  $\beta_{\text{tempo}_5}$ ,  $\tau_\delta$ ,  $\delta_1$  e  $\delta_{38}$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ , ambas convergem para a mesma região, apesar de terem pontos iniciais diferentes. Para além disso, o teste de Gelman e Rubin confirma que existe convergência das cadeias. No gráfico do factor de redução<sup>10</sup> ( $R$ ), na Figura 2.3, verifica-se que  $R$  converge para 1 para todos os respectivos parâmetros e por isso indicando convergência para as respectivas cadeias.

O diagnóstico de autocorrelação foi realizado para todas as cadeias, indicando que um espaçamento de tamanho 40 entre iterações diminuía a autocorrelação para valores próximos de zero. O teste de convergência de Raftery e Lewis (1995) indicou que existe convergência para todos os parâmetros, e tal acontece ao fim de poucas iterações (Tabela 2.1). Entretanto, o teste de convergência Gelman e Rubin (Figura 2.3), levanta alguma dúvida se o período de aquecimento é realmente curto.

Assim, para as estimativas bayesianas adoptou-se para os parâmetros apenas uma das cadeias, tendo-se verificado que 500,000 iterações, um período de aquecimento de 5000 iterações e um espaçamento de tamanho 40 entre iterações era adequado para assegurar convergência, remover a influência dos valores iniciais e diminuir a autocorrelação para valores próximos de zero. Tendo em conta uma amostra final de tamanho 12,375, foram seleccionados a mediana da distribuição *a posteriori* como estimativa pontual dos parâmetros e um intervalo de credibilidade/HPD de 95% para os parâmetros.

---

<sup>10</sup> $R$  é a raiz quadrada da razão entre a variância entre as cadeias e a variância dentro das cadeias. Um valor grande (maior do que 1) para  $R$  indica que a variância entre as cadeias é substancialmente maior que a variância dentro das cadeias, o que quer dizer que serão necessárias mais iterações.

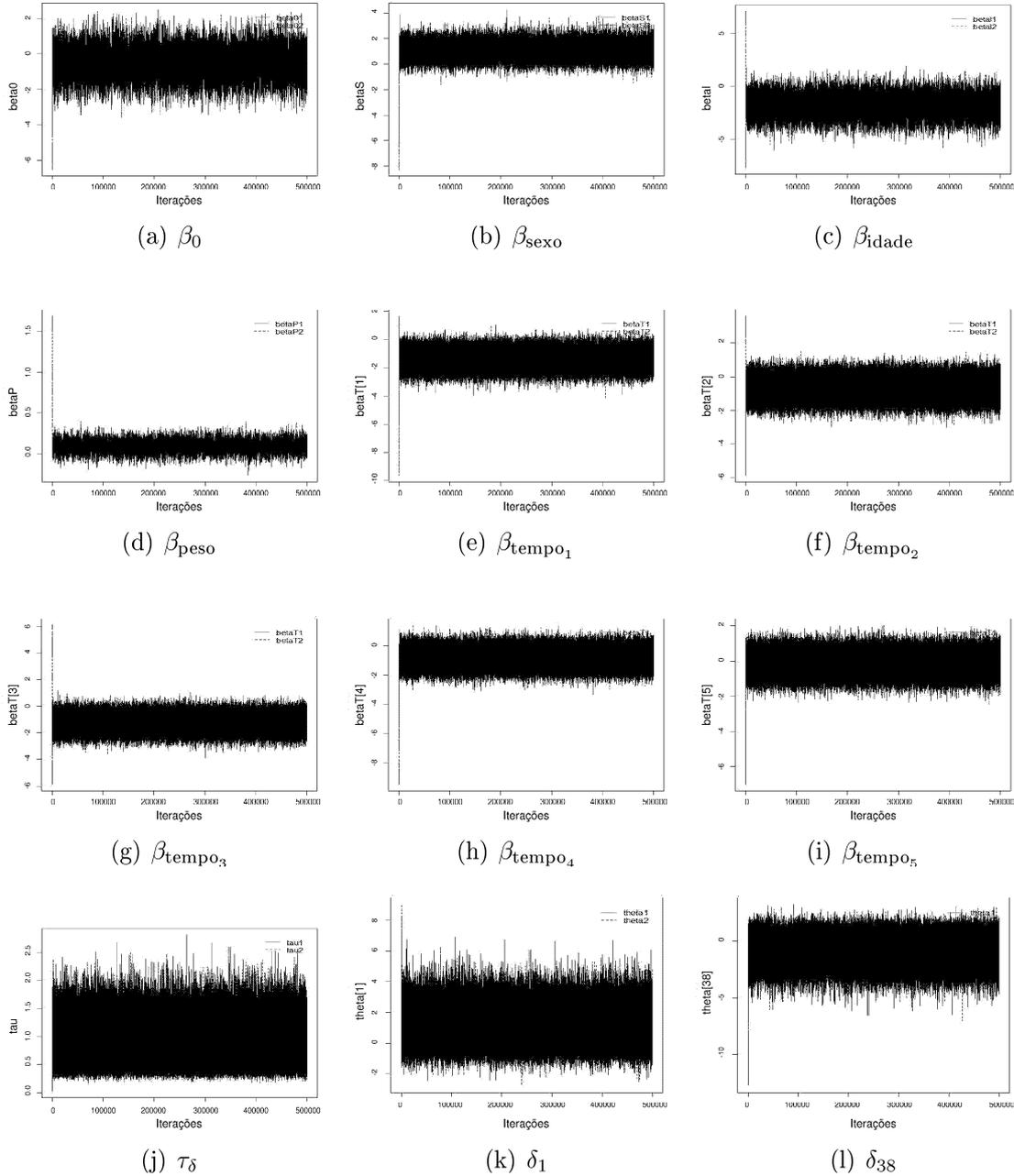


Figura 2.2: Traço das duas cadeias de Markov, com pontos iniciais diferentes, consideradas para os parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$ ,  $\beta_{\text{tempo}_5}$ ,  $\tau_\delta$ ,  $\delta_1$  e  $\delta_{38}$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ .

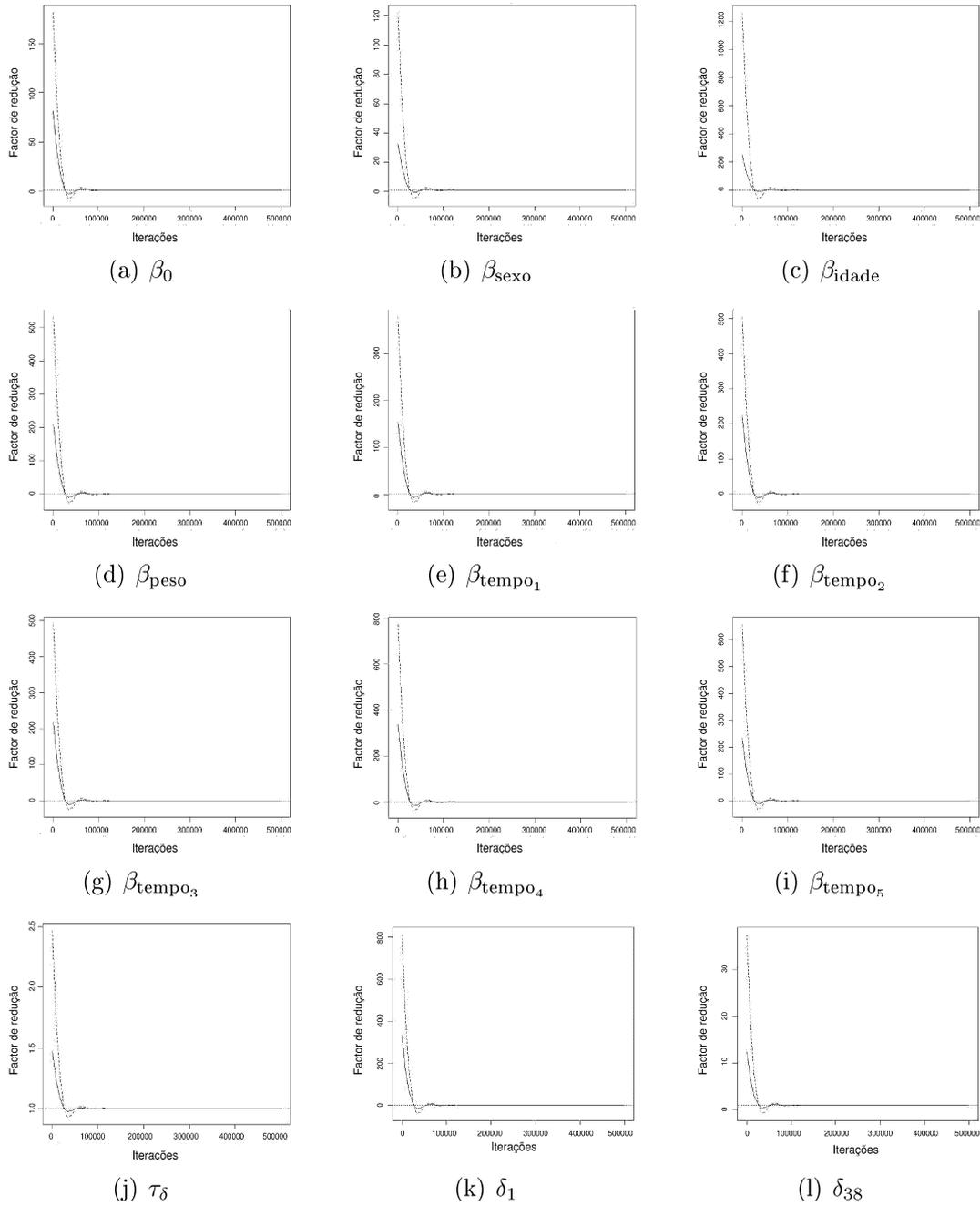


Figura 2.3: Teste de convergência Gelman e Rubin, utilizando duas cadeias de Markov com pontos iniciais diferentes, para os parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$ ,  $\beta_{\text{tempo}_5}$ ,  $\tau_\delta$ ,  $\delta_1$  e  $\delta_{38}$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th}^*$ . - - - - 97.5%, ——— mediana.

Tabela 2.1: Estimativa do período de aquecimento e da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50) para as cadeias de Markov dos parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$ ,  $\beta_{\text{tempo}_5}$ ,  $\tau_\delta$ ,  $\delta_1$ ,  $\delta_5$ ,  $\delta_{10}$ ,  $\delta_{15}$ ,  $\delta_{20}$ ,  $\delta_{25}$ ,  $\delta_{30}$  e  $\delta_{38}$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ .

Parâmetros	Período de aquecimento	Função de autocorrelação			
		Espaçamento			
		5	10	20	50
$\beta_0$	276	0.9237	0.8600	0.7532	0.5348
$\beta_{\text{sexo}}$	34	0.7719	0.6190	0.4150	0.1536
$\beta_{\text{idade}}$	62	0.8451	0.7490	0.6214	0.4300
$\beta_{\text{peso}}$	366	0.9570	0.9181	0.8471	0.6684
$\beta_{\text{tempo}_1}$	44	0.4888	0.3338	0.1914	0.0494
$\beta_{\text{tempo}_2}$	40	0.4727	0.3224	0.1877	0.0502
$\beta_{\text{tempo}_3}$	38	0.4792	0.3268	0.1857	0.0488
$\beta_{\text{tempo}_4}$	40	0.4918	0.3337	0.1883	0.0472
$\beta_{\text{tempo}_5}$	38	0.4506	0.3055	0.1755	0.0453
$\tau_\delta$	42	0.3828	0.2267	0.0884	0.0139
$\delta_1$	18	0.2015	0.0589	0.0150	0.0050
$\delta_5$	20	0.1878	0.0603	0.0185	0.0024
$\delta_{10}$	22	0.1901	0.0709	0.0306	0.0127
$\delta_{15}$	20	0.1849	0.0598	0.0257	0.0087
$\delta_{20}$	22	0.2013	0.08001	0.0410	0.0126
$\delta_{25}$	22	0.3244	0.1689	0.0850	0.0284
$\delta_{30}$	20	0.1633	0.0408	0.01236	-0.0006
$\delta_{38}$	20	0.1797	0.0433	0.0038	-0.0016

## 2.4.2 Análise de sensibilidade

Consideraram-se quatro cenários para análise de sensibilidade, admitindo diferentes valores que definem as distribuições *a priori* dos respectivos parâmetros:

$$(1) \beta_0 \sim N(0, 1.0), \beta_{\text{sexo}} \sim N(0, 0.1), \beta_{\text{idade}} \sim N(0, 0.1), \beta_{\text{peso}} \sim N(0, 0.1), \beta_{\text{tempo}_j} \overset{i.i.d}{\sim}$$

$N(0, 0.1)$ , para  $j = 1, \dots, t - 1$  e  $\delta_i | \tau_\delta \stackrel{i.i.d}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim Gama(10.0, 10.0)$ ,  
para  $i = 1, \dots, n$ ;

(2)  $\beta_0 \sim N(0, 0.1)$ ,  $\beta_{\text{sexo}} \sim N(0, 0.1)$ ,  $\beta_{\text{idade}} \sim N(0, 0.1)$ ,  $\beta_{\text{peso}} \sim N(0, 0.1)$ ,  $\beta_{\text{tempo}_j} \stackrel{i.i.d}{\sim}$   
 $N(0, 0.1)$ , para  $j = 1, \dots, t - 1$  e  $\delta_i | \tau_\delta \stackrel{i.i.d}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim Gama(10.0, 10.0)$ ,  
para  $i = 1, \dots, n$ ;

(3)  $\beta_0 \sim N(0, 1.0)$ ,  $\beta_{\text{sexo}} \sim N(0, 1.0)$ ,  $\beta_{\text{idade}} \sim N(0, 1.0)$ ,  $\beta_{\text{peso}} \sim N(0, 1.0)$ ,  $\beta_{\text{tempo}_j} \stackrel{i.i.d}{\sim}$   
 $N(0, 1.0)$ , para  $j = 1, \dots, t - 1$  e  $\delta_i | \tau_\delta \stackrel{i.i.d}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim Gama(10.0, 10.0)$ ,  
para  $i = 1, \dots, n$ ;

(4)  $\beta_0 \sim N(0, 1.0)$ ,  $\beta_{\text{sexo}} \sim N(0, 0.1)$ ,  $\beta_{\text{idade}} \sim N(0, 0.1)$ ,  $\beta_{\text{peso}} \sim N(0, 0.1)$ ,  $\beta_{\text{tempo}_j} \stackrel{i.i.d}{\sim}$   
 $N(0, 0.1)$ , para  $j = 1, \dots, t - 1$  e  $\delta_i | \tau_\delta \stackrel{i.i.d}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim Gama(0.5, 0.5)$ ,  
para  $i = 1, \dots, n$ .

No primeiro cenário, admite-se maior precisão, embora ainda pequena, para a distribuição *a priori* atribuída ao parâmetro intersecção na regressão logística,  $\beta_0$ , e menor precisão nos restantes cenários.

Para os parâmetros  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$  e  $\beta_{\text{tempo}}$  escolheu-se o terceiro cenário, como sendo aquele em que todos apresentam uma distribuição *a priori* com muita precisão, e menor precisão nos outros cenários (Figura 2.4(a)). Os valores que definem a distribuição *a priori* hierárquica para o efeito aleatório,  $\delta$ , representam todos uma precisão igual a 1.0 para a distribuição Gaussiana. No entanto o parâmetro  $\tau_\delta$  tem maior precisão (=10.0) nos três primeiros cenários e menor precisão (=0.5) no último cenário (Figura 2.4(b)).

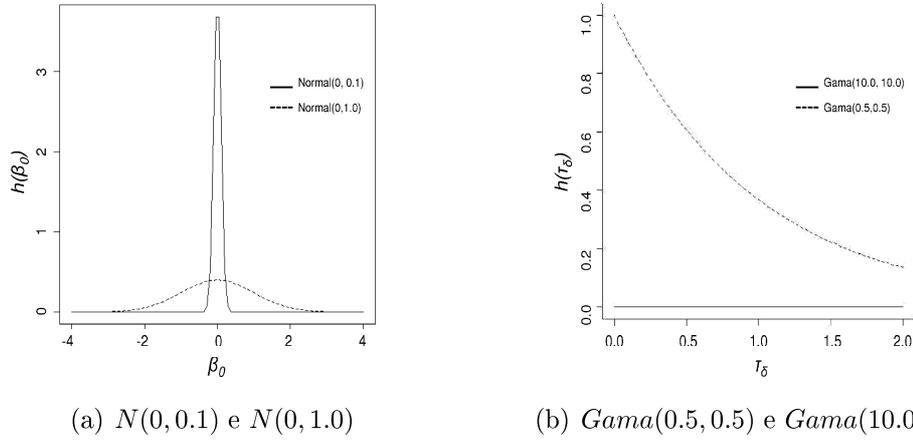


Figura 2.4: Gráfico das distribuições *a priori* escolhidas para análise de sensibilidade para os parâmetros:  $(\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{peso}}, \beta_{\text{tempo}}, \delta)$  a) e  $\tau_\delta$  b).

O desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) foi pequeno, indicando que as distribuições *a posteriori* foram adequadamente amostradas. As medianas *a posteriori* para os parâmetros de interesse são coerentes entre todas as análises (Tabelas 2.2 e 2.3 e Figura 2.5). Nos quatro cenários, a mediana *a posteriori* (e o desvio padrão) para  $N$  foram, respectivamente, iguais a 46.62 (10.08), 47.12 (11.10), 45.67 (8.04), 52.28 (260.50).

A inferência sobre  $N$  é muito sensível à escolha dos valores que definem a distribuição *a priori* de  $\delta$ , sobretudo, no que diz respeito à variabilidade. Quanto menor for a precisão da distribuição *a priori* definida para  $\delta$  maior será a variância associada à estimação *a posteriori* de  $N$ . Nos cenários 1, 2 e 3, os parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$  e  $\beta_{\text{tempo}}$  obtiveram diferentes estimativas *a posteriori*, de acordo com a escolha de diferentes valores para as respectivas distribuições *a priori*, mas, entretanto, a estimativa *a posteriori* para  $N$  não foi tão sensível, quando comparada nos quatro cenários.

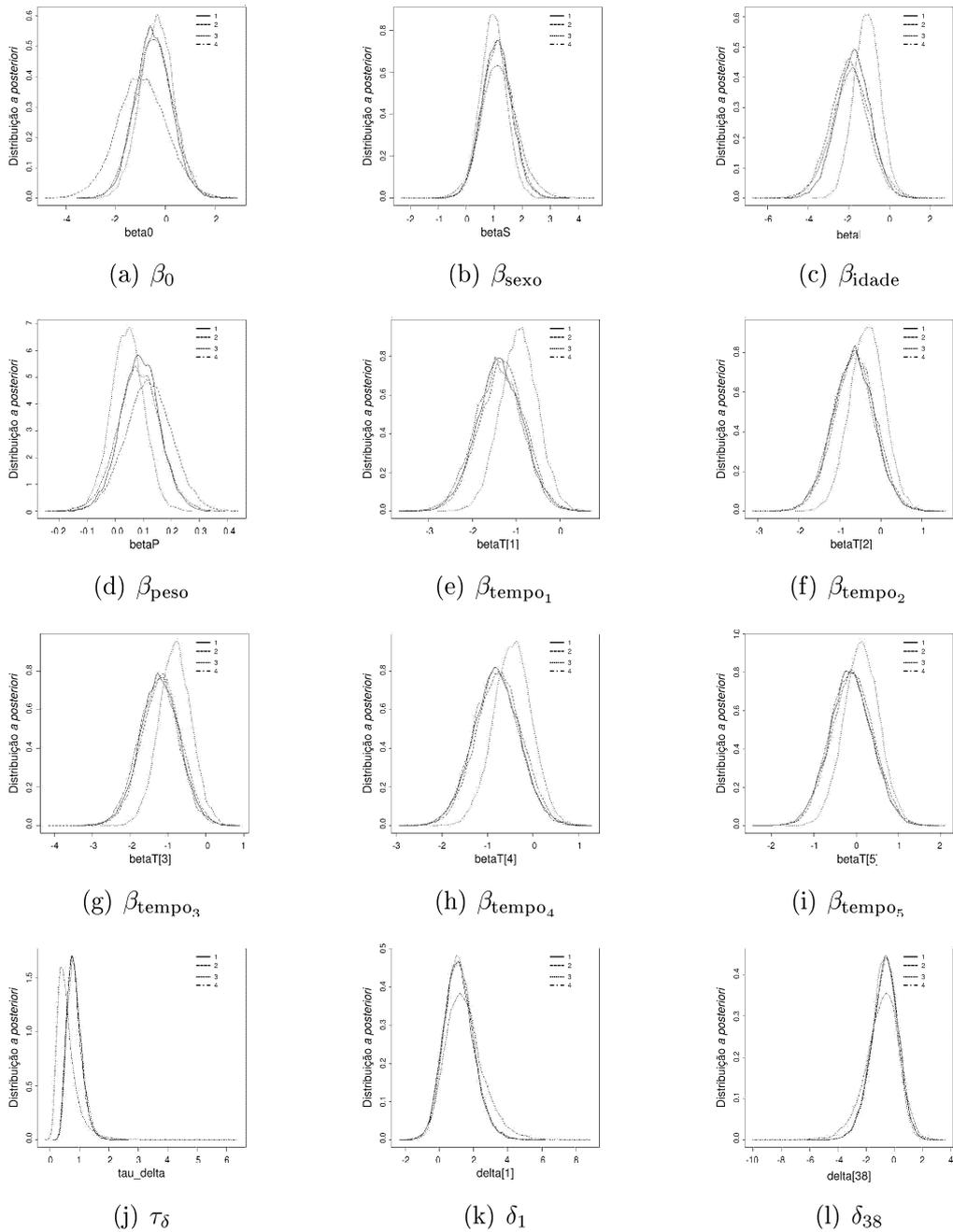


Figura 2.5: Gráfico das distribuições *a posteriori* para os parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$ ,  $\beta_{\text{tempo}_5}$ ,  $\tau_{\delta}$ ,  $\delta_1$  e  $\delta_{38}$ , obtidas tendo em conta as distribuições *a priori* dos quatro cenários.

Tabela 2.2: Resultado da análise de sensibilidade para os parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$  e  $\beta_{\text{tempo}_2}$ , tendo em conta as distribuições *a priori* nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição *a posteriori* é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD.

Distribuição <i>a priori</i>	Características da cadeia de Markov					Distribuição <i>a posteriori</i>			
	$se_b$	Espaçamento				Mediana	$sd$	I.C./HPD	
		5	10	20	50				
$\beta_0$	1	0.0071	0.9237	0.8600	0.7532	0.5348	-0.5056	0.7230	-1.883, 0.948
	2	0.0099	0.9598	0.9246	0.8625	0.7100	-0.9534	1.0015	-2.875, 1.069
	3	0.0067	0.9205	0.8544	0.7450	0.5231	-0.3370	0.6824	-1.739, 0.959
	4	0.0075	0.9263	0.8636	0.7570	0.5321	-0.4539	0.7668	-1.883, 1.118
$\beta_{\text{sexo}}$	1	0.0051	0.7719	0.6190	0.4150	0.1536	1.0850	0.5334	[0.030, 2.112]
	2	0.0052	0.7755	0.6248	0.4178	0.1442	1.1080	0.5430	[0.020, 2.140]
	3	0.0044	0.7237	0.5480	0.3270	0.0862	0.9587	0.4607	[0.055, 1.867]
	4	0.0062	0.8293	0.7062	0.5224	0.2421	1.1460	0.6505	[-0.095, 2.501]
$\beta_{\text{idade}}$	1	0.0082	0.8451	0.7490	0.6214	0.4300	-1.7800	0.8437	[-3.397, -0.077]
	2	0.0088	0.8674	0.7834	0.6742	0.5089	-2.0050	0.9093	[-3.760, -0.204]
	3	0.0062	0.7663	0.6352	0.4843	0.2952	-1.1190	0.6434	[-2.450, 0.069]
	4	0.0095	0.8788	0.7957	0.6741	0.4679	-1.8590	0.9757	[-3.850, 0.021]
$\beta_{\text{peso}}$	1	0.0007	0.9570	0.9181	0.8471	0.6684	0.0883	0.0686	[-0.051, 0.223]
	2	0.0008	0.9715	0.9456	0.8975	0.7744	0.1193	0.0833	[-0.045, 0.284]
	3	0.0006	0.9471	0.9004	0.8156	0.6131	0.0439	0.0585	[-0.074, 0.156]
	4	0.0008	0.9615	0.9267	0.8628	0.7014	0.0835	0.0759	[-0.069, 0.232]
$\beta_{\text{tempo}_1}$	1	0.0049	0.4888	0.3338	0.1915	0.0494	-1.3710	0.5198	[-2.360, -0.319]
	2	0.0049	0.4960	0.3446	0.2073	0.0533	-1.3080	0.5232	[-2.386, -0.351]
	3	0.0041	0.3777	0.2090	0.0890	0.0112	-0.9322	0.4317	[-1.748, -0.065]
	4	0.0050	0.4913	0.3267	0.1846	0.0479	-1.4090	0.5313	[-2.464, -0.411]
$\beta_{\text{tempo}_2}$	1	0.0048	0.4727	0.3224	0.1877	0.0502	-0.6661	0.5045	[-1.673, 0.310]
	2	0.0048	0.4903	0.3410	0.2009	0.0594	-0.6242	0.5084	[-1.677, 0.320]
	3	0.0040	0.3699	0.2091	0.0908	0.0076	-0.3371	0.4218	[-1.188, 0.462]
	4	0.0048	0.4630	0.3150	0.1784	0.0437	-0.6788	0.5068	[-1.627, 0.376]

Tabela 2.3: Resultado da análise de sensibilidade para os parâmetros  $\beta_{tempo_3}$ ,  $\beta_{tempo_4}$ ,  $\beta_{tempo_5}$ ,  $\tau_\delta$ ,  $\delta_1$  e  $\delta_{38}$ , tendo em conta as distribuições *a priori* nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição *a posteriori* é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD.

Distribuição <i>a priori</i>	Características da cadeia de Markov						Distribuição <i>a posteriori</i>		
	$se_b$	Espaçamento				Mediana	$sd$	I.C./HPD	
		5	10	20	50				
$\beta_{tempo_3}$	1	0.0049	0.4792	0.3268	0.1857	0.0488	-1.2190	0.5212	[-2.280,-0.224]
	2	0.0049	0.4891	0.3389	0.1996	0.0545	-1.1670	0.5198	[-2.209,-0.176]
	3	0.0040	0.3635	0.2031	0.0857	0.0129	-0.8115	0.4287	[-1.663,0.027]
	4	0.0049	0.4575	0.3073	0.1741	0.0417	-1.2480	0.5245	[-2.301,-0.262]
$\beta_{tempo_4}$	1	0.0047	0.4918	0.3337	0.1883	0.0472	-0.8058	0.5025	[-1.816,0.156]
	2	0.0048	0.5094	0.3567	0.2093	0.0557	-0.7469	0.5063	[-1.722,0.245]
	3	0.0040	0.3735	0.2076	0.0866	0.0090	-0.4486	0.4228	[-1.314,0.351]
	4	0.0047	0.4786	0.3209	0.1783	0.0425	-0.8174	0.5037	[-1.802,0.167]
$\beta_{tempo_5}$	1	0.0047	0.4506	0.3055	0.1755	0.0453	-0.1344	0.5010	[-1.115,0.850]
	2	0.0048	0.4742	0.3239	0.1960	0.0562	-0.0880	0.5054	[-1.049,0.908]
	3	0.0040	0.3559	0.1971	0.0882	0.0093	0.1249	0.4221	[-0.709,0.954]
	4	0.0047	0.4405	0.2907	0.1631	0.0451	-0.1402	0.4982	[-1.129,0.820]
$\tau_\delta$	1	0.0024	0.3828	0.2267	0.0884	0.0139	0.8010	0.2547	[0.388,1.335]
	2	0.0024	0.3873	0.2300	0.0928	0.0206	0.7933	0.2229	[0.384,1.328]
	3	0.0024	0.3710	0.2120	0.0780	0.0132	0.8374	0.2602	[0.402,1.377]
	4	0.0038	0.7048	0.5744	0.3855	0.1349	0.5015	0.3901	[0.101,1.291]
$\delta_1$	1	0.0083	0.2015	0.0590	0.0150	0.0050	1.0850	0.8878	[-0.591,2.843]
	2	0.0083	0.2056	0.0654	0.0140	0.0033	1.1140	0.8846	[-0.493,2.957]
	3	0.0080	0.2262	0.0749	0.0194	0.0017	1.1430	0.8580	[-0.413,2.926]
	4	0.0106	0.2771	0.1319	0.0637	0.0185	1.3710	1.1272	[-0.562,3.777]
$\delta_{38}$	1	0.0090	0.2072	0.0790	0.0297	0.0033	-0.6472	0.9537	[-2.675,1.065]
	2	0.0089	0.1849	0.0464	0.0041	0.0011	-0.6282	0.9523	[-2.654,1.074]
	3	0.0087	0.2098	0.0594	0.0101	0.0003	-0.7165	0.9203	[-2.600,1.010]
	4	0.0114	0.2306	0.0878	0.0309	0.0029	-0.7859	1.2160	[-3.378,1.352]

### 2.4.3 Selecção de modelos

Diferentes modelos proporcionam diferentes estimativas, por isso a escolha entre os modelos pode ser na maior parte das vezes importante. O objectivo da selecção de modelos é encontrar um meio termo entre a qualidade de ajustamento dos dados observados e a complexidade do modelo. Tem sido um desafio grande encontrar um bom método para a selecção de modelos para populações fechadas. Tipicamente, o melhor modelo, i.e. aquele que for seleccionado, será utilizado para se fazer a inferência, embora ultimamente também tem sido utilizado mais do que um modelo para se fazer inferência, designado por ponderação de modelos (e.g., Burnham e Anderson, 2004; Hoeting *et al.*, 1999).

Existem dois grupos de métodos para selecção de modelos: um grupo que produz um valor que serve de critério de selecção para cada modelo, que será comparado entre um conjunto de modelos (e.g., média quadrática do erro predictivo (Gelfand e Ghosh, 1998), DIC (Critério de Informação da Desviância (*Deviance Information Criteria*), Spiegelhalter *et al.*, 2002), BIC (Critério de Informação Bayesiano (*Bayesian Information Criteria*), por exemplo, Link e Barker, 2006)) e outro grupo que efectua a exploração automática do espaço dos modelos (e.g., selecção de variável Gibbs (George e McCulloch, 1993), método de Monte Carlo baseado em cadeias de Markov com saltos reversíveis (Green, 1995)).

Entre os vários métodos para a selecção de modelos, na literatura da estatística em ecologia, o Critério de Informação de Akaike (AIC) é sem dúvida o mais consensual (Burnham e Anderson, 2002).

Quando se estimam os parâmetros utilizando o método de máxima verosimi-

lhança, é possível aumentar o valor da função de verosimilhança, acrescentando parâmetros ao modelo, o que pode resultar em sobre-ajustamento. Deste modo, Akaike (1974) propôs um critério para a selecção de modelos, que penaliza aqueles modelos com maior número de parâmetros:

$$AIC = -2 \log f(\mathbf{y}|\hat{\boldsymbol{\theta}}) + 2p_{\theta}, \quad (2.17)$$

onde  $f(\mathbf{y}|\hat{\boldsymbol{\theta}})$  é a função verosimilhança,  $p_{\theta}$  é a dimensão do vector parâmetro  $\boldsymbol{\theta}$  e  $\hat{\boldsymbol{\theta}}$  é o vector estimador de máxima verosimilhança para os parâmetros do modelo candidato. Burnham e Anderson (2002, 2004) têm chamado a atenção para a importância de se utilizar uma versão corrigida para o AIC, quando o tamanho da amostra ( $n$ ) é pequena quando comparado com o número de parâmetros, e.g.  $n/p_{\theta} < 40$ , que foi desenvolvida por Sugiura (1978) como sendo:

$$AIC_c = AIC + \frac{2p_{\theta}(p_{\theta} + 1)}{n - p_{\theta} - 1}.$$

O Critério de Informação Bayesiano ou o Critério de Schwarz, BIC (*Bayesian Information Criteria*), e também por vezes designado por SBC ou SBIC, foi desenvolvido por Schwarz (1978), propondo assim um critério bayesiano, muito similar ao AIC, para a selecção de modelos:

$$BIC = -2 \log f(\mathbf{y}|\hat{\boldsymbol{\theta}}) + p_{\theta} \log n. \quad (2.18)$$

O BIC penaliza de forma mais acentuada os modelos com maior número de parâmetros do que o AIC (comparar 2.17 com 2.18). Esses dois critérios, AIC e

BIC, tendem a seleccionar modelos menos complexos, i.e. aqueles que têm menor número de parâmetros, e o modelo seleccionado será aquele que apresentar o menor valor para o respectivo critério. Contudo, tanto o AIC como o BIC não funcionam muito bem para modelos que envolvem variáveis latentes, uma vez que nesses casos a dimensão do vector de parâmetros não é bem definida.

Do ponto de vista frequencista, a avaliação do modelo pode ser feita com base na desviância ( $D$ , *deviance*), que é dada pela diferença do logaritmo da verosimilhança entre o modelo ajustado e o modelo saturado. O modelo saturado refere-se ao modelo cujo número de parâmetro corresponde ao número de observações, o que resulta num ajuste perfeito aos dados. Dempster (1974) sugeriu examinar a distribuição *a posteriori* da desviância clássica definida por:

$$D(\boldsymbol{\theta}) = -2 \log f(\mathbf{y}|\boldsymbol{\theta}) + 2 \log h(\mathbf{y}),$$

onde  $h(\mathbf{y})$  é uma determinada função padronizada especificada que só depende dos dados. Não obstante, a desviância não é recomendada para a selecção de modelos complexos, uma vez que os modelos complexos têm tendência para se ajustarem melhor aos dados e terão por isso sempre o menor valor para a desviância. Assim, é necessário medir essa “complexidade” e fazer uma selecção adequada dos modelos. A complexidade pode ser medida através do “número efectivo de parâmetros”, designado por  $p_D$  e também conhecido por *complexidade bayesiana*:

$$p_D = \overline{D(\boldsymbol{\theta})} - D(\tilde{\boldsymbol{\theta}}),$$

onde  $\tilde{\theta}$  é uma estimativa de  $\theta$  que depende de  $\mathbf{y}$ . Geralmente,  $\bar{\theta} = E[\theta|\mathbf{y}]$  é a escolha natural para  $\tilde{\theta}$ , mas a mediana ou a moda *a posteriori* são outras possíveis alternativas.

Spiegelhalter *et al.* (2002) adoptaram essa ideia como base para desenvolver o Critério de Informação da Desviância (DIC, *Deviance Information Criteria*). Este critério consiste em duas componentes, uma que mede a qualidade do ajuste e outra que penaliza o modelo pelo facto de aumentar a complexidade do modelo:

$$\begin{aligned} DIC &= \text{“qualidade de ajustamento”} + \text{“complexidade”} \\ &= D(\tilde{\theta}) + 2p_D \\ &= \overline{D(\theta)} + p_D. \end{aligned}$$

Basicamente, o DIC é um critério de selecção de modelos que considera os requisitos do ajuste do modelo e o nível baixo de complexidade. Geralmente, quanto mais parâmetros se acrescentam aos modelos, mais complexos ficam, e melhor será o ajuste. O Critério de Informação da Desviância, conseqüentemente, penaliza os modelos com mais parâmetros e, por conseguinte, o modelo que satisfaz o princípio de parcimónia é escolhido, apresentando neste caso o menor valor para o DIC.

A vantagem do DIC sobre os outros critérios, acima apresentados, para a selecção bayesiana de modelos, é que pode ser facilmente calculado no programa WINBUGS, a partir de amostras das cadeias geradas através do método MCMC. O DIC é obtido mediante o cálculo de  $\overline{D(\theta)}$  como sendo uma média de  $D(\theta)$  através das amostras de  $\theta$ , e  $D(\bar{\theta})$  como sendo o valor de  $D(\theta)$  obtido da média das amostras  $\theta$ . Todavia, o DIC também tem sido alvo de alguma controvérsia por parte de Spiegelhalter *et al.*

(2002) e de Celeux *et al.* (2006), no sentido de não ser o critério mais adequado para selecção de modelos hierárquicos e/ou mistos.

Considerou-se o DIC como um critério preliminar para a selecção de modelos. Note-se que, quando se utiliza a inferência bayesiana, os modelos mais complexos,  $\mathcal{M}_{th}$  e  $\mathcal{M}_{th^*}$ , apresentam menor valor (em termos absolutos) para a desviância, respectivamente iguais a 290.70 e 248.00, quando comparados com os outros dois modelos mais simples,  $\mathcal{M}_t$  e  $\mathcal{M}_h$ , respectivamente iguais a 311.90 e 295.70. Analogamente, os modelos mais complexos são os mais penalizados, apresentando  $p_D$  iguais a 8.55 e 21.57, respectivamente.

Entre os quatro modelos, quando se utiliza a inferência bayesiana, o modelo  $\mathcal{M}_{th^*}$  é o que melhor se ajusta aos dados, tendo em conta o DIC, pois é o modelo que apresenta o menor valor para o DIC, DIC=291.15 (Tabela 2.4).

Tabela 2.4: Comparação entre  $p_\theta$ ,  $p_D$  e os resultados obtidos para os diferentes critérios (AIC, BIC, D e DIC) para selecção dos modelos  $\mathcal{M}_t$ ,  $\mathcal{M}_h$ ,  $\mathcal{M}_{th}$  e  $\mathcal{M}_{th^*}$ , considerando o método de máxima verosimilhança (MV) e a inferência bayesiana (IB).

	$\mathcal{M}_t$		$\mathcal{M}_h$		$\mathcal{M}_{th}$		$\mathcal{M}_{th^*}$
	MV	IB	MV	IB	MV	IB	IB
$p_\theta$	6	6	4	4	9	9	47
$p_D$	-0.003	5.87	-4.65	3.73	-21.018	8.55	21.57
AIC	317.97	323.30	295.86	303.00	296.23	308.00	341.50
BIC	342.70	348.63	313.24	321.08	335.34	347.81	546.24
D	305.97	311.90	287.86	295.70	278.23	290.70	- 248.00
DIC	305.96	323.64	278.56	303.16	236.20	307.80	291.15

Entretanto, independentemente do método de inferência utilizado, se se levar em conta os outros critérios para a selecção de modelos, a desviância considera o modelo mais complexo  $\mathcal{M}_{th^*}$ , que tem maior número de parâmetros, e os critérios AIC e BIC seleccionam o modelo  $\mathcal{M}_h$ .

### 2.4.4 Resultados e análise do modelo

Na Tabela 2.5, apresentam-se as estimativas para os parâmetros que definem as probabilidades de captura, o tamanho da população, bem como as respectivas estimativas do desvio padrão, considerando os 4 modelos e as duas técnicas de inferência.

Tabela 2.5: Estimativas, coeficiente de variação [CV%] e desvio padrão (*sd*) obtidos com os métodos de máxima verosimilhança (MV) e inferência bayesiana (IB), para os parâmetros  $N$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$  e  $\beta_{\text{tempo}_5}$ , considerando os modelos  $\mathcal{M}_t$ ,  $\mathcal{M}_h$ ,  $\mathcal{M}_{th}$  e  $\mathcal{M}_{th^*}$ .

	$\mathcal{M}_t$		$\mathcal{M}_h$		$\mathcal{M}_{th}$		$\mathcal{M}_{th^*}$
	MV	EB	MV	EB	MV	EB	EB
N	38.44 [11.06]	38.43 [0.55]	40.10 [28.60]	39.77 [2.39]	39.91 [27.19]	39.56 [2.20]	46.62 [21.62]
$\beta_0$	0.62 (2.10)	1.18 (0.31)	-2.02 (4.44)	-1.01 (0.52)	0.88 (1.97)	-1.01 (0.52)	-0.51 (0.72)
$\beta_{\text{sexo}}$	-	-	0.85 (1.91)	0.80 (0.32)	-2.04 (3.58)	0.86 (0.32)	1.09 (0.53)
$\beta_{\text{idade}}$	-	-	-1.97 (3.51)	-1.45 (0.53)	0.16 (0.37)	-1.58 (0.55)	-1.78 (0.84)
$\beta_{\text{peso}}$	-	-	0.16 (0.37)	0.08 (0.05)	-1.18 (3.08)	0.10 (0.05)	0.22 (0.01)
$\beta_{\text{tempo}_1}$	-1.07 (2.90)	-0.99 (0.45)	-	-	-0.59 (3.02)	-1.22 (0.48)	-1.37 (0.52)
$\beta_{\text{tempo}_2}$	-0.54 (2.90)	-0.43 (0.45)	-	-	-1.06 (3.08)	-0.61 (0.47)	-0.67 (0.50)
$\beta_{\text{tempo}_3}$	-0.96 (2.90)	-0.87 (0.45)	-	-	-0.71 (3.02)	-1.11 (0.48)	-1.22 (0.52)
$\beta_{\text{tempo}_4}$	-0.64 (2.90)	-0.56 (0.45)	-	-	-0.12 (3.02)	-0.74 (0.47)	-0.81 (0.50)
$\beta_{\text{tempo}_5}$	-0.11 (2.90)	0.001 (0.45)	-	-	-1.46 (4.81)	-0.17 (0.48)	-0.13 (0.50)

Os resultados para as estimativas dos parâmetros dos diferentes submodelos do modelo  $\mathcal{M}_{th^*}$ , considerando o método de máxima verosimilhança, foram obtidos utilizando o programa MARK 4.3 (White e Burnham, 1999).

Verificou-se que as estimativas bayesianas para o tamanho da população são

mais precisas. O coeficiente de variação associado à estimativa bayesiana para  $N$  é, pelo menos, dez vezes inferior ao CV obtido utilizando o método de máxima verossimilhança, quando se utilizam os modelos  $\mathcal{M}_t$ ,  $\mathcal{M}_h$  e  $\mathcal{M}_{th}$ .

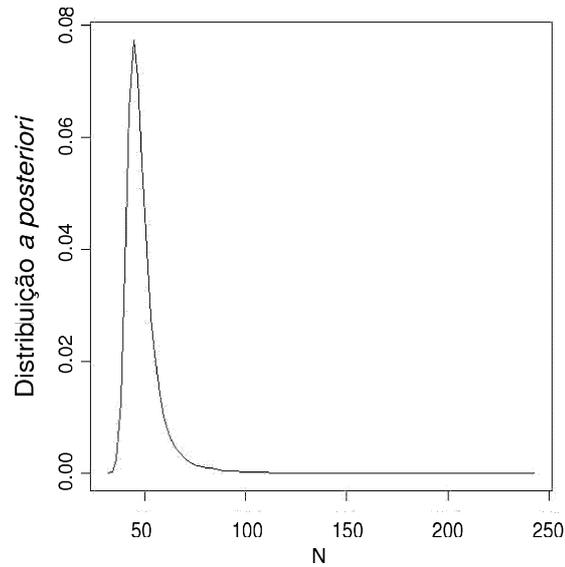


Figura 2.6: Função distribuição *a posteriori* do parâmetro  $N$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ .

Apesar de as estimativas para os parâmetros que definem as probabilidades de captura serem diferentes quando se utilizam as duas técnicas de inferência, não existe muita diferença entre as estimativas para o tamanho da população. O modelo  $\mathcal{M}_t$  apresenta a menor estimativa para  $N$ , sendo igual a 38.44 com um CV igual a 11.06%, quando se utiliza o método de máxima verossimilhança, e igual a 38.43 com um CV de 0.55%, para inferência bayesiana.

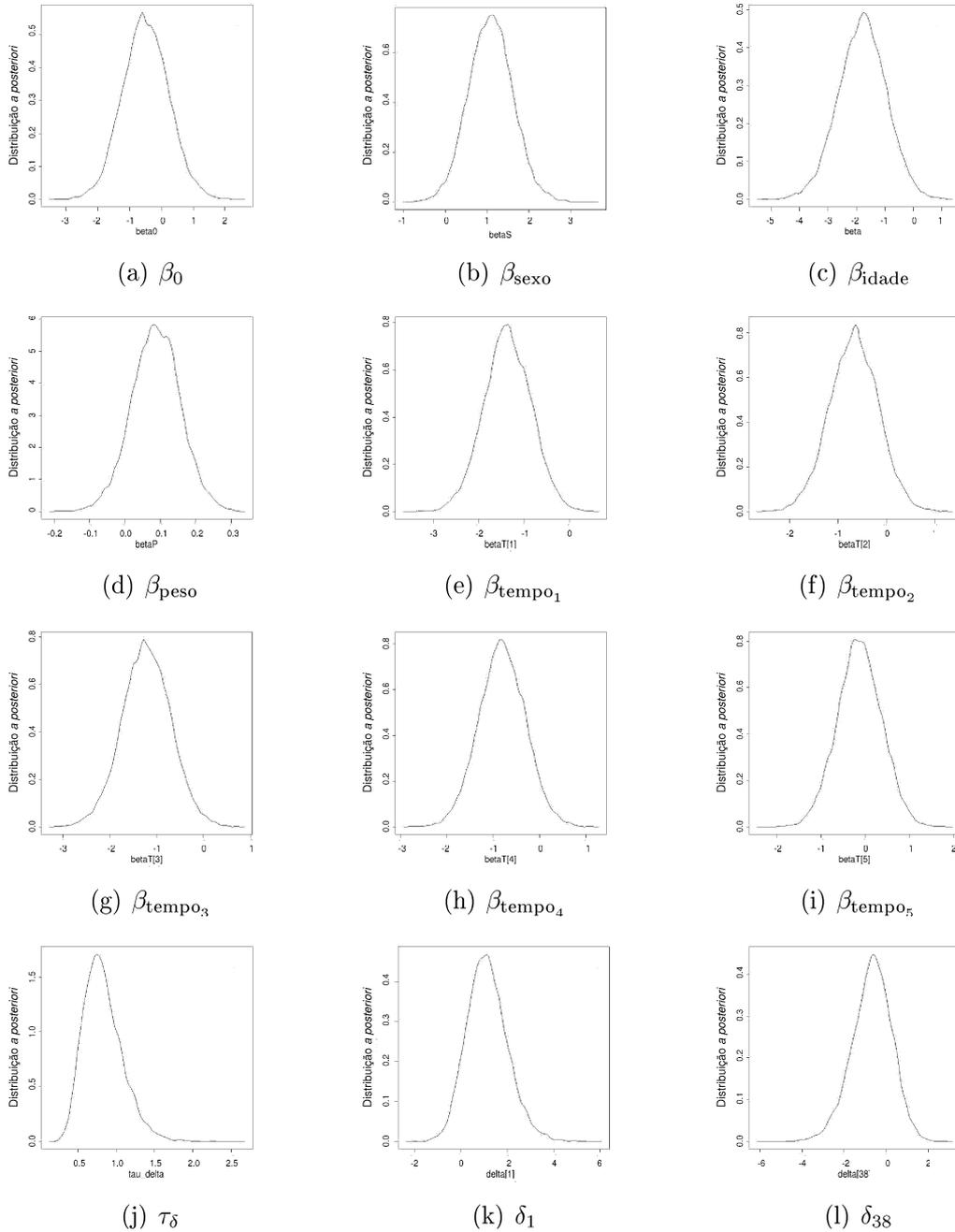


Figura 2.7: Função distribuição *a posteriori* dos parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\beta_{\text{tempo}_1}$ ,  $\beta_{\text{tempo}_2}$ ,  $\beta_{\text{tempo}_3}$ ,  $\beta_{\text{tempo}_4}$ ,  $\beta_{\text{tempo}_5}$ ,  $\tau_{\delta}$ ,  $\delta_1$  e  $\delta_{38}$ , quando o modelo utilizado é do tipo  $\mathcal{M}_{th^*}$ .

A estimativa obtida para o tamanho da população com o modelo  $\mathcal{M}_{th^*}$  (46.62) é superior à obtida com o método de máxima verosimilhança, utilizando o modelo  $\mathcal{M}_{th}$  (39.91). O intervalo HPD de 95% para o parâmetro  $N$ , utilizando o modelo  $\mathcal{M}_{th^*}$ , é  $[39.85, 65.38]$ , apresentando uma assimetria positiva, como se pode verificar na distribuição *a posteriori* de  $N$  (Figura 2.6). Os outros parâmetros, com a exceção de  $\tau_\delta$ , que apresenta uma ligeira assimetria positiva, apresentam uma distribuição *a posteriori* aproximadamente simétrica (Figura 2.7).

## Capítulo 3

# Estimación bayesiana de $g_0$ em amostragem por distâncias

Parte deste capítulo encontra-se sob a forma de publicações:

Monteiro, J. F. G. e Alpizar-Jara, R. (2010). Bayesian estimation of probability at distance zero using a double platform. (*em preparação*)

Monteiro, J. F. G., Alpizar-Jara, R., Serrano, E., Crampe, J. P. e Pérez, J. M. (2008). Estimación bayesiana de  $g_0$  usando el muestreo por distancias y su aplicación en las estimas de densidad de ungulados de montaña. Em *Tendencias Actuales en el Estudio y Conservación de los Caprinos Europeos*. (Torres, J. E. G., León, J. C.-M., Paris, P. F. e de Llano Aguilar, R. C. eds). II Congreso Internacional del género CAPRA en Europa, Jaén, 207–217 pp.

Monteiro, J. F. G. e Alpizar-Jara, R. (2006). Estimación bayesiana de  $g_0$  em amostragem por distâncias. Em *Ciência Estatística - Actas do XIII Congresso Anual da SPE*. (Canto e Castro, L., Martins, E. G., Rocha, C., Oliveira, M. F., Leal, M. M. e Rosado, F. eds). Sociedade Portuguesa de Estatística, Ericeira, 501–510 pp.

### 3.1 Introdução

Um dos métodos mais utilizados para estimar a densidade de populações animais é a amostragem por distâncias (Buckland *et al.*, 2001). Alpizar-Jara e Pollock (1996), entre outros autores, têm generalizado o modelo de amostragem por distâncias convencional, considerando detecções de animais por mais do que um observador. Estes autores consideraram um modelo que combina a informação das distâncias obtidas por amostragem de transectos lineares com as detecções dos animais pelos vários observadores, como analogia às múltiplas capturas nos modelos de captura-recaptura (Borchers *et al.*, 2002). Este modelo permite diferentes probabilidades de detecção entre os observadores e relaxar o pressuposto chave, i.e., todos os animais na linha do transecto são detectados com probabilidade 1,  $g_0 = 1$ . Também permite estimar o parâmetro  $g_0$  e testar a hipótese de que  $g_0 = 1$ .

O modelo combinado de transectos lineares e captura-recaptura tem sido estudado por alguns autores para estimar  $g_0$  (e.g., Alpizar-Jara e Pollock, 1996; Borchers *et al.*, 1998b; Chen, 2000; Manly *et al.*, 1996; Okamura *et al.*, 2003; Quang e Becker, 1997; Schweder *et al.*, 1999), mas nenhum deles utilizou a abordagem bayesiana.

Neste capítulo, o parâmetro  $g_0$  é estimado através do modelo combinado (Alpizar-Jara e Pollock, 1999), designado a partir daqui por MCAP, utilizando a inferência bayesiana. Estima-se a mediana *a posteriori* dos parâmetros que definem a função de detecção, utilizando o método Monte Carlo, baseado em cadeias de Markov implementado no WINBUGS.

A metodologia é ilustrada com dois exemplos. O primeiro exemplo utiliza os dados da experiência realizada por Otto e Pollock (1990) com latas de cervejas e o

segundo exemplo baseia-se nos dados, do ano 2002, de rebecos no Parque Nacional dos Perinéus (Monteiro *et al.*, 2008).

### 3.2 Modelo de transectos lineares

O modelo de transectos lineares convencional permite estimar o tamanho de uma determinada população fechada sob os pressupostos de que: todos os indivíduos em cima da linha de observação são detectados com probabilidade 1 ( $g_0 = 1$ ), os animais são detectados antes de efectuarem qualquer movimento em relação ao observador, a detecção de um indivíduo é independente do outro e não existem erros de medição.

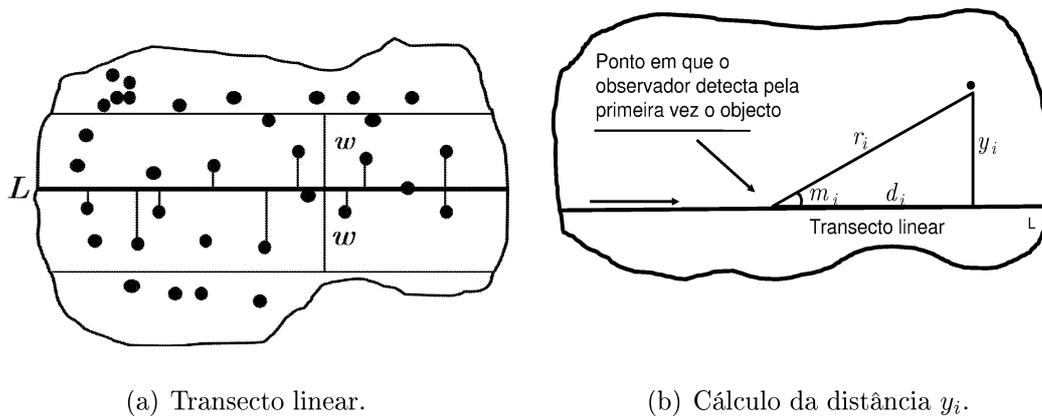


Figura 3.1: Transecto linear ao longo da faixa de largura  $2w$  e comprimento  $L$ , a). Os objectos estão representados por  $\bullet$  e os segmentos de recta ilustram as distâncias perpendiculares,  $y_i$ , aos objectos detectados. Ao longo do transecto medem-se as distâncias  $r_i$  e os ângulos  $m_i$  para calcular as distâncias perpendiculares  $y_i$ , b). Note-se que  $d_i$  é a distância horizontal entre o observador e o indivíduo  $i$ . Adaptado de Buckland *et al.* (1993, p. 5).

De uma forma aleatória, traçam-se linhas rectas ao longo de uma área de interesse. Cada uma dessas linhas é então percorrida de uma ponta à outra, à medida

que se vão detectando os objectos de ambos os lados da linha central. Tipicamente, os animais são detectados ao longo de uma área de forma rectangular, de dimensão  $a = 2Lw$ , onde  $L$  representa o comprimento e  $2w$  a largura do transecto (Buckland *et al.*, 2001). Aos  $n_2$  indivíduos detectados, mede-se a distância ( $y_i$ ,  $i = 1, \dots, n_2$  e  $\mathbf{y} = (y_1, \dots, y_{n_2})$ ) perpendicular à linha do transecto (Figura 3.1(a)). As distâncias podem também ser obtidas através das duas grandezas conhecidas  $r_i$  e  $m_i$  (Figura 3.1(b)), dado que:

$$y_i = r_i \operatorname{sen}(m_i)$$

$$d_i = r_i \operatorname{cos}(m_i).$$

Após a obtenção das distâncias perpendiculares, constrói-se um histograma de frequências e é ajustada uma função de detecção,  $g^*(y)$ , que representa a probabilidade de detectar um indivíduo, dado que este se encontra a uma distância perpendicular  $y$  da linha central do transecto.

Uma função de detecção, diga-se  $g^*(y|\sigma)$ , definida por um único parâmetro,  $\sigma$ , e truncada no intervalo  $[0, w]$ , não é necessariamente uma função densidade de probabilidade, porque pode não integrar 1 no seu domínio. Assim, a função  $g^*(y|\sigma)$  é normalizada por  $\mu^* = \int_0^w g^*(y|\sigma) dy$ , para se obter uma função densidade de probabilidade,  $f(y)$ , das distâncias observadas,  $y$ , dada por:

$$f(y) = \frac{g^*(y)}{\mu^*}.$$

A partir da função densidade de probabilidade das distâncias detectadas é possível estimar  $\mu^*$ , e conseqüentemente a probabilidade de observar um indivíduo no

transecto linear,  $p_2^* = \mu^*/w$ . Assim, um estimador da abundância na área coberta,  $a$ , baseado unicamente nas distâncias, é dado por  $\hat{N}_{TL} = n_2/\hat{p}_2^*$ . A forma deste estimador supõe que os pressupostos básicos da metodologia de transecto linear se cumprem (Buckland *et al.*, 2001, p. 29-39).

### 3.2.1 Pressuposto crítico ( $g_0 = 1$ )

O pressuposto mais crítico na amostragem por distâncias convencional é que a probabilidade de detectar um indivíduo, dado que este está sobre a linha central do transecto, é igual a 1, i.e.  $g_0 = g^*(0) = 1$ . A violação deste pressuposto pode ocorrer devido a pelo menos duas fontes de enviesamento: percepção (visibilidade e ambiente) e disponibilidade (Marsh e Sinclair, 1989). Neste trabalho, restringe-se a atenção apenas ao enviesamento por visibilidade, que pode ocorrer em consequência da capacidade do observador em detectar os animais. Por exemplo, o observador pode ser inexperiente na área de estudo, pode não estar munido de material óptico adequado, ter deficiência na visão ou na audição, etc. O enviesamento devido ao ambiente deve-se às condições do habitat e do ambiente (e.g. topografia, vegetação, vento, precipitação, grau de luminosidade, etc). O enviesamento por causa da disponibilidade pode ocorrer pelo facto de o animal se encontrar por detrás de uma árvore ou reagir antes da presença do observador.

### 3.3 Combinaco dos modelos de Lincoln-Petersen e transectos lineares

Alpizar-Jara e Pollock (1999) propuseram um modelo que combina os mtodos de amostragem captura-recaptura e transectos lineares, e que permite relaxar o pressuposto crtico da amostragem por transectos lineares (i.e.,  $g_0 = 1$ ). O estimador da abundncia para o modelo combinado mostrou ser mais eficiente e menos enviesado que o estimador baseado em um s dos mtodos. O modelo combinado assume que os indivduos marcados so identificados univocamente; que os indivduos tm igual probabilidade de serem marcados ou avistados dentro de cada amostra, e que existe independncia entre os animais marcados e avistados. Para alm desses, o modelo tambm considera os outros pressupostos usuais do modelo de transectos lineares.

O modelo combinado considera dois momentos de amostragem para estimar o tamanho da populao, anlogo aos modelos do tipo  $\mathcal{M}_t$ , para quando  $t = 2$ . Num primeiro momento, assume-se que  $n_1$  indivduos foram previamente marcados, e durante um segundo momento de amostragem,  $n_2$  indivduos so avistados. No segundo momento de amostragem, medem-se as distncias, utilizando o protocolo de amostragem por transectos lineares. Designa-se por  $n_{11}$  o nmero de indivduos avistados que foram marcados previamente;  $n_{01}$  o nmero de indivduos avistados mas que no foram marcados previamente, e  $n_{10}$  o nmero de indivduos conhecidos como marcados mas que no foram avistados. O total dos indivduos, marcados ou avistados, designa-se por  $n (= n_{11} + n_{10} + n_{01})$ , e  $\mathbf{n} = (n_{10}, n_{01}, n_{11})$  corresponde ao vector que contm a informao sobre as capturas.

Define-se a função de detecção por:

$$g(y|\sigma) = g_0 g^*(y|\sigma), \quad 0 \leq y \leq w, \quad \sigma > 0, \quad (3.1)$$

onde os indivíduos na distância zero não são necessariamente avistados com probabilidade 1 (i.e.,  $0 < g_0 \leq 1$ ). O parâmetro fulcral é agora um parâmetro de escala considerando que  $g^*(0|\sigma) = 1$ .

### 3.3.1 Função de verosimilhança

A função de verosimilhança para o modelo combinado (Alpizar-Jara e Pollock, 1999) é dada por:

$$\begin{aligned} f(\mathbf{n}, \mathbf{y}|N, p_1, g_0, \sigma) &= \frac{N!}{n_{10}!n_{01}!n_{11}!(N-n)!} [p_1 (1 - p_2(g_0, \sigma))]^{n_{10}} [(1 - p_1)p_2(g_0, \sigma)]^{n_{01}} \\ &\times [p_1 p_2(g_0, \sigma)]^{n_{11}} [(1 - p_1)(1 - p_2(g_0, \sigma))]^{N-n} \prod_{i=1}^{n_2} \frac{g^*(y_i|\sigma)}{\mu^*}, \\ &0 < p_1, p_2(g_0, \sigma) < 1, \quad N \geq n, \end{aligned} \quad (3.2)$$

onde  $p_1$  e  $p_2(g_0, \sigma)$  correspondem, respectivamente, à probabilidade de um indivíduo ter sido marcado e à probabilidade de um indivíduo ter sido avistado.

Os autores acima referidos definiram (3.2) como sendo um produto de três funções, definidas a seguir por (3.3), (3.4) e (3.5) :

$$f(\mathbf{n}, \mathbf{y}|N, p_1, g_0, \sigma) = f_{LP}^1(N, p(g_0, \sigma)|n) f_{LP}^2(p_1, p_2(g_0, \sigma)|n_{10}, n_{01}, n_{11}) f_{TL}(\sigma|\mathbf{y}, n_2),$$

onde  $p(g_0, \sigma) = p_1 + p_2(g_0, \sigma) - p_1p_2(g_0, \sigma)$ ,

$$f_{LP}^1(N, p(g_0, \sigma)|n) = \frac{N!}{n!(N-n)!} p(g_0, \sigma)^n (1 - p(g_0, \sigma))^{(N-n)},$$

$$0 < p(g_0, \sigma) < 1, \quad N \geq n, \quad (3.3)$$

$$f_{LP}^2(p_1, p_2(g_0, \sigma)|n_{10}, n_{01}, n_{11}) = \frac{n!}{n_{10}!n_{01}!n_{11}!} \left[ \frac{p_1(1 - p_2(g_0, \sigma))}{p(g_0, \sigma)} \right]^{n_{10}}$$

$$\times \left[ \frac{(1 - p_1)p_2(g_0, \sigma)}{p(g_0, \sigma)} \right]^{n_{01}} \left[ \frac{p_1p_2(g_0, \sigma)}{p(g_0, \sigma)} \right]^{n_{11}},$$

$$0 < p(g_0, \sigma) < 1, \quad (3.4)$$

e

$$f_{TL}(\sigma|\mathbf{y}, n_2) = \prod_{i=1}^{n_2} \frac{g^*(y_i|\sigma)}{\mu^*}. \quad (3.5)$$

O estimador para o tamanho da população, obtido a partir da expressão (3.2), designa-se por estimador de máxima verosimilhança completa,  $\hat{N}_{MCAP}$ . Aquele que for obtido exclusivamente a partir de (3.3) ou (3.4), sem tomar em conta a informação proporcionada pelas distâncias, denomina-se por estimador de Lincoln-Petersen,  $\hat{N}_{LP}$ , e o que for meramente estimado por (3.5) entende-se como  $\hat{N}_{TL}$ .

Assim, o estimador para a probabilidade de avistar um indivíduo, utilizando MCAP, é dado por  $\hat{p}_2(\hat{g}_0, \hat{\sigma}) = \hat{g}_0\hat{p}_2^*(\hat{\sigma})$ , e o estimador do tamanho da população é dado por  $\hat{N}_{MCAP} = \hat{N}_{TL}/\hat{g}_0$ . Se  $g_0 = 1$ , o estimador obtido com o modelo combinado é equivalente ao estimador da amostragem por distâncias convencional. Para estimar  $g_0$ , ambas as partes da informação são necessárias: dados da captura-recaptura e as

distâncias dos indivíduos avistados.

A função de verossimilhança apresentada em (3.2) não é sobreparametrizada. Uma estatística suficiente mínima é obtida a partir de  $(n_{10}, n_{01}, n_{11}, y_1, \dots, y_{n_2})$ , que tem dimensão  $n_2 + 3$ . Uma vez que no modelo existem apenas 4 parâmetros desconhecidos  $(N, p_1, g_0, \sigma)$ , todos podem ser estimados.

Habitualmente, o modelo de transecto linear é usado para estimar densidades de populações, enquanto que o modelo de captura-recaptura tem sido mais utilizado para estimar a abundância populacional. Quando os dois modelos são combinados, o tamanho da população pode ser estimado através da função de verossimilhança dada pela equação (3.2).

A função de verossimilhança pode ser achatada na vizinhança de 1, e podem surgir problemas de convergência quando se maximiza a equação (3.2) para valores de  $g_0$  perto de 1 (Monteiro, 2001; Monteiro e Alpizar-Jara, 2006). Propõe-se uma abordagem bayesiana de estimação para evitar os problemas relacionados com a otimização computacional. A inferência bayesiana permite incorporar informação *a priori* nos parâmetros de interesse e a distribuição *a posteriori* dá uma interpretação probabilística dos parâmetros de interesse.

### 3.4 Estimação bayesiana de $g_0$

Apresenta-se uma abordagem alternativa sob a perspectiva bayesiana para estimar  $g_0$ , usando informação das marcas e distâncias, i.e.  $(\mathbf{n}, \mathbf{y})$ . Pretende-se estimar, para além de  $g_0$ , o tamanho da população,  $N$ , bem como a probabilidade de um indivíduo ter sido marcado,  $p_1$ , e a probabilidade de um indivíduo ter sido avistado

$p_2(g_0, \sigma)$ . A probabilidade de um indivíduo ser avistado define-se em função dos parâmetros  $g_0$  e  $\sigma$ , onde o último define a função de detecção, que neste caso se assume como sendo a Semi-Normal, isto é:

$$g(y|g_0, \sigma) = g_0 e^{-\frac{y^2}{2\sigma^2}}, \quad 0 \leq y \leq w, \quad \sigma > 0, \quad 0 < g_0 \leq 1. \quad (3.6)$$

### 3.4.1 Distribuições *a priori*

Na inferência bayesiana é necessário definir uma distribuição *a priori* para o vector de parâmetros desconhecidos,  $\boldsymbol{\theta} = (N, \lambda, g_0, \sigma^2, p_1)$ , onde  $\lambda$  é o hiperparâmetro que define a distribuição *a priori* para  $N$ , que será apresentada posteriormente. Assumindo que as componentes do vector são independentes, a seguir formulam-se distribuições *a priori* para cada um dos parâmetros envolvidos.

Primeiro, há que ter em conta que  $0 < p_1, p_2 \leq 1$  e  $0 < g_0 \leq 1$ , pois tratam-se de probabilidades. Para  $N$  fixo, a partir da função de verosimilhança dada em (3.2) tem-se que:

$$\frac{\partial^2 \log f(p_1, p_2|N)}{\partial p_1^2} = -\frac{n_1}{p_1^2} - \frac{N - n_1}{(1 - p_1)^2} \quad (3.7)$$

$$\frac{\partial^2 \log f(p_1, p_2|N)}{\partial p_2^2} = -\frac{n_2}{p_2^2} - \frac{N - n_2}{(1 - p_2)^2} \quad (3.8)$$

$$\frac{\partial^2 \log f(p_1, p_2|N)}{\partial p_1 \partial p_2} = 0, \quad (3.9)$$

e tendo em conta que  $E[n_1] = Np_1$  e  $E[n_2] = Np_2$ , de (3.7), (3.8) e (3.9) tem-se que

$$I(p_1, p_2) = \text{Diag} \left( \frac{1}{p_1(1 - p_1)}, \frac{1}{p_2(1 - p_2)} \right),$$

e portanto, a distribuição de Jeffreys, tal e qual expressa em (1.3), é dada por

$$h(p_1|N, g_0, \sigma) \propto p_1^{-\frac{1}{2}}(1 - p_1)^{-\frac{1}{2}},$$

que, de uma forma geral, pode-se considerar a seguinte classe de distribuição *a priori*:

$$h(p_1|N, g_0, \sigma) \propto p_1^{a-1}(1 - p_1)^{b-1}, \quad (3.10)$$

onde  $a$  e  $b \in \mathbb{R}$ . Em particular, quando  $a, b > 0$ , (3.10) representa o núcleo de uma distribuição  $Beta(a, b)$ .

Wang (2002) sugeriu que uma distribuição *a priori* plausível para  $p_1$  é a Uniforme,

$$p_1 \sim Unif(0, 1), \quad (3.11)$$

que é um caso particular da distribuição Beta, para  $a = b = 1$ . Assim, assume-se uma distribuição não informativa para  $p_1$ .

Para o tamanho da população, assumiu-se uma distribuição *a priori* hierárquica do tipo Poisson-Gama, também proposta por George e Robert (1992). No primeiro nível da distribuição *a priori*, assume-se que  $N$  tem distribuição Poisson com média  $\lambda$ ,  $\lambda > 0$ , i.e.:

$$N|\lambda \sim Poi(\lambda), \quad N = 1, 2, \dots, \quad (3.12)$$

e no segundo nível tem-se que  $\lambda$  tem distribuição *a priori* Gama com hiperparâmetros  $\tau$  e  $\eta$ , i.e.:

$$\lambda \sim Gama(\tau, \eta), \lambda > 0, \quad (3.13)$$

onde os hiperparâmetros  $\tau$  e  $\eta$  são ambos reais positivos e predeterminados a partir do valor esperado e da variância de  $\lambda$ , tendo em conta a informação disponível. Seja  $\nu_0$  um valor que se pensa ser o tamanho da população tendo em conta a informação disponível. Considerando um CV de 10% para  $\nu_0$ , tem-se que  $E[\lambda] = 100/\nu_0$  e  $Var[\lambda] = 0.01 (E[\lambda])^2$ . Os valores para os hiperparâmetros  $\tau$  e  $\eta$  são calculados a partir destas duas últimas igualdades.

A distribuição *a priori* marginal de  $N$ , que se obtém tendo em conta as equações (3.12) e (3.13) e integrando em  $\lambda$ , é dada por:

$$\begin{aligned} \int_0^\infty \frac{\lambda^N}{N!} e^{-\lambda} \frac{\eta^\tau \lambda^{\tau-1}}{\Gamma(\tau)} e^{-\eta\lambda} d\lambda &= \frac{\eta^\tau}{N! \Gamma(\tau)} \int_0^\infty \lambda^{N+\tau-1} e^{-\lambda(\eta+1)} d\lambda \\ &\propto \frac{1}{N!} \int_0^\infty \lambda^{N+\tau-1} e^{-\lambda(\eta+1)} d\lambda \\ &= \frac{\Gamma(N+\tau)}{(\eta+1)^{N+\tau} N!} \int_0^\infty \frac{(\eta+1)^{N+\tau} \lambda^{N+\tau-1} e^{-\lambda(\eta+1)}}{\Gamma(N+\tau)} d\lambda \\ &= \frac{\Gamma(N+\tau)}{(\eta+1)^{N+\tau} N!}. \end{aligned} \quad (3.14)$$

Quando a informação sobre o parâmetro  $\lambda$  é escassa, sugere-se uma distribuição *a priori* não informativa, pois da distribuição *a priori* de  $\lambda$ , dado pela equação (3.13),

tem-se que:

$$h(\lambda) \propto \lambda^{\tau-1} e^{-\eta\lambda} \text{ e } \lim_{\substack{\tau \rightarrow 0 \\ \eta \rightarrow 0}} \lambda^{\tau-1} e^{-\eta\lambda} = \frac{1}{\lambda},$$

e da distribuição *a priori* marginal de  $N$ , (3.14), tem-se que  $h(N) \rightarrow 1/N$  quando  $\tau, \eta \rightarrow 0$ , que é uma distribuição *a priori* imprópria. Nos dois exemplos que se apresentam nas secções 3.6 e 3.7, não se põe esse tipo de problema, porque existe informação suficiente sobre o tamanho da população. Mas há que levar em conta que esta distribuição *a priori*, sugerida para  $N$ , pode ser questionada quando há pouca informação sobre o tamanho da população.

Considerando  $N$ ,  $p_1$  e  $\sigma$  como sendo fixos, tendo em conta que  $E[n_2] = Ng_0\mu^*/w$  e a função de verosimilhança dada em (3.2) vem que:

$$\frac{\partial^2 \log f(g_0|N, p_1, \sigma)}{\partial g_0^2} = -\frac{n_2}{g_0^2} - \frac{(N - n_2)(\frac{\mu^*}{w})^2}{(1 - \frac{\mu^*}{w}g_0)^2}, \quad (3.15)$$

e de (3.15) conclui-se que a distribuição *a priori* de Jeffreys pode-se escrever como sendo

$$h(g_0|N, p_1, \sigma) \propto g_0^{-\frac{1}{2}} \left(1 - \frac{\mu^*}{w}g_0\right)^{-\frac{1}{2}}.$$

Para  $g_0$  adoptou-se a distribuição Uniforme,

$$g_0 \sim Unif(0, 1). \quad (3.16)$$

Utilizou-se a distribuição Gama Invertida (GI) para definir a distribuição *a priori*

para  $\sigma^2$ , i.e.:

$$\sigma^2 \sim GI(\alpha, \beta), \quad (3.17)$$

porque é a distribuição *a priori* conjugada para a inversa da variância da distribuição Normal (Gelman *et al.*, 2004). Esta distribuição também tem sido proposta por Eguchi e Gerrodette (2009) e Karunamuni e Quinn (1995), na estimação bayesiana da densidade populacional em amostragem por distâncias. Os hiperparâmetros  $\alpha$  e  $\beta$  são determinados a partir da informação disponível sobre variabilidade das distâncias, uma vez que  $\sigma$  é o parâmetro de escala da função de detecção. Para um valor estipulado,  $\xi_0$ , a partir da informação sobre a variabilidade das distâncias, e considerando um CV de 10% para  $\xi_0$ , tem-se que  $E[\sigma^2] = 100/\xi_0$  e  $Var[\sigma^2] = 0.01 (E[\sigma^2])^2$ . Os valores para os hiperparâmetros  $\alpha$  e  $\beta$  são calculados a partir destas duas últimas igualdades.

### 3.4.2 Distribuições *a posteriori*

Assumiu-se que  $\mu^* \simeq \sigma K$  tendo em conta que a função de detecção é a Semi-Normal definida em (3.6), onde  $K$  é constante e é calculada com base na informação do valor esperado de  $p_2^*$ . A função de verosimilhança dada em (3.2) pode ser reescrita

da seguinte maneira:

$$\begin{aligned}
 f(\mathbf{n}, \mathbf{y} | \boldsymbol{\theta}) &= \frac{N!}{n_{10}!n_{01}!n_{11}!(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \\
 &\times \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} \\
 &\propto \frac{N!}{(N-n)!} p_1^{n_1} (1-p_1)^{N-n_1} (\sigma g_0)^{n_2} \left( \frac{w}{K} - \sigma g_0 \right)^{N-n_2} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{\sigma^{n_2}}, \\
 N &\geq n.
 \end{aligned} \tag{3.18}$$

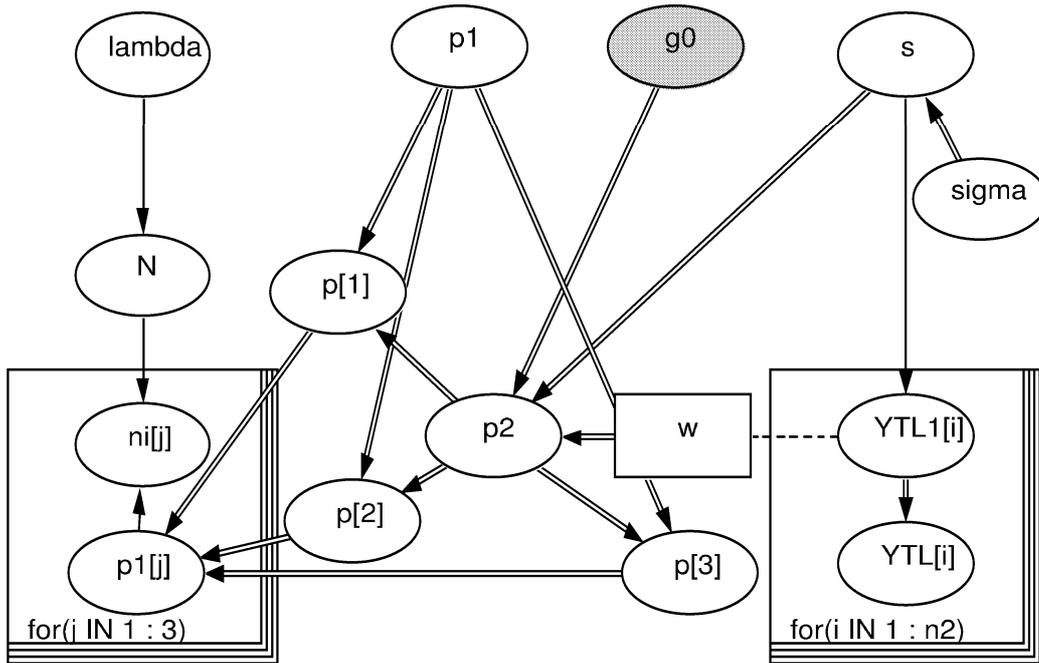


Figura 3.2: DAG do modelo combinado de captura-recaptura e transectos lineares (MCNH), para estimar a probabilidade de detectar um indivíduo na distância zero,  $g_0$ .

Seja  $\boldsymbol{\theta}$  o vector de parâmetros desconhecidos,  $\boldsymbol{\theta} = (N, \lambda, g_0, \sigma^2, p_1)$ , e sejam  $\mathbf{n}$  e

$\mathbf{y}$  as observações (ver DAG do modelo combinado na Figura 3.2), sob o pressuposto de independência entre as componentes do vector  $\boldsymbol{\theta}$ , a distribuição *a posteriori* de  $\boldsymbol{\theta}$  dado  $(\mathbf{n}, \mathbf{y})$ , é proporcional ao produto entre a função de verosimilhança (3.18) e as distribuições *a priori* (3.11), (3.12), (3.13), (3.16) e (3.17):

$$\begin{aligned}
 h(\boldsymbol{\theta}|\mathbf{n}, \mathbf{y}) &\propto f(\mathbf{n}, \mathbf{y}|\boldsymbol{\theta})h(N|\lambda)h(\lambda)h(p_1)h(g_0)h(\sigma^2) \propto \frac{N!}{(N-n)!} p_1^{n_1} (1-p_1)^{N-n_1} \\
 &\times (\sigma g_0)^{n_2} \left( \frac{w}{K} - \sigma g_0 \right)^{N-n_2} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{\sigma^{n_2}} \frac{e^{-\lambda} \lambda^N}{N!} \frac{\eta^\tau \lambda^{\tau-1} e^{-\eta\lambda}}{\Gamma(\tau)} \\
 &\times \frac{\beta^\alpha \left(\frac{1}{\sigma^2}\right)^{\alpha-1} e^{-\frac{\beta}{\sigma^2}}}{\Gamma(\alpha)}. \tag{3.19}
 \end{aligned}$$

O método de amostragem Gibbs baseia-se num algoritmo de simulação, que permite gerar amostras da distribuição *a posteriori*, (3.19), utilizando distribuições *a posteriori* condicionais completas de cada um dos parâmetros não observados.

**Proposição 3.1.** *A distribuição a posteriori condicional completa de  $(N, \lambda, g_0, \sigma^2, p_1)$  dada a informação sobre as capturas,  $\mathbf{n}$ , e as distâncias,  $\mathbf{y}$ , é dada por:*

- a)  $h(g_0|\mathbf{n}, \mathbf{y}, N, \lambda, \sigma^2, p_1) \propto g_0^{n_2} \left( \frac{w}{K\sigma} - g_0 \right)^{N-n_2}$ ;
- b)  $N \sim Poi(\lambda(1-p_1)(1-p_2(g_0, \sigma)))$ ;
- c)  $\lambda \sim Gama(N + \tau, 1 + \eta)$ ;
- d)  $h(\sigma^2|\mathbf{n}, \mathbf{y}, g_0, N, \lambda, p_1) \propto \left( \frac{w}{Kg_0} - \sigma \right)^{N-n_2} \frac{e^{-\frac{\beta + \sum_{i=1}^{n_2} \frac{y_i^2}{2}}{\sigma^2}}}{\sigma^{2(\alpha-1)}}$ ;
- e)  $p_1 \sim Beta(n_1 + 1, N - n_1 + 1)$ ;

*Demonstração.*

a) Tendo em conta a função de verosimilhança dada em (3.18) e a distribuição *a priori* (3.16) tem-se que:

$$h(g_0|\mathbf{n}, \mathbf{y}, N, \lambda, \sigma^2, p_1) \propto f(\mathbf{n}, \mathbf{y}|\boldsymbol{\theta})h(g_0)$$

$$\begin{aligned} &\propto \frac{N!}{(N-n)!} p_1^{n_1} (1-p_1)^{N-n_1} (\sigma g_0)^{n_2} \left(\frac{w}{K} - \sigma g_0\right)^{N-n_2} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{\sigma^{n_2}} \\ &\propto g_0^{n_2} \left(\frac{w}{K\sigma} - g_0\right)^{N-n_2}; \end{aligned}$$

b) Tendo em conta a função de verosimilhança dada em (3.18) e as distribuições *a priori* (3.12) e (3.13) tem-se que:

$$h(N|\mathbf{n}, \mathbf{y}, g_0, \lambda, \sigma^2, p_1) \propto f(\mathbf{n}, \mathbf{y}|\boldsymbol{\theta})h(N|\lambda)h(\lambda)$$

$$\begin{aligned} &\propto \frac{N!}{(N-n)!} p_1^{n_1} (1-p_1)^{N-n_1} (\sigma g_0)^{n_2} \left(\frac{w}{K} - \sigma g_0\right)^{N-n_2} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{\sigma^{n_2}} \\ &\times e^{-\lambda} \frac{\lambda^N \eta^\tau \lambda^{\tau-1} e^{-\eta\lambda}}{N! \Gamma(\tau)} \\ &\propto \frac{\lambda^N}{(N-n)!} (1-p_1)^{N-n_1} \left(\frac{w}{K} - \sigma g_0\right)^{N-n_2} \\ &\propto \frac{\lambda^N}{(N-n)!} (1-p_1)^N (1-p_2(g_0, \sigma))^N \\ &\propto \frac{\lambda^N}{(N-n)!} ((1-p_1)(1-p_2(g_0, \sigma)))^N \\ &\propto e^{\lambda(1-p_1)(1-p_2(g_0, \sigma))} \frac{(\lambda(1-p_1)(1-p_2(g_0, \sigma)))^{N-n}}{(N-n)!}; \\ &N \geq n; \end{aligned}$$

c) Tendo em conta a função de verosimilhança dada em (3.18) e as distribuições *a priori* (3.12) e (3.13) tem-se que:

$$h(\lambda|\mathbf{n}, \mathbf{y}, g_0, N, \sigma^2, p_1) \propto f(\mathbf{n}, \mathbf{y}|\boldsymbol{\theta})h(N|\lambda)h(\lambda)$$

$$\begin{aligned} &\propto \frac{N!}{(N-n)!} p_1^{n_1} (1-p_1)^{N-n_1} (\sigma g_0)^{n_2} \left(\frac{w}{K} - \sigma g_0\right)^{N-n_2} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{\sigma^{n_2}} \\ &\times e^{-\lambda} \frac{\lambda^N \eta^\tau \lambda^{\tau-1} e^{-\eta\lambda}}{N! \Gamma(\tau)} \\ &\propto e^{-\lambda(1+\eta)} \lambda^{N+\tau-1} \\ &\propto \frac{(1+\eta)^{N+\tau} \lambda^{N+\tau-1} e^{-\lambda(1+\eta)}}{\Gamma(N+\tau)}; \end{aligned}$$

d) Tendo em conta a função de verosimilhança dada em (3.18) e a distribuição *a priori* (3.17) tem-se que:

$$h(\sigma^2|\mathbf{n}, \mathbf{y}, g_0, N, \lambda, p_1) \propto f(\mathbf{n}, \mathbf{y}|\boldsymbol{\theta})h(\sigma^2)$$

$$\begin{aligned} &\propto \frac{N!}{(N-n)!} p_1^{n_1} (1-p_1)^{N-n_1} (\sigma g_0)^{n_2} \left(\frac{w}{K} - \sigma g_0\right)^{N-n_2} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{\sigma^{n_2}} \frac{\beta^\alpha \left(\frac{1}{\sigma^2}\right)^{\alpha-1} e^{-\frac{\beta}{\sigma^2}}}{\Gamma(\alpha)} \\ &\propto \left(\frac{w}{Kg_0} - \sigma\right)^{N-n_2} \frac{e^{-\frac{\beta + \sum_{i=1}^{n_2} \frac{y_i^2}{\sigma^2}}{\sigma^2}}}{\sigma^{2(\alpha-1)}}; \end{aligned}$$

e) Tendo em conta a função de verosimilhança dada em (3.18) e a distribuição *a priori* (3.11) tem-se que:

$$h(p_1|\mathbf{n}, \mathbf{y}, N, \lambda, g_0, \sigma^2) \propto f(\mathbf{n}, \mathbf{y}|\boldsymbol{\theta})h(p_1)$$

$$\begin{aligned} &\propto \frac{N!}{(N-n)!} p_1^{n_1} (1-p_1)^{N-n_1} (\sigma g_0)^{n_2} \left(\frac{w}{K} - \sigma g_0\right)^{N-n_2} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{\sigma^{n_2}} \\ &\propto p_1^{n_1} (1-p_1)^{N-n_1} \frac{\Gamma(N+2)}{\Gamma(n_1+1)\Gamma(N-n_1+1)}. \end{aligned}$$

□

## 3.5 Estudo de simulação

Para a simulação de dados foram considerados diferentes tamanhos para a população; pequeno ( $N = 150$ ), médio ( $N = 500$ ) e grande ( $N = 1000$ ). Os objectos foram aleatoriamente e homogeneamente distribuídos dentro do transecto com 200 metros de comprimento e 40 metros de largura. Para a probabilidade de captura,  $p_1$ , assumiram-se valores iguais a 0.10, 0.20 e 0.40. Foi considerada a função Semi-Normal para a função de detecção, com o parâmetro  $\sigma = 7$ . Para cada valor assumido para o parâmetro  $g_0$  (0.50, 0.70, 0.90 e 1.00) está associado um valor para a probabilidade de avistar um objecto dentro do transecto ( $p_2=0.10, 0.22, 0.39, 0.44$ ).

Com base nos valores predeterminadas para  $N$ ,  $p_1$ ,  $\sigma$ ,  $g_0$ ,  $L$ , e  $w$ , foi realizada a simulação de dados para cada caso de estudo. Os valores para  $n_{10}$ ,  $n_{01}$ , e  $n_{11}$  foram gerados de uma multinomial  $(N, p_1, p_2)$ , onde a probabilidade de avistar um objecto,  $p_2$ , é dada por  $g_0 p_2^*$ . As  $n_2$  distâncias perpendiculares foram geradas da função Semi-Normal truncada na distância  $w$ .

### 3.5.1 Análises de diagnóstico

Na Figura 3.3 verifica-se que, em diferentes cenários, as duas cadeias para o parâmetro  $g_0$  convergem ambas para a mesma região, apesar de terem pontos iniciais diferentes. Para além disso, é possível ver no gráfico do factor de redução, na Figura 3.4, que  $R$  converge para 1 em todos os cenários e por isso indica convergência para a respectiva cadeia.

O diagnóstico de autocorrelação foi realizado para todas as cadeias, indicando que um espaçamento de tamanho 20 entre iterações diminuía a autocorrelação para

valores próximos de zero. O teste de convergência de Raftery e Lewis (1995) indica que existe convergência e que tal acontece ao fim de poucas iterações, o que implica um período de aquecimento relativamente curto (Tabela 3.1).

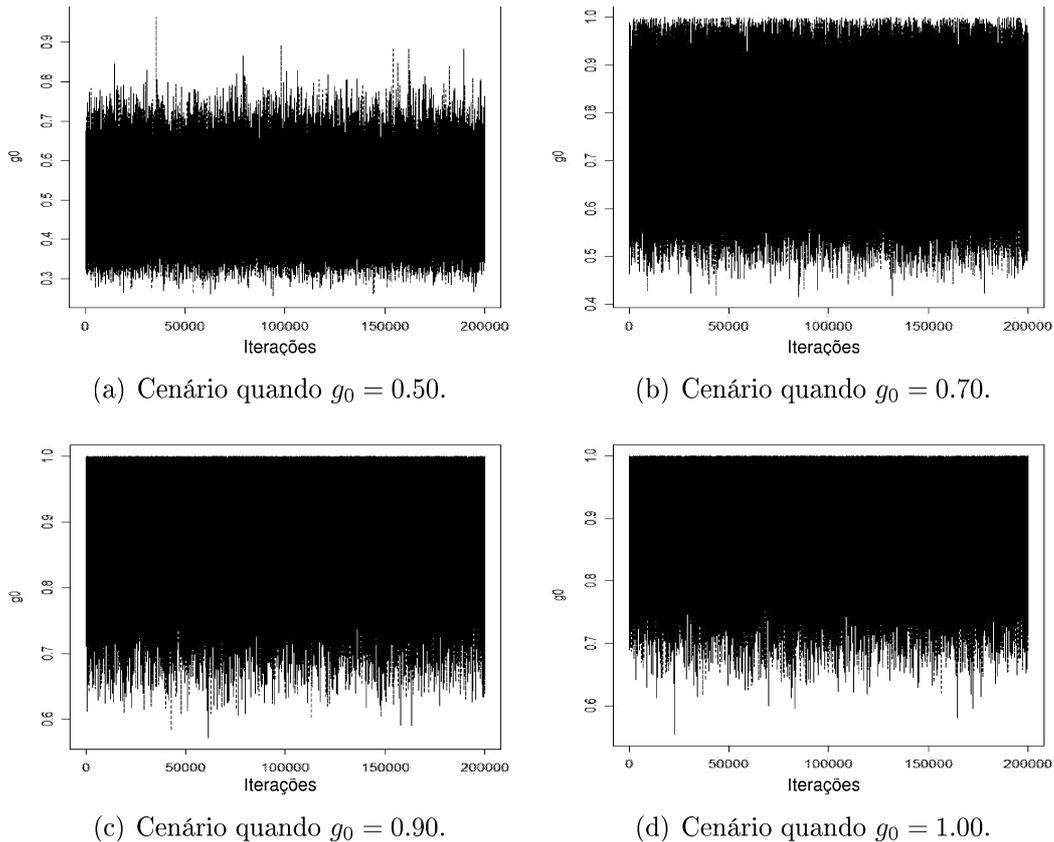


Figura 3.3: Traço das duas cadeias de Markov, com pontos iniciais diferentes, consideradas para o parâmetro  $g_0$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários considerados para a simulação de dados para  $g_0$ , i.e.  $g_0 = 0.50, 0.70, 0.90, 1.00$ .

Assim, para as estimativas bayesianas dos parâmetros, adoptou-se apenas uma das cadeias, tendo-se verificado que para as 200,000 iterações, um período de aquecimento de 1,000 iterações e um espaçamento de tamanho 20 entre iterações era adequado para assegurar convergência, remover a influência dos valores iniciais e diminuir a autocorrelação para valores próximos de zero. Adoptou-se, com base

na amostra final de tamanho 9,950, a mediana da distribuição *a posteriori* como estimativa pontual dos parâmetros e um intervalo de credibilidade de 95% para os parâmetros.

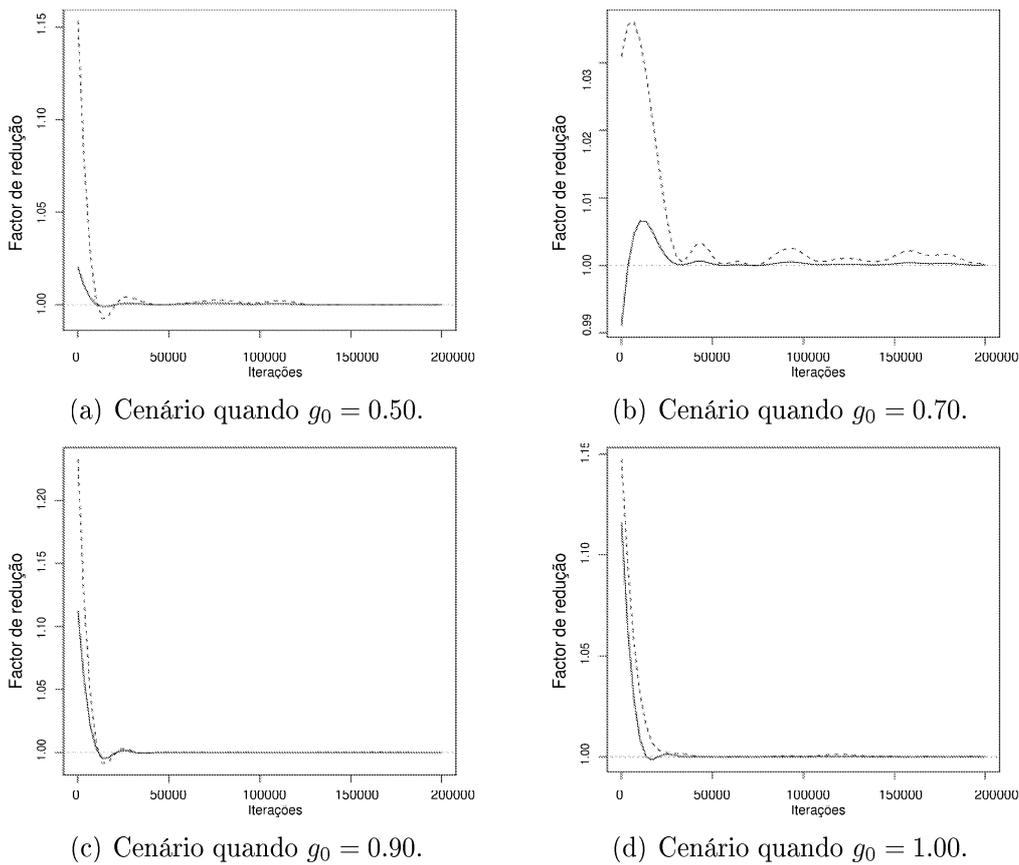


Figura 3.4: Teste de convergência Gelman e Rubin, utilizando duas cadeias de Markov com pontos iniciais diferentes, para o parâmetro  $g_0$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários considerados para a simulação de dados para  $g_0$ , i.e.  $g_0 = 0.50, 0.70, 0.90, 1.00$ . - - - - 97.5%, ——— mediana.

Tabela 3.1: Estimativa do período de aquecimento e da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50), para as cadeias de Markov dos parâmetros  $\lambda$ ,  $g_0$ ,  $\sigma^2$  e  $p_1$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários (1,2,3,4) considerados para a simulação de dados para o parâmetro  $g_0$ , respectivamente para  $g_0 = 0.50$ ,  $0.70$ ,  $0.90$  e  $1.00$ .

Parâmetros		Período de aquecimento	Função de autocorrelação			
			Espaçamento			
			5	10	20	50
$\lambda$	1	18	0.4007	0.1741	0.0340	-0.0019
	2	16	0.3666	0.1575	0.0302	-0.0003
	3	10	0.2408	0.0655	0.0051	-0.0079
	4	12	0.2459	0.0711	0.0057	0.0012
$g_0$	1	14	0.2486	0.1068	0.0231	0.0008
	2	20	0.2811	0.1198	0.0178	-0.0049
	3	16	0.1697	0.0456	0.0038	-0.0061
	4	18	0.1690	0.0457	0.0037	-0.0015
$\sigma^2$	1	6	0.0040	0.0009	0.0028	0.0007
	2	6	0.0054	-0.0001	0.0026	-0.0012
	3	8	0.0143	-0.0007	-0.0017	0.0018
	4	8	0.0162	-0.0036	-0.0040	-0.0004
$p_1$	1	27	0.2713	0.1202	0.0268	0.0014
	2	16	0.2480	0.1041	0.0226	-0.0025
	3	12	0.1327	0.0406	0.0015	-0.0031
	4	10	0.1157	0.0332	0.0039	0.0004

### 3.5.2 Resultados e discussão

A Tabela 3.2 apresenta as estimativas, coeficiente de variação (CV) e as percentagens do enviesamento relativo (%ER) para os estimadores  $\hat{N}_{MCAP}$ ,  $\hat{p}_1$  e  $\hat{p}_2$ .

De um modo geral, o estimador proposto para o tamanho da população,  $\hat{N}_{MCAP}$ , não é enviesado para a maioria dos cenários considerados, com ligeiro enviesamento

relativo entre 0% e 12%. Quando  $g_0 = 1.00$  a percentagem do enviesamento para  $\hat{N}_{MCAP}$  é  $< 4\%$ . O enviesamento é maior quando o tamanho da população é menor ( $N = 150$ ). O tamanho da população é quase sempre sobrestimado para todas as situações, excepto quando ( $g_0 = 1.00, p_1 = 0.20, p_2 = 0.44$  e  $N = 500$  ou  $N = 1000$ ) e ( $g_0 = 1.00, p_1 = 0.40, p_2 = 0.44$  e  $N = 500$ ), nos quais foi subestimado, com um enviesamento igual a -2.40, -0.30 e -0.80 respectivamente. O estimador para o tamanho da população tende a ser mais preciso nas situações onde  $N$  é grande.

O estimador  $\hat{p}_1$  apresenta uma amplitude grande para o %ER, variando entre 0.05% a 40.27%. Quando se considera  $p_1 = 0.10, p_2 = 0.22$  e  $g_0 = 0.50$  resulta que o %ER é maior para todos os estimadores. E quando o tamanho da população é grande ( $N = 1000$ ), o estimador dos parâmetros apresentam menor %ER ( $< 4\%$ ).

O estimador  $\hat{p}_1$  é menos preciso quando  $p_1 = 0.10$  e o tamanho da população é pequeno,  $N = 150$ , apresentando um CV entre 27% e 35%, e mais preciso quando  $p_1 = 0.40$  e o tamanho da população é grande,  $N = 1000$ . O estimador de  $p_2$  é menos preciso quando  $p_2 = 0.22$ , sobretudo para populações pequenas, e mais preciso quando  $p_2 = 0.44$ .

O desempenho do estimador  $\hat{g}_0$  foi analisado considerando diferentes cenários. A Tabela 3.3 resume as estimativas, o primeiro e o terceiro quartil ( $Q_1$  e  $Q_3$ ), a mediana ( $me$ ), a percentagem raiz do erro quadrático médio (%REQM), a percentagem do enviesamento relativo (%ER) e o coeficiente de variação.

O estimador do parâmetro  $g_0$  é mais preciso quando a probabilidade de detecção na distância zero é elevada e menos preciso quando  $g_0 = 0.50$ . O estimador de  $g_0$  tem mais variabilidade quando a probabilidade de captura e o tamanho da população são pequenos ( $p_1 = 0.10$  e  $N = 150$ ).



Tabela 3.3: Resumo da distribuição *a posteriori* do parâmetro  $g_0$  nos diferentes cenários considerados para a simulação de dados. Apresentam-se o primeiro quartil ( $Q_1$ ), a mediana ( $me$ ), o terceiro quartil ( $Q_3$ ), o desvio padrão ( $sd$ ), a percentagem do enviesamento relativo ( $\%ER$ ), a percentagem da raiz do erro quadrático médio ( $\%REQM$ ) e o coeficiente de variação ( $CV\%$ ).

	$p_1$	$N$	$Q_1$	$me$	$Q_3$	$sd$	$\%ER$	$\%REQM$	$CV\%$
$g_0 = 0.50$	0.10	150	0.39	0.59	0.86	0.12	-18.46	0.15	20.31
	0.10	500	0.40	0.53	0.71	0.08	-6.72	0.09	14.68
	0.10	1000	0.42	0.53	0.67	0.06	-6.46	0.07	12.09
$p_2 = 0.22$	0.20	150	0.34	0.51	0.75	0.11	-2.74	0.11	20.52
	0.20	500	0.37	0.49	0.64	0.07	2.58	0.07	14.23
	0.20	1000	0.39	0.49	0.60	0.05	2.56	0.06	11.23
$g_0 = 0.70$	0.40	150	0.30	0.46	0.67	0.10	8.96	0.11	21.13
	0.40	500	0.42	0.53	0.68	0.07	-6.60	0.07	12.23
	0.40	1000	0.41	0.50	0.60	0.05	0.96	0.05	10.03
$p_2 = 0.31$	0.10	150	0.51	0.73	0.96	0.12	-4.16	0.12	16.38
	0.10	500	0.54	0.69	0.90	0.09	1.06	0.09	13.17
	0.10	1000	0.57	0.71	0.88	0.08	-1.60	0.08	10.87
$g_0 = 0.90$	0.20	150	0.48	0.69	0.94	0.12	0.74	0.12	17.01
	0.20	500	0.57	0.73	0.92	0.09	-4.40	0.09	11.90
	0.20	1000	0.57	0.70	0.85	0.07	-0.13	0.07	10.06
$p_2 = 0.39$	0.40	150	0.51	0.72	0.95	0.11	-3.23	0.12	15.61
	0.40	500	0.59	0.73	0.89	0.08	-4.41	0.08	10.56
	0.40	1000	0.61	0.73	0.85	0.06	-3.59	0.07	8.47
$g_0 = 1.00$	0.10	150	0.66	0.88	0.99	0.09	2.06	0.09	10.30
	0.10	500	0.75	0.92	1.00	0.07	-1.97	0.07	7.30
	0.10	1000	0.76	0.91	1.00	0.06	-1.36	0.07	7.05
$p_2 = 0.44$	0.20	150	0.63	0.86	0.99	0.10	4.73	0.11	11.46
	0.20	500	0.75	0.91	1.00	0.07	-1.21	0.07	7.18
	0.20	1000	0.75	0.89	0.99	0.07	1.49	0.07	7.39
$g_0 = 1.00$	0.40	150	0.64	0.85	0.99	0.09	5.22	0.11	11.07
	0.40	500	0.72	0.88	0.99	0.07	2.64	0.07	8.11
	0.40	1000	0.76	0.89	0.99	0.06	1.41	0.06	6.68
$p_2 = 0.44$	0.10	150	0.71	0.92	1.00	0.08	8.35	0.11	8.37
	0.10	500	0.80	0.94	1.00	0.05	5.52	0.08	5.68
	0.10	1000	0.82	0.95	1.00	0.05	5.32	0.07	5.24
$g_0 = 1.00$	0.20	150	0.69	0.89	1.00	0.08	10.55	0.13	9.41
	0.20	500	0.76	0.92	1.00	0.06	8.30	0.10	6.88
	0.20	1000	0.82	0.94	1.00	0.05	5.94	0.08	5.26
$p_2 = 0.44$	0.40	150	0.68	0.89	0.99	0.08	11.16	0.14	9.49
	0.40	500	0.77	0.91	0.99	0.06	8.98	0.11	6.82
	0.40	1000	0.86	0.96	1.00	0.04	4.27	0.06	4.02

Quando se assume que  $g_0 = 1.00$ , o estimador  $\hat{g}_0$  apresenta um enviesamento positivo. Nos outros cenários considerados para  $g_0$ , ocorrem ambas situações de enviesamento positivo ou negativo.

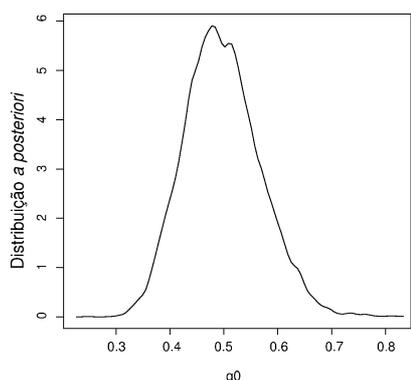
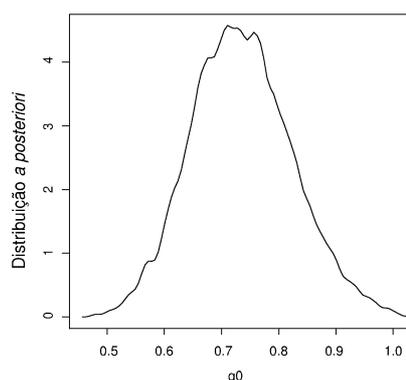
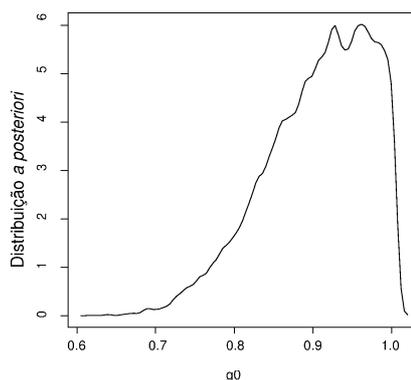
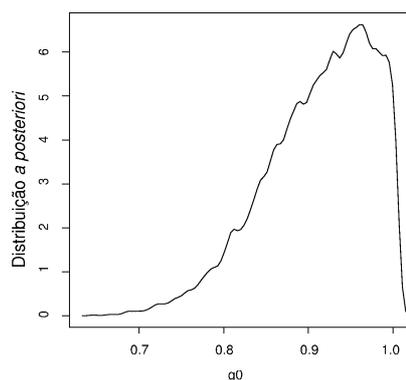
(a) Cenário quando  $g_0 = 0.50$ .(b) Cenário quando  $g_0 = 0.70$ .(c) Cenário quando  $g_0 = 0.90$ .(d) Cenário quando  $g_0 = 1.00$ .

Figura 3.5: Distribuição *a posteriori* para  $g_0$ , quando  $N = 500$  e  $p_1 = 0.20$ , para os quatro cenários considerados para a simulação de dados para  $g_0$ , i.e.  $g_0 = 0.50, 0.70, 0.90, 1.00$ .

De uma forma geral, quando  $g_0 = 0.50$  o estimador do parâmetro  $g_0$  não é enviesado, embora apresenta um enviesamento de  $-18.46\%$  quando o tamanho da população é igual a 150. Quando  $g_0 = 0.70$  e  $g_0 = 0.90$  o estimador  $\hat{g}_0$  tem um enviesamento baixo, quando comparado com os outros cenários. Quando  $g_0 = 1.00$ ,

o estimador do parâmetro  $g_0$  é mais enviesado quando o tamanho da população é menor, i.e.  $N = 150$ .

Na Figura 3.5 apresenta-se a distribuição *a posteriori* do parâmetro  $g_0$ , quando o tamanho da população é igual a 500 e a probabilidade de captura é igual a 0.20, para todos os cenários de  $g_0$ . Quando o estimador do parâmetro  $g_0$  assume valores baixos ( $g_0 = 0.50$  e  $g_0 = 0.70$ ), as distribuições *a posteriori* são aproximadamente simétricas, e apresentam uma assimetria negativa para os cenários com valores elevados ( $g_0 = 0.90$  e  $g_0 = 1.00$ ).

### 3.6 Experiência com latas de cervejas

Este exemplo é baseado na experiência feita por Otto e Pollock (1990) com 495 latas castanhas de cervejas. Os autores utilizaram latas distribuídas em quatro grupos de tamanhos diferentes: 1, 2, 4 e 8, respectivamente. As latas foram distribuídas aleatoriamente ao longo de um transecto linear com 200 metros de comprimento e 40 metros de largura. Foram detectadas 202 latas, independentemente, por 9 observadores (ver dados na Tabela B.2 no Apêndice B.1). Tendo em conta que o modelo combinado apresentado neste capítulo não leva em conta o factor tamanho de grupo e diferentes observadores, consideraram-se as latas como unidades individuais, vistas por apenas um observador. Nos próximos capítulos serão abordados modelos que incorporam a heterogeneidade (observável e não observável) individual e consideram a hipótese de existirem vários observadores na linha central do transecto.

Considerando a função Semi-Normal como função de detecção, estimaram-se, utilizando o método de transectos lineares, os seguintes parâmetros:  $N_{TL}$ ,  $p_2^*$  e  $\mu^*$ .

Com o objectivo de aplicar os dados utilizando o modelo combinado, assumiu-se sem perda de generalidade que  $p_1=0.30$ , estimou-se  $p_2^* = 0.38$  utilizando o número de latas de cervejas detectadas,  $n_2 = 202$ , e por isso, parece razoável considerar  $p_2=0.40$ . Com base nos valores de  $p_1$  e  $p_2$  e dado que  $N = 495$ , recorrendo ao programa MAPLE V (Char *et al.*, 1991), gerou-se uma amostra aleatória de uma distribuição multinomial, do tipo referida na equação (3.2), obtendo-se  $n_1=149$  e  $n_{11}=60$ .

### 3.6.1 Análises de diagnóstico

Tendo em conta os dados das distâncias das latas de cervejas (Tabela B.3), e que  $n_{10} = 89$ ,  $n_{01} = 142$  e  $n_{11} = 60$ , utilizando o método de máxima verosimilhança e o método bayesiano, foram estimados os parâmetros  $g_0$ ,  $N$ ,  $p_1$ ,  $\sigma$ .

Todas as cadeias de Markov dos parâmetros passaram no teste de convergência de Gelman e Rubin (1992), a autocorrelação aproxima-se para valores perto de zero, quando o espaçamento é de 20 (Figura 3.6 a) e b)), e o teste de Raftery e Lewis (1995) indica um período de aquecimento relativamente curto. Considerou-se uma amostra de tamanho 9950, a mediana da distribuição *a posteriori* como estimativa pontual dos parâmetros e um intervalo de credibilidade de 95% para os parâmetros, com excepção do parâmetro  $g_0$ . Para  $g_0$  utilizou-se um intervalo de maior densidade *a posteriori* de 95%, porque a respectiva distribuição é assimétrica, como foi observado na Figura 3.5(d). Para as distribuições unimodais e simétricas, os intervalos de credibilidade e HPD são idênticos (Gelman *et al.*, 2004).

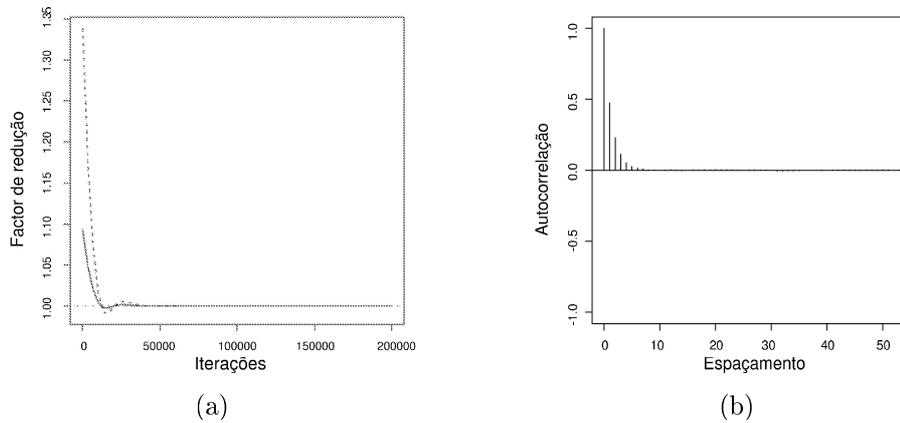


Figura 3.6: Gráfico do teste de convergência de Gelman e Rubin (a), utilizando duas cadeias de Markov com pontos iniciais diferentes, e gráfico da estimativa da função de autocorrelação (b), utilizando uma cadeia de Markov, para o parâmetro  $g_0$ . - - - 97.5%, ——— mediana.

### 3.6.2 Análise de sensibilidade

Para a análise de sensibilidade consideraram-se as seguintes distribuições *a priori*:

- (1)  $g_0 \sim \text{Beta}(1.00, 1.00)$ ,  $E[g_0] = 0.50$ ,  $\text{Var}[g_0] = 0.083$  ;
- (2)  $g_0 \sim \text{Beta}(0.50, 0.50)$ ,  $E[g_0] = 0.50$ ,  $\text{Var}[g_0] = 0.125$  ;
- (3)  $g_0 \sim \text{Beta}(2.00, 0.50)$ ,  $E[g_0] = 0.20$ ,  $\text{Var}[g_0] = 0.046$  ;
- (4)  $g_0 \sim \text{Beta}(0.50, 2.00)$ ,  $E[g_0] = 0.80$ ,  $\text{Var}[g_0] = 0.046$  ;
- (5)  $g_0 \sim \text{Beta}(2.00, 2.00)$ ,  $E[g_0] = 0.50$ ,  $\text{Var}[g_0] = 0.050$  .

Todas elas expressam pouca variabilidade para  $g_0$ , sendo a segunda a que tem maior variabilidade. A primeira, a segunda e a última distribuição foram escolhidas para expressar um valor esperado da probabilidade de detecção na distância 0 igual a 0.50. A terceira e a quarta distribuições representam um valor extremo para  $g_0$ , sendo o valor esperado, respectivamente, igual a 0.20 e a 0.80 (Figura 3.7(a)). Na

Figura 3.7(b) verifica-se que a inferência sobre  $g_0$  depende da escolha da distribuição *a priori*.

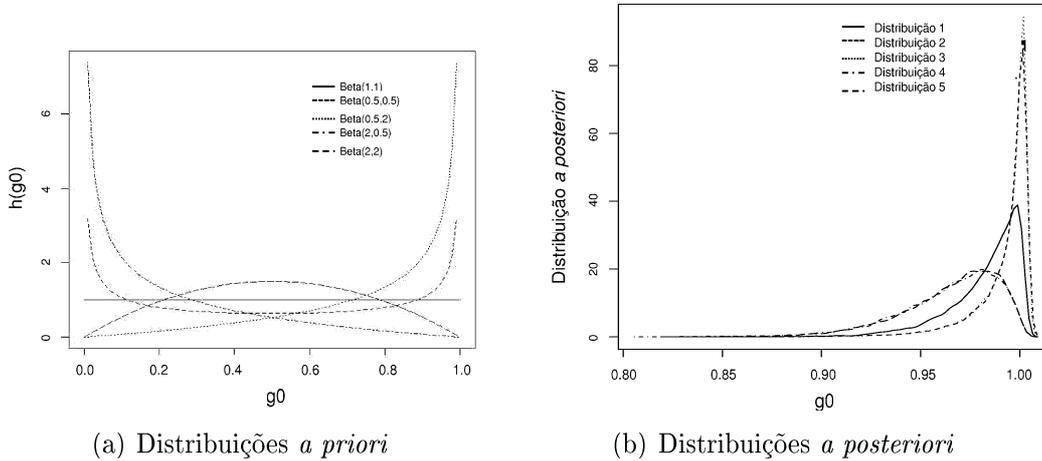


Figura 3.7: Gráfico das distribuições *a priori* (a), escolhidas para análise de sensibilidade para o parâmetro  $g_0$ , e as respectivas distribuições *a posteriori* (b).

### 3.6.3 Resultados

Escolheu-se a primeira distribuição,  $g_0 \sim Unif(0, 1)$ , para representar a informação *a priori*, atribuindo-se assim a todos os indivíduos a mesma função densidade de probabilidade (igual a 1) de serem vistos na distância 0. A partir da informação sobre o tamanho da população e sobre a variabilidade das distâncias obtidas, neste exemplo, considerou-se, para a função de detecção, a seguinte aproximação para  $\mu^* \approx 1.2534\sigma$ . Os hiperparâmetros  $\tau$  e  $\eta$  foram considerados iguais a 100 e 0.2 e  $\alpha$  e  $\beta$  são iguais a 100 e 1321.276, respectivamente.

Tabela 3.4: Quadro comparativo entre as estimativas obtidas através do método de Monte Carlo, utilizando o estimador de máxima verosimilhança, com as estimativas da distribuição *a posteriori*, utilizando o método de Monte Carlo baseado em cadeias de Markov, para os parâmetros  $g_0$ ,  $N$ ,  $p_1$ ,  $p_2$  e  $\mu^*$ . Foram utilizados três modelos: Lincoln-Petersen, transectos lineares e combinado. Apresentam-se as respectivas estimativas pontuais, o intervalo de confiança ou credibilidade/HPD e o desvio padrão (*sd*).

Estimador de Lincoln-Petersen			Estimador de transectos lineares		
	MV	EB		MV	EB
$\hat{p}_1$	0.297	0.300	$\hat{\mu}^*$	7.641	8.043
$sd(\hat{p}_1)$	0.032	0.028	$sd(\hat{\mu}^*)$	0.393	0.626
$\hat{p}_2$	0.403	0.405	$\hat{p}_2^*$	0.382	0.402
$sd(\hat{p}_2)$	0.040	0.034	$sd(\hat{p}_2^*)$	0.020	0.031
$\hat{N}_{LP}$	501.6	498.0	$\hat{N}_{TL}$	529.0	501.0
$sd(\hat{N}_{LP})$	41.96	32.47	$sd(\hat{N}_{TL})$	46.12	39.61
Estimador do modelo combinado					
	MV	EB		MV	EB
$\hat{g}_0$	1.050	0.923	$\hat{N}_{MCAP}$	501.7	513.0
$sd(\hat{g}_0)$	0.124	0.066	$sd(\hat{N}_{MCAP})$	41.61 <sup>a</sup>	31.57
$IC/HPD(\hat{g}_0)$	[0.79,1.28]	[0.78,1.00]	$IC(\hat{N})$	[420.14,583.26]	[457.00,581.00]

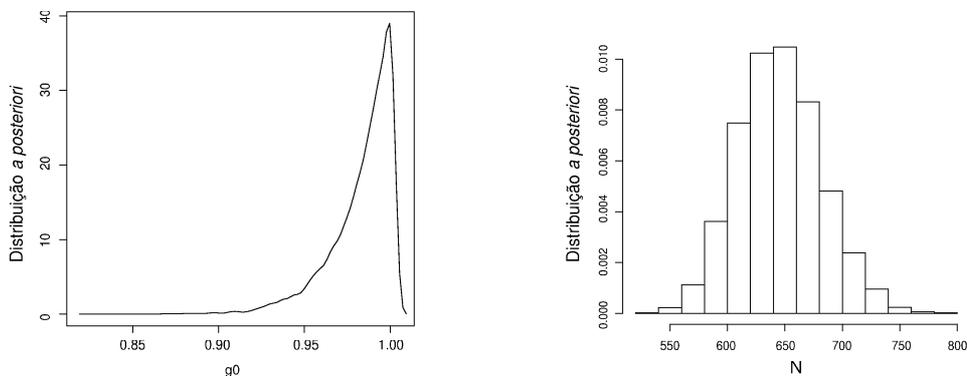
<sup>a</sup>Esta estimativa foi obtida recorrendo as propriedades da variância e utilizando a expansão da serie de Taylor de primeira ordem.

Na Tabela 3.4 apresentam-se estimativas para os parâmetros de interesse na estimação da abundância, utilizando a inferência bayesiana, tendo em conta os modelos de Lincoln-Petersen, transectos lineares e combinado, com as respectivas funções de verosimilhança dadas nas equações 3.3, 3.4 e 3.2, respectivamente.

As estimativas bayesianas, utilizando os modelos de transecto linear e captura-recaptura, foram obtidas utilizando o mesmo procedimento aplicado para o modelo combinado com a particularidade de usar a função de verosimilhança com apenas informação relevante para cada um dos respectivos modelos.

A população é sobrestimada em quase todos os métodos de estimação e modelos utilizados. O maior enviesamento verificou-se quando o modelo utilizado foi o de transectos lineares e o método foi o de máxima verosimilhança, com um enviesamento relativo de 6.87%. As estimativas bayesianas são quase sempre mais precisas. Por exemplo, o desvio padrão da distribuição *a posteriori* de  $g_0$  é igual a 0.066 (CV=7%), em comparação com 0.124 (CV=12%) obtido utilizando o método de máxima verosimilhança.

Na Figura 3.8(a), pode-se verificar que a distribuição *a posteriori* de  $g_0$  apresenta uma assimetria negativa, o que se traduz por um intervalo HPD de 95% para o parâmetro  $g_0$  de [0.78,1]. O método de máxima verosimilhança estima um valor superior a 1 para o parâmetro  $g_0$ , o que carece de significado sob o ponto de vista biológico e estatístico.



(a) Função distribuição *a posteriori* de  $g_0$ . (b) Função distribuição *a posteriori* de  $N$ .

Figura 3.8: Função distribuição *a posteriori* dos parâmetros  $g_0$ , (a), e  $N$ , (b).

A distribuição *a posteriori* do tamanho da população é dada na Figura 3.8(b), e apresenta uma mediana de 513.0 com um intervalo de credibilidade a 95% igual a [457.0,581.0]. Nota-se que a distribuição *a posteriori* do tamanho da população é praticamente simétrica.

## 3.7 Aplicação nas estimação da densidade de ungulados de montanha

### 3.7.1 Área de estudo

Neste exemplo, escolheu-se uma população de rebecos (*Rupicapra p. pyrenaica*), no Parque Nacional dos Pirenéus (sul de França), considerando que o tamanho da população na área de estudo é praticamente conhecido, valor este que servirá como ponto de referência para avaliar a metodologia.

No mês de Outubro do ano 2002, entre as 8:35 e 11:46 horas, a uma altitude entre 1527 e 1641 metros foi realizado um transecto linear com 2.0 km de comprimento. Apenas um dos lado do transecto foi amostrado, uma vez que a visibilidade foi dificultada num dos lados, devido a presença de muita vegetação. Durante a amostragem, mediu-se a distância perpendicular dos animais detectados na linha do transecto, e registraram-se o sexo e a idade de cada um deles para obter uma estimativa do tamanho da população, através do método de amostragem por distâncias (ver dados na Tabela B.3 no Apêndice B.2). As distâncias mínimas e máximas detectadas foram respectivamente iguais a 464 e 1580 metros. Os dados das distâncias foram truncadas à esquerda na distância igual a 450 metros.

A população de rebecos não é caçada desde a criação do parque em 1967, e tem uma estrutura de idades considerada como estável (Crampe, 1996). Os animais são marcados e monitorizados periodicamente, desde o início dos anos 90. Os animais são capturados, com laços para as extremidades e imobilização química, e marcados mediante colares e marcas auriculares.

Durante o período de estudo, a zona de *Clot-Cayan* albergava 230 rebecos adul-

tos, com uma densidade populacional de 62.5 rebecos/km<sup>2</sup>. Mais informação sobre a área de estudo e a população de rebecos pode ser encontrada em Crampe (1996) e Crampe *et al.* (2002).

### 3.7.2 Análises de diagnóstico

Todos os parâmetros passaram no teste de convergência e poucas iterações foram necessárias para que as cadeia de Markov atingissem a convergência. A autocorrelação entre as cadeias é baixa e um espaçamento de tamanho 20 é o suficiente para aproximar a autocorrelação para valores próximos de 0 (Figura 3.9(a)).

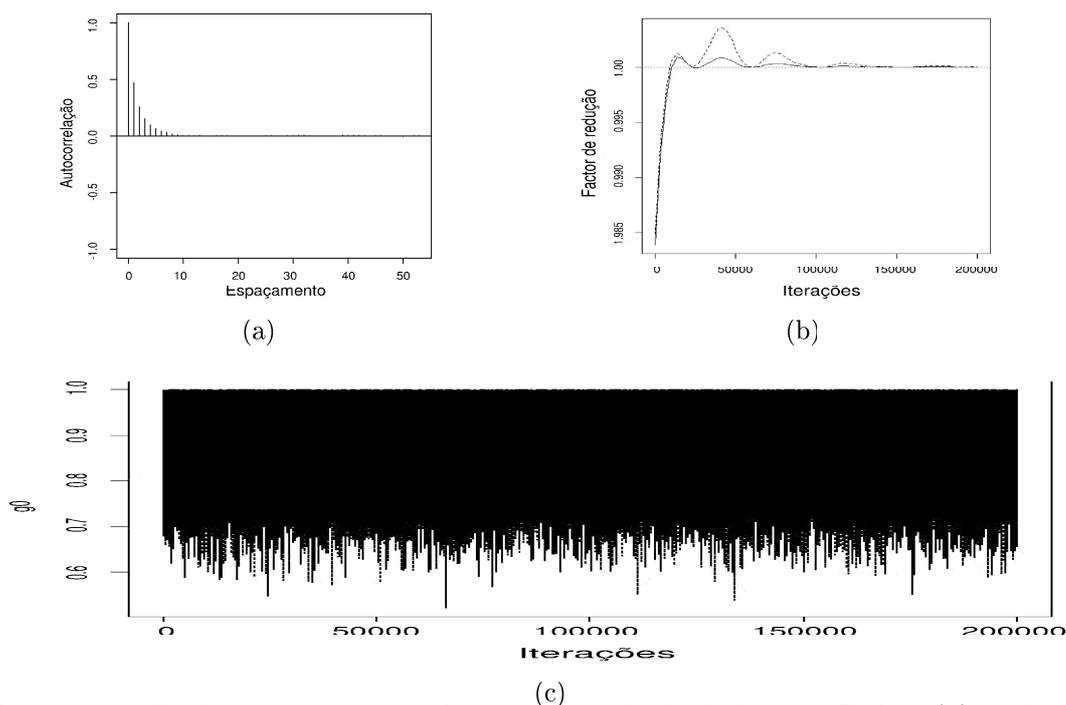


Figura 3.9: Gráficos para o teste de convergência de Gelman e Rubin (a), utilizando duas cadeias de Markov com pontos iniciais diferentes, a estimativa da função de autocorrelação (b) e o traço (c), utilizando uma cadeia de Markov, para o parâmetro  $g_0$ . - - - - 97.5%, ——— mediana.

Na Figura 3.9, pode-se verificar que a convergência para o parâmetro  $g_0$  ocorre para cadeias com pontos iniciais diferentes (Figura 3.9(c)). Apesar de o teste de Gelman e Rubin parecer indicar virtualmente uma convergência lenta (atenção à escala) (Figura 3.9(b)), outros testes indicaram convergência e um período de aquecimento de tamanho igual a 10 iterações.

### 3.7.3 Resultados

Para a abordagem clássica, os dados foram analisados com recurso ao programa DISTANCE, e considerou-se o Critério de Informação de Akaike, AIC (Akaike's Information Criteria) para a selecção de modelos. O coeficiente de variação e um intervalo de confiança de 95% para a densidade foram estimados, considerando diferentes funções de detecção.

Tabela 3.5: Estimativa obtida da análise da densidade, do coeficiente de variação (CV%) e do intervalo de confiança (IC) para a população de rebecos em *Clot-Cayan*, Parque Nacional de Pirenéus, utilizando dados de amostragem realizada no mês de Outubro do ano 2002. Foram consideradas quatro funções de detecção com ajustamento de co-seno: Taxa de risco, Exponencial negativa, Uniforme e Semi-Normal. A selecção de modelos é feita usando o Critério de Informação de Akaike (AIC).

Modelo	AIC	Densidade <sup>a</sup>	CV%	I.C. (95%)
Taxa de risco – co-seno	1363.93	34,39	7.93	[29.4,40.2]
Exponencial negativa – co-seno	1371.70	40.45	4,46	[37.0,44.2]
Uniforme – co-seno	1369.69	40.44	4,46	[37.0,44.2]
Semi-Normal – co-seno	1371.74	36.83	15.93	[26.9,50.4]

<sup>a</sup>Apenas um lado do transecto foi amostrado (lado direito).

O modelo Taxa de risco (Tabela 3.5), foi seleccionado como o que melhor se ajusta aos dados. A densidade estimada, tendo em conta apenas um lado do transecto, foi de 34.39 rebecos/km<sup>2</sup>, com um CV de 7.93% e um intervalo de confiança de 95% igual a [29.40,40.20]. Ao ter-se em conta ambos os lados, a densidade é de 68.77 rebecos/km<sup>2</sup>. Claro que nesta estimativa se supõe que a densidade no lado não amostrado do transecto é idêntica e este pressuposto pode ser questionado.

Foram detectados 100 rebecos, dos quais 53 estavam marcados durante o período de estudo. Usando o estimador de Lincoln-Petersen, obteve-se que o número de animais na área é 231, o que corresponde a uma densidade de 62.60 rebecos/km<sup>2</sup>, CV = 13.8 % e I.C. (95%) = [45.7,79.5].

Estudos anteriores da mesma população (Pérez *et al.*, em preparação), utilizando o MCAP, indicam que a probabilidade de detecção de um indivíduo na linha do transecto não é inferior a 1, i.e.  $g_0 \approx 1$ , mas os desvios padrões dos estimadores são muito elevados, apesar de combinar informação dos dois métodos de amostragem.

Com a informação disponível do número de rebecos marcados, estimou-se  $g_0$  utilizando o MCAP. O estimador MCAP para  $g_0$ , quando a técnica utilizada é a máxima verosimilhança, é um valor superior a 1 com um coeficiente de variação associado não inferior a 14% (Tabela 3.6).

Tabela 3.6: Estimativa obtida da análise do tamanho da população ( $N$ ) e da estimativa para  $g_0$  e o respectivo coeficiente de variação (CV%) e intervalo de confiança (IC), utilizando o estimador de máxima verosimilhança do modelo combinado proposto por Alpizar-Jara e Pollock (1999), para a população de rebecos em *Clot-Cayan*, Parque Nacional de Pirenéus, utilizando dados de amostragem realizada no mês de Outubro do ano 2002. Foram consideradas quatro funções de detecção com ajustamento de co-seno: Taxa de risco, Exponencial negativa, Uniforme e Semi-Normal.

Modelo	$N$	$g_0$	CV% [ $g_0$ ]	I.C. (95%) [ $g_0$ ]
Taxa de risco – co-seno	253	1.10	15.89	[0.76,1.44]
Exponencial Negativa – co-seno	298	1.29	14.47	[0.93,1.66]
Uniforme – co-seno	298	1.29	14.47	[0.93,1.66]
Semi-Normal – co-seno	271	1.18	21.05	[0.69,1.66]

Na Tabela 3.7 apresentam-se as estimativas para os parâmetros de interesse na estimação da abundância, utilizando a inferência bayesiana. O coeficiente de variação da distribuição *a posteriori* de  $g_0$  é igual a 7.25%, em comparação com 15.89% para CV da estimativa de  $g_0$ , obtida mediante o método de máxima verosimilhança.

A distribuição *a posteriori* apresenta um grande assimetria negativa, com um intervalo HPD de 95% para o parâmetro  $g_0$  igual a [0.78,1] (Tabela 3.7 e Figura 3.10). Considerando que 230 é o valor que se acredita ser mais próximo do tamanho real da população de rebecos, a estimativa obtida para  $N$ , utilizando o modelo combinado e a abordagem bayesina é menos enviesada, se se comparar com a estimativa obtida utilizando o modelo de transectos lineares e o método de máxima verosimilhança.

Tabela 3.7: Estimativa, intervalo de credibilidade/HPD e desvio padrão (*sd*) obtidos para os parâmetros  $g_0$ ,  $N$ ,  $p_1$  e  $p_2$ , utilizando o modelo combinado, para a população de rebecos em *Clot-Cayan*, Parque Nacional de Pirenéus, utilizando dados de amostragem do ano 2002.

Parâmetro	Mediana	CV%	I.C. (95%)
$N$	242.0	7.50	[210.0,281.0]
$g_0$	0.922	7.25	[0.78,1.00] <sup>a</sup>
$p_1$	0.22	14.13	[0.17,0.29]
$p_2$	0.40	7.87	[0.33,0.45]

<sup>a</sup>Intervalo HPD (95%).

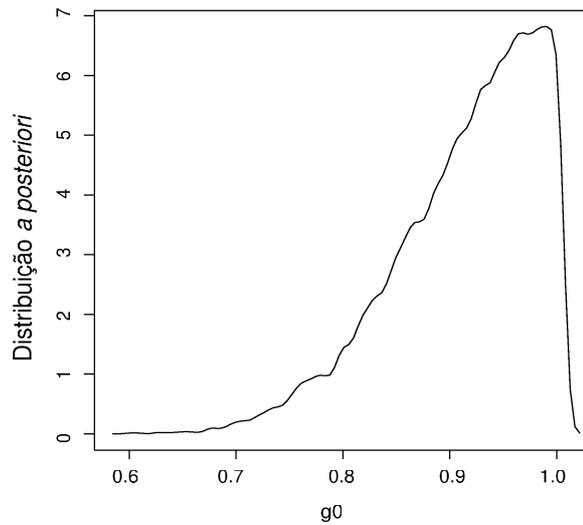


Figura 3.10: Distribuição *a posteriori* para  $g_0$ .

## Capítulo 4

# Estimação bayesiana da probabilidade na distância zero utilizando o modelo de dupla plataforma e heterogeneidade

Parte deste capítulo encontra-se sob a forma de publicação:

Monteiro, J. F. G., Alpizar-Jara, R., Serrano, E., Campre, J.-P. e Pérez, J. M. (2009). A Bayesian estimation approach for  $g_0$  estimation in line transect mark-resighting framework. (*em revisão*)

## 4.1 Introdução

A heterogeneidade nas probabilidades de detecção ou avistamento é outra fonte comum de enviesamento nos estimadores da abundância ou densidade. A heterogeneidade pode ser devida a várias variáveis observáveis, tais como distância, idade, sexo e tamanho de grupo, bem como a outras possíveis características individuais não observáveis (Pollock, 2002). Este tipo de enviesamento pode ser severo, particularmente se for negligenciado quando se estima  $g_0$ .

Borchers *et al.* (1998a,b) propuseram modelos de dupla plataforma, combinando transectos lineares e captura-recaptura, considerando heterogeneidade nas probabilidades de detecção, incorporando outras covariáveis para além das distâncias perpendiculares. O modelo proposto apenas considera covariáveis observáveis relacionadas com características individuais.

Pollock e Kendall (1987) foram os primeiros investigadores a indicar o problema da estimação de  $g_0$  e sugeriram métodos que combinam contagem no terreno com avistamento aéreo. Outros autores (e.g., Alpizar-Jara e Pollock, 1999; Borchers *et al.*, 1998a,b; Manly *et al.*, 1996; Okamura *et al.*, 2003; Schweder *et al.*, 1999) utilizaram e recomendaram a combinação de captura-recaptura e métodos de transectos lineares para estimar  $g_0$ , mas nenhum deles utilizou inferência bayesiana.

A correcção de estimadores da abundância ou da densidade estimando  $g_0$  foi aplicada com sucesso em amostragens para estimar abundância de cervos do pantanal (*Blastocerus dichotomus*) (Andriolo *et al.*, 2005). A amostragem de transectos lineares de dupla plataforma também foi realizada em Cardigan Bay SAC, nos Estados Unidos, para estimar  $g_0$  em golfinhos nariz de garrafa (*Tursiops truncatus*),

focas cinzentas (*Halichoerus grypus*) e toninhas do cais (*Phocoena phocoena*) (Reay, 2005). Laake *et al.* (1997) também estimaram  $g_0$  para toninhas do cais baseando-se em dados de amostragens aéreas.

O modelo proposto por Alpizar-Jara e Pollock (1999), tal como foi definido no Capítulo 3, é utilizado como base para modelar a heterogeneidade, nas probabilidades de detecção na distância zero, relativa às características inerentes aos indivíduos detectados. A modelação da heterogeneidade observável, tem sido tradicionalmente abordado utilizando modelos de captura-recaptura para populações fechadas, designados por  $\mathcal{M}_h$  e  $\mathcal{M}_{th}$  (Otis *et al.*, 1978).

À semelhança do Capítulo 2, aqui utilizou-se a regressão logística para modelar a heterogeneidade observável nas probabilidades de detecção usando covariáveis individuais. A heterogeneidade não observável é modelada como efeito aleatório através da abordagem bayesiana.

Neste capítulo, propõe-se um modelo que incorpora heterogeneidade em  $g_0$ . Apresenta-se a função de verosimilhança deste modelo e especificam-se as respectivas distribuições *a priori* para os parâmetros desconhecidos envolvidos na análise bayesiana utilizando MCMC. O desempenho do estimador da probabilidade de detecção na distância zero, baseada no modelo combinado, é analisada e comparada apenas quando a heterogeneidade observável é modelada (*MCH*), ou quando ambas as heterogeneidades (observável e não observável) são modeladas (*MCH\**). A título ilustrativo será utilizado um exemplo com ungulados de meio porte, baseado em dados, do ano 2007, da população de rebecos (*Rupicapra p. pyrenaica*), no Parque Nacional de Perinéus.

## 4.2 Modelando heterogeneidade observável e não observável em $g_0$

A seguir, apresenta-se o modelo combinado marca-recaptura e transecto linear, quando a probabilidade de detecção na distância zero,  $g_0$ , depende das características inerentes aos indivíduos observados e das covariáveis individuais.

A heterogeneidade é um facto na vida das populações selvagens. Muitos factores desconhecidos podem influenciar a probabilidade de detecção dos indivíduos na linha central. Se esta heterogeneidade imensurável for ignorada, os animais detectados na linha do transecto podem ser aqueles que são mais facilmente observados, o que provoca um enviesamento em  $\hat{g}_0$ . Consequentemente, o estimador para o tamanho da população proposto por Alpizar-Jara e Pollock (1999), que é inversamente proporcional a  $g_0$ , será também enviesado.

Quando os modelos têm muitos parâmetros, usualmente, são adoptados modelos hierárquicos. Assim, a incerteza nos hiperparâmetros é modelada assumindo diferentes níveis para as distribuições *a priori*. Os modelos hierárquicos têm sido utilizados para estimar parâmetros em métodos de captura-recaptura, como por exemplo Johnson e Hoeting (2003) e Link e Barker (2005).

Aqui, apresenta-se uma abordagem para modelar a probabilidade na distância zero, baseado no trabalho apresentado no Capítulo 2, onde se define um modelo  $\mathcal{M}_{th}^*$  que incorpora heterogeneidade observável e não observável em probabilidades de capturas individuais:

$$\ln \left( \frac{g_{0i}}{1 - g_{0i}} \right) = \beta_0 + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{grupo}} \times \text{grupo}_i + \delta_i,$$

e

$$g_0 = \sum_{i=1}^{n_2} \frac{g_{0i}}{n_2},$$

para

$$g_{0i} = \frac{e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{grupo}} \times \text{grupo}_i + \delta_i}}{1 + e^{\beta_0 + \beta_{\text{sexo}} \times \text{sexo}_i + \beta_{\text{idade}} \times \text{idade}_i + \beta_{\text{grupo}} \times \text{grupo}_i + \delta_i}}, \quad (4.1)$$

para  $i = 1, \dots, n_2$ . O coeficiente  $\beta_{\text{grupo}}$  e a variável  $\text{grupo}_i$  representam respectivamente, o efeito do tamanho do grupo e o tamanho do grupo a que o indivíduo  $i$  pertence.

Estimou-se  $g_0$  como sendo uma média, tendo em conta a informação dos indivíduos, como se cada indivíduo estivesse na linha do transecto. A heterogeneidade não observável inerente ao indivíduo  $i$  é modelada como sendo um efeito aleatório  $\delta_i$ , que poderá explicar variáveis não mensuráveis que afectam  $g_0$  ao nível individual.

### 4.2.1 Função de verosimilhança

Seja  $\mathbf{z} = (\text{sexo}_1, \dots, \text{sexo}_{n_2}, \text{idade}_1, \dots, \text{idade}_{n_2}, \text{grupo}_1, \dots, \text{grupo}_{n_2})$  o vector que contém a informação sobre as características observáveis dos indivíduos, tais como o sexo, idade e tamanho do grupo. Sejam  $\boldsymbol{\beta} = (\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}})$  e  $\boldsymbol{\delta} = (\delta_1, \dots, \delta_{n_2})$  os vectores com os parâmetros que correspondem aos efeitos fixos e aleatórios, res-

pectivamente, que definem  $g_0$ . A função de verosimilhança para o modelo combinado é dada por:

$$\begin{aligned}
 f(\mathbf{n}, \mathbf{y}, \mathbf{z} | N, \lambda, p_1, \sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= \frac{N!}{n_{10}!n_{01}!n_{11}!(N-n)!} [p_1 (1 - p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))]^{n_{10}} \\
 &\times [(1 - p_1)p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)]^{n_{01}} [p_1p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)]^{n_{11}} \\
 &\times [(1 - p_1)(1 - p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))]^{N-n} \prod_{i=1}^{n_2} \frac{g^*(y_i | \sigma)}{\mu^*}, \\
 &0 < p_1, p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) < 1, \quad N \geq n. \tag{4.2}
 \end{aligned}$$

A função de detecção tal como foi definida em (3.1) depende agora de outros parâmetros para além de  $\sigma$ :

$$g(y | \sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) = g_0(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times g^*(y | \sigma), \quad 0 \leq y \leq w, \quad \sigma > 0, \tag{4.3}$$

onde  $\boldsymbol{\beta}$ ,  $\boldsymbol{\delta}$  e  $\tau_\delta$  são os vectores parâmetros que definem  $g_{0_i}$  na equação (4.1). Tendo em conta a função de detecção, definida como sendo uma Semi-Normal em (3.6), a função de detecção em (4.3) pode ser escrita como  $g(y | \sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) = g_0(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times e^{-\frac{1}{2}(\frac{y}{\sigma})^2}$ , e um estimador para a probabilidade de avistar um indivíduo, utilizando o modelo  $MCH^*$ , é dado por  $\widehat{p}_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) = \widehat{g}_0(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \widehat{p}_2^*(\sigma)$ .

A função de verosimilhança acima referida tem um total de  $7 + n_2$  parâmetros,  $(N, p_1, \sigma, \beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}}, \delta_1, \dots, \delta_{n_2})$ , para serem estimados. Uma vez que a estatística suficiente mínima é obtida a partir de  $(\mathbf{n}, \mathbf{y}, \mathbf{z})$ , cujo vector tem dimensão  $3 + 4n_2$ , então todos os  $7 + n_2$  parâmetros são estimáveis.

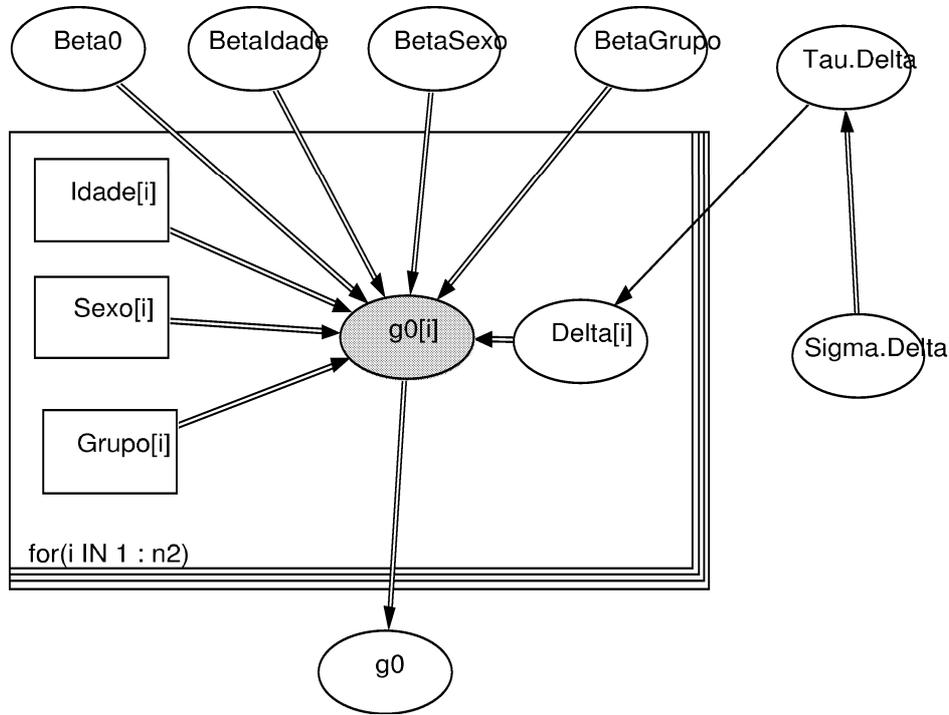


Figura 4.1: DAG do modelo combinado de captura-recaptura e transectos lineares, quando as probabilidades de captura dependem de características individuais e de outros factores não observáveis ( $MCH^*$ ), para estimar a probabilidade de detectar um indivíduo na distância zero,  $g_0$ . O rectângulo maior representa o ciclo em torno dos  $g_0$  individuais ( $g_{0_i}, i = 1, \dots, n_2$ ).

### 4.2.2 Distribuição *a priori*

Nesta secção, apresenta-se um modelo bayesiano com estrutura hierárquica para a distribuição *a priori* dos parâmetros. Assumindo que os parâmetros desconhecidos são independentes, define-se uma distribuição *a priori* para cada parâmetro. De forma análoga ao que se apresentou no Capítulo 2, definiram-se distribuições *a priori* Gaussiana, com média igual a 0 e precisão igual a 1.0, para os parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$  e distribuições *a priori* hierárquicas do tipo Normal-Gama para

os efeitos aleatórios individuais. Os valores para  $\xi$  e  $v$  são ambos iguais a 10.0, o que corresponde a um valor esperado para  $\tau_\delta$  igual a 1.0, mas com precisão de 0.1. Para os parâmetros  $N$ ,  $p_1$ ,  $\sigma$  foram consideradas as mesmas distribuições *a priori* definidas no Capítulo 3.

### 4.2.3 Distribuições *a posteriori*

Seja  $\boldsymbol{\theta} = (N, \lambda, p_1, \sigma^2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$  o vector dos parâmetros desconhecidos, e  $\mathbf{n}$ ,  $\mathbf{y}$  e  $\mathbf{z}$  os respectivos vectores com a informação observada (ver DAG do modelo combinado na Figura 4.1), sobre o pressuposto de independência entre os parâmetros, a distribuição *a posteriori* completa de  $\boldsymbol{\theta}$  é proporcional ao produto entre a função de verosimilhança (4.2) e as respectivas distribuições *a priori* dos parâmetros:

$$\begin{aligned}
 h(\boldsymbol{\theta}|\mathbf{n}, \mathbf{y}, \mathbf{z}) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta})h(N|\lambda)h(\lambda)h(p_1)h(\sigma^2) \\
 &\times h(\beta_0)h(\beta_{\text{sexo}})h(\beta_{\text{idade}})h(\beta_{\text{grupo}}) \prod_{i=1}^{n_2} h(\delta_i)h(\tau_\delta) \\
 &\propto \frac{N!}{n_{10}!n_{01}!n_{11}!(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\
 &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} \frac{e^{-\lambda} \lambda^N}{N!} \frac{\eta^\tau \lambda^{\tau-1} e^{-\eta\lambda}}{\Gamma(\tau)} \frac{\beta^\alpha (\frac{1}{\sigma^2})^{\alpha-1} e^{-\frac{\beta}{\sigma^2}}}{\Gamma(\alpha)} \\
 &\times \frac{1}{4\pi^2} e^{-\frac{\beta_0^2 + \beta_{\text{sexo}}^2 + \beta_{\text{idade}}^2 + \beta_{\text{grupo}}^2}{2}} \left( \frac{\tau_\delta}{2\pi} \right)^{\frac{n_2}{2}} e^{-\frac{\tau_\delta \sum_{i=1}^{n_2} \delta_i^2}{2}} \frac{10^{10} \tau_\delta^9 e^{-10\tau_\delta}}{\Gamma(10)}.
 \end{aligned}$$

**Proposição 4.1.** *A distribuição a posteriori condicional completa de  $(N, \lambda, p_1, \sigma^2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$  dada a informação sobre o histórico das capturas, as distâncias perpendiculares dos indivíduos detectados e as informações sobre as características individuais,  $(\mathbf{n}, \mathbf{y}, \mathbf{z})$ , é dada por:*

a)  $h(g_0 | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2)$

$$\propto \left(1 - \frac{\sigma g_0 K}{w}\right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_0^2 + \beta_{sexo}^2 + \beta_{idade}^2 + \beta_{grupo}^2}{2} + \tau_\delta \left(\sum_{i=1}^{n_2} \delta_i^2 + 10\right)} \left(\frac{\tau_\delta}{2\pi}\right)^{\frac{n_2}{2}} \tau_\delta^9;$$

b)

$$N \sim Poi(\lambda(1 - p_1)(1 - p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)));$$

c)

$$\lambda \sim Gama(N + \tau, 1 + \eta);$$

d)  $h(\sigma^2 | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$

$$\propto \left(\frac{w}{K g_0} - \sigma\right)^{N-n_2} \frac{e^{-\frac{\beta + \sum_{i=1}^{n_2} \frac{y_i^2}{\sigma^2}}{2}}}{\sigma^{2(\alpha-1)}};$$

e)

$$p_1 \sim Beta(n_1 + 1, N - n_1 + 1);$$

f) Para  $i = 1, \dots, n_2$ ,

$$h(\delta_i | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \boldsymbol{\beta}, \delta_1, \dots, \delta_{i-1}, \delta_{i+1}, \dots, \delta_{n_2}, \tau_\delta) \\ \propto \left(1 - \frac{\sigma g_0 K}{w}\right)^{N-n_2} g_0^{n_2} e^{-\frac{\tau_\delta \delta_i^2}{2}};$$

g)  $h(\tau_\delta | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \boldsymbol{\beta}, \boldsymbol{\delta})$

$$\propto \left(1 - \frac{\sigma g_0 K}{w}\right)^{N-n_2} g_0^{n_2} \left(\frac{\tau_\delta}{2\pi}\right)^{\frac{n_2}{2}} e^{-\sum_{i=1}^{n_2} \frac{\tau_\delta \delta_i^2}{2}} \tau_\delta^9 e^{-10\tau_\delta};$$

h)  $h(\beta_0 | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_{sexo}, \beta_{idade}, \beta_{grupo}, \boldsymbol{\delta}, \tau_\delta)$

$$\propto \left(1 - \frac{\sigma g_0 K}{w}\right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_0^2}{2}};$$

i)  $h(\beta_{sexo} | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_0, \beta_{idade}, \beta_{grupo}, \boldsymbol{\delta}, \tau_\delta)$

$$\propto \left(1 - \frac{\sigma g_0 K}{w}\right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_{sexo}^2}{2}};$$

j)  $h(\beta_{idade} | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_0, \beta_{sexo}, \beta_{grupo}, \boldsymbol{\delta}, \tau_\delta)$

$$\propto \left(1 - \frac{\sigma g_0 K}{w}\right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_{idade}^2}{2}};$$

k)  $h(\beta_{grupo} | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_0, \beta_{sexo}, \beta_{idade}, \boldsymbol{\delta}, \tau_\delta)$

$$\propto \left(1 - \frac{\sigma g_0 K}{w}\right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_{grupo}^2}{2}}.$$

*Demonstração.*

a) Tendo em conta a função de verosimilhança dada em (4.2) e as distribuições *a priori* de  $\boldsymbol{\beta}$ ,  $\boldsymbol{\delta}$  e  $\tau_\delta$ , tem-se que:

$$\begin{aligned}
 h(g_0|\mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta})h(\beta_0)h(\beta_{\text{sexo}})h(\beta_{\text{idade}})h(\beta_{\text{grupo}})\prod_{i=1}^{n_2} h(\delta_i)h(\tau_\delta) \\
 &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\
 &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} \frac{1}{4\pi^2} e^{-\frac{\beta_0^2 + \beta_{\text{sexo}}^2 + \beta_{\text{idade}}^2 + \beta_{\text{grupo}}^2}{2}} \\
 &\times \left( \frac{\tau_\delta}{2\pi} \right)^{\frac{n_2}{2}} e^{-\frac{\tau_\delta \sum_{i=1}^{n_2} \delta_i^2}{2}} \frac{10^{10} \tau_\delta^9 e^{-10\tau_\delta}}{\Gamma(10)} \\
 &\propto \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_0^2 + \beta_{\text{sexo}}^2 + \beta_{\text{idade}}^2 + \beta_{\text{grupo}}^2 + \tau_\delta (\sum_{i=1}^{n_2} \delta_i^2 + 10)}{2}} \left( \frac{\tau_\delta}{2\pi} \right)^{\frac{n_2}{2}} \tau_\delta^9;
 \end{aligned}$$

b) Considerando a função de verosimilhança dada em (4.2) e as distribuições *a priori* de  $N$  e  $\lambda$ , tem-se que:

$$\begin{aligned}
 h(N|\mathbf{n}, \mathbf{y}, \mathbf{z}, \lambda, p_1, \sigma^2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta})h(N|\lambda)h(\lambda) \\
 &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\
 &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} \frac{e^{-\lambda} \lambda^N}{N!} \frac{\eta^\tau \lambda^{\tau-1} e^{-\eta\lambda}}{\Gamma(\tau)} \\
 &\propto \frac{\lambda^N}{(N-n)!} (1-p_1)^{N-n_1} \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} \\
 &\propto \frac{\lambda^N}{(N-n)!} (1-p_1)^N (1-p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))^N \\
 &\propto \frac{\lambda^N}{(N-n)!} ((1-p_1)(1-p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)))^N
 \end{aligned}$$

$$\propto e^{\lambda(1-p_1)(1-p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))} \frac{(\lambda(1-p_1)(1-p_2(\sigma, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)))^{N-n}}{(N-n)!}, \quad N \geq n;$$

c) Dada a função de verosimilhança definida em (4.2) e as distribuições *a priori* de  $N$  e  $\lambda$ , tem-se que:

$$\begin{aligned} h(\lambda | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, p_1, \sigma^2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) h(N | \lambda) h(\lambda) \\ &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\ &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}} e^{-\lambda} \lambda^N \eta^\tau \lambda^{\tau-1} e^{-\eta\lambda}}{(\sigma K)^{n_2} N! \Gamma(\tau)} \\ &\propto e^{-\lambda(1+\eta)} \lambda^{N+\tau-1} \propto \frac{(1+\eta)^{N+\tau} \lambda^{N+\tau-1} e^{-\lambda(1+\eta)}}{\Gamma(N+\tau)}; \end{aligned}$$

d) Pela função de verosimilhança definida em (4.2) e a distribuição *a priori* de  $\sigma^2$ , tem-se que:

$$\begin{aligned} h(\sigma^2 | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) h(\sigma^2) \\ &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\ &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}} \beta^\alpha \left( \frac{1}{\sigma^2} \right)^{\alpha-1} e^{-\frac{\beta}{\sigma^2}}}{(\sigma K)^{n_2} \Gamma(\alpha)} \\ &\propto \left( \frac{w}{K g_0} - \sigma \right)^{N-n_2} \frac{e^{-\frac{\beta + \sum_{i=1}^{n_2} y_i^2}{\sigma^2}}}{\sigma^{2(\alpha-1)}}; \end{aligned}$$

e) Tendo em conta a função de verosimilhança dada em (4.2) e a distribuição *a priori* de  $p_1$ , tem-se que:

$$\begin{aligned} h(p_1|\mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, \sigma^2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta})h(p_1) \\ &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\ &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} \\ &\propto p_1^{n_1} (1-p_1)^{N-n_1} \frac{\Gamma(N+2)}{\Gamma(n_1+1)\Gamma(N-n_1+1)}; \end{aligned}$$

f) Considerando a função de verosimilhança dada em (4.2) e as distribuições *a priori* de  $\delta_i$  e  $\tau_\delta$ , tem-se que para  $i = 1, \dots, n_2$ :

$$\begin{aligned} h(\delta_i|\mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \boldsymbol{\beta}, \delta_1, \dots, \delta_{i-1}, \delta_{i+1}, \dots, \delta_{n_2}, \tau_\delta) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta})h(\delta_i|\tau_\delta)h(\tau_\delta) \\ &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\ &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} e^{-\frac{\tau_\delta \delta_i^2}{2}} \frac{10^{10} \tau_\delta^9 e^{-10\tau_\delta}}{\Gamma(10)} \\ &\propto \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} g_0^{n_2} e^{-\frac{\tau_\delta \delta_i^2}{2}}; \end{aligned}$$

g) Tendo em conta a função de verosimilhança dada em (4.2) e as distribuições *a priori* de  $\boldsymbol{\delta}$  e  $\tau_\delta$ , tem-se que:

$$\begin{aligned} h(\tau_\delta|\mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \boldsymbol{\beta}, \boldsymbol{\delta}) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta})h(\delta_1|\tau_\delta) \dots h(\delta_{n_2}|\tau_\delta)h(\tau_\delta) \\ &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \end{aligned}$$

$$\begin{aligned} & \times \left[ (1 - p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} \left( \frac{\tau_\delta}{2\pi} \right)^{\frac{n_2}{2}} e^{-\sum_{i=1}^{n_2} \frac{\tau_\delta \delta_i^2}{2}} \frac{10^{10} \tau_\delta^9 e^{-10\tau_\delta}}{\Gamma(10)} \\ & \propto \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} g_0^{n_2} \left( \frac{\tau_\delta}{2\pi} \right)^{\frac{n_2}{2}} e^{-\sum_{i=1}^{n_2} \frac{\tau_\delta \delta_i^2}{2}} \tau_\delta^9 e^{-10\tau_\delta}; \end{aligned}$$

h) A partir da função de verosimilhança dada em (4.2) e a distribuição *a priori* de  $\beta_0$ , tem-se que:

$$\begin{aligned} & h(\beta_0 | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}}, \boldsymbol{\delta}, \tau_\delta) \propto f(\mathbf{n}, \mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) h(\beta_0) \\ & \propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1 - p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\ & \times \left[ (1 - p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} e^{-\frac{\beta_0^2}{2}} \\ & \propto \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_0^2}{2}}; \end{aligned}$$

i) Tendo em conta a função de verosimilhança dada em (4.2) e a distribuição *a priori* de  $\beta_{\text{sexo}}$ , tem-se que:

$$\begin{aligned} & h(\beta_{\text{sexo}} | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_0, \beta_{\text{idade}}, \beta_{\text{grupo}}, \boldsymbol{\delta}, \tau_\delta) \propto f(\mathbf{n}, \mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) h(\beta_{\text{sexo}}) \\ & \propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1 - p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\ & \times \left[ (1 - p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} e^{-\frac{\beta_{\text{sexo}}^2}{2}} \\ & \propto \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_{\text{sexo}}^2}{2}}; \end{aligned}$$

j) Da função de verosimilhança dada em (4.2) e a distribuição *a priori* de  $\beta_{\text{idade}}$ , tem-se que:

$$\begin{aligned}
 h(\beta_{\text{idade}} | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_0, \beta_{\text{sexo}}, \beta_{\text{grupo}}, \boldsymbol{\delta}, \tau_{\delta}) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) h(\beta_{\text{idade}}) \\
 &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\
 &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} e^{-\frac{\beta_{\text{idade}}^2}{2}} \\
 &\propto \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_{\text{idade}}^2}{2}};
 \end{aligned}$$

k) Considerando a função de verosimilhança dada em (4.2) e a distribuição *a priori* de  $\beta_{\text{grupo}}$ , tem-se que:

$$\begin{aligned}
 h(\beta_{\text{grupo}} | \mathbf{n}, \mathbf{y}, \mathbf{z}, N, \lambda, p_1, \sigma^2, \beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \boldsymbol{\delta}, \tau_{\delta}) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z} | \boldsymbol{\theta}) h(\beta_{\text{grupo}}) \\
 &\propto \frac{N!}{(N-n)!} \left[ p_1 \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{n_{10}} \left[ (1-p_1) \frac{\sigma g_0 K}{w} \right]^{n_{01}} \left[ p_1 \frac{\sigma g_0 K}{w} \right]^{n_{11}} \\
 &\times \left[ (1-p_1) \left( 1 - \frac{\sigma g_0 K}{w} \right) \right]^{N-n} \frac{e^{-\frac{\sum_{i=1}^{n_2} y_i^2}{2\sigma^2}}}{(\sigma K)^{n_2}} e^{-\frac{\beta_{\text{grupo}}^2}{2}} \\
 &\propto \left( 1 - \frac{\sigma g_0 K}{w} \right)^{N-n_2} g_0^{n_2} e^{-\frac{\beta_{\text{grupo}}^2}{2}}.
 \end{aligned}$$

□

### 4.3 Estimação da densidade de ungulados de montanha

Escolheu-se, a título de exemplo, a mesma população de rebecos (*Rupicapra p. pyrenaica*) de *Cauterets*, no Parque Nacional de Pirenéus, cuja área de estudo foi descrita no Capítulo 3.

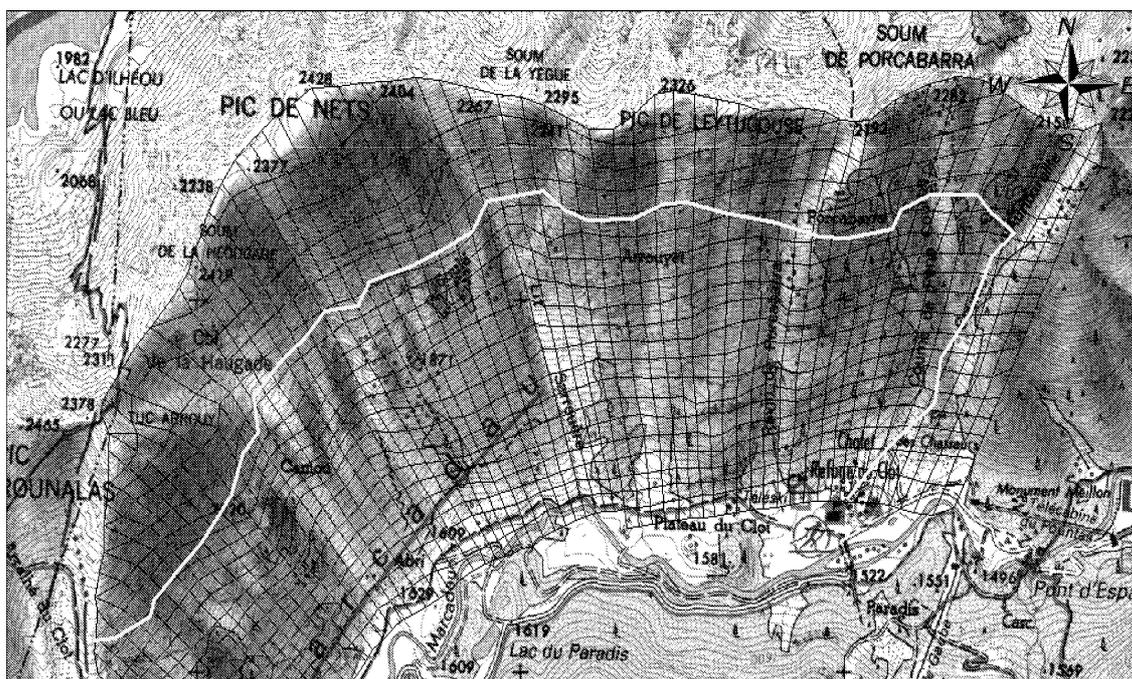


Figura 4.2: Transecto linear, ver linha traçada a branco, com 4.471 km de comprimento, realizado em “Clot-Cayan”, em Outubro de 2007, entre as 08:55 horas e 12:50 horas.

Em Outubro de 2007, entre as 08:55 horas e as 12:50 horas, foi realizado um transecto linear com 4.471 km de comprimento. Foram registadas as respectivas distâncias perpendiculares dos animais avistados, utilizando um telémetro a laser (Bushnell<sup>TM</sup>, Denver, Colorado), e os ângulos foram medidos com uma bússola. Os indivíduos foram observados a distâncias compreendidas entre os 50 e os 265 metros

(Figure 4.2).

Durante a amostragem, para além das distâncias perpendiculares, registaram-se o sexo, a idade, o tamanho do grupo em que o animal se encontrava, e também se este tinha sido previamente marcado ou não. Foram detectados 72 rebecos, 21 dos quais estavam previamente marcados e 34 eram machos. Uma vez que existiam apenas alguns rebecos crias, foram consideradas apenas duas categorias para a idade: 1-crias (0 a 1 ano)/juvenis (1 a 2 anos) e 2-adultos (mais de 2 anos). O tamanho do grupo variou entre 1 a 7 animais. Durante a realização da amostragem, sabe-se que no total haviam 57 animais marcados na área de estudo. No Apêndice B.4. (Tabela B.4) são apresentados mais detalhes sobre os dados.

### 4.3.1 Análises de diagnóstico

As duas cadeias relativamente aos parâmetros  $N$ ,  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\tau_\delta$ ,  $\delta_1$ ,  $\delta_{36}$  e  $\delta_{72}$ , para o modelo  $MCH^*$ , convergem ambas para a mesma região, apesar de terem pontos iniciais diferentes (Figura 4.3). Para além disso, é possível ver no gráfico do factor de redução, na Figura 4.4, que  $R$  converge para 1, para todos os respectivos parâmetros, o que indica a convergência das respectivas cadeias.

O diagnóstico de autocorrelação foi realizado para todas as cadeias dos parâmetros desconhecidos, indicando que um espaçamento de tamanho 50 entre iterações diminuía a autocorrelação para valores próximos de zero. O teste de convergência de Raftery e Lewis (1995) indicou que existe convergência para todos os parâmetros.

Apesar de o teste de Raftery e Lewis sugerir um período de aquecimento relativamente curto (Tabela 4.1), o teste de convergência Gelman e Rubin (Figura 4.4), indica que a convergência acontece após várias iterações.

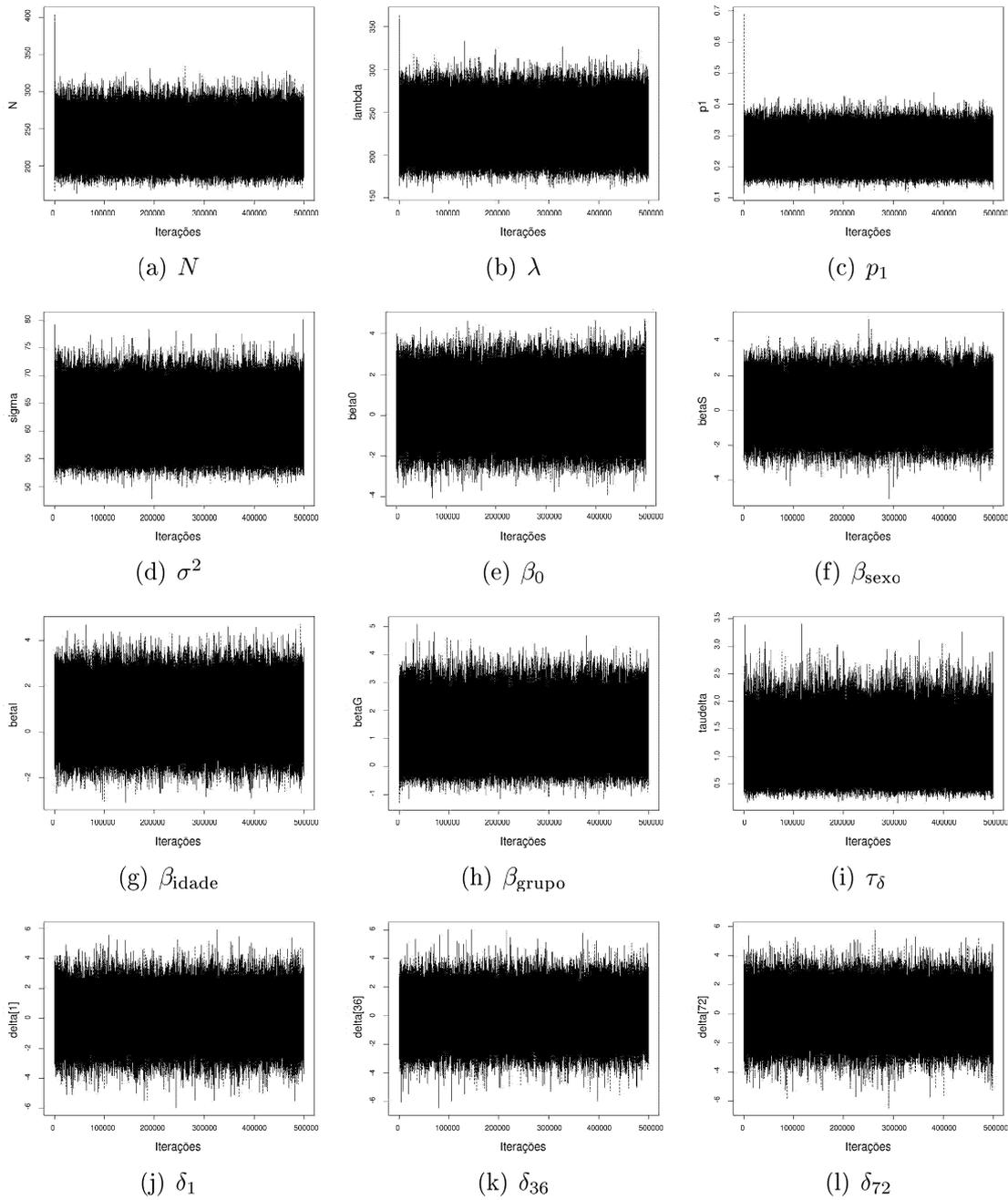


Figura 4.3: Traço das duas cadeias de Markov, com pontos iniciais diferentes, consideradas para os parâmetros  $N$ ,  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$ ,  $\tau_{\delta}$ ,  $\delta_1$ ,  $\delta_{36}$  e  $\delta_{72}$ , para o modelo  $MCH^*$ .

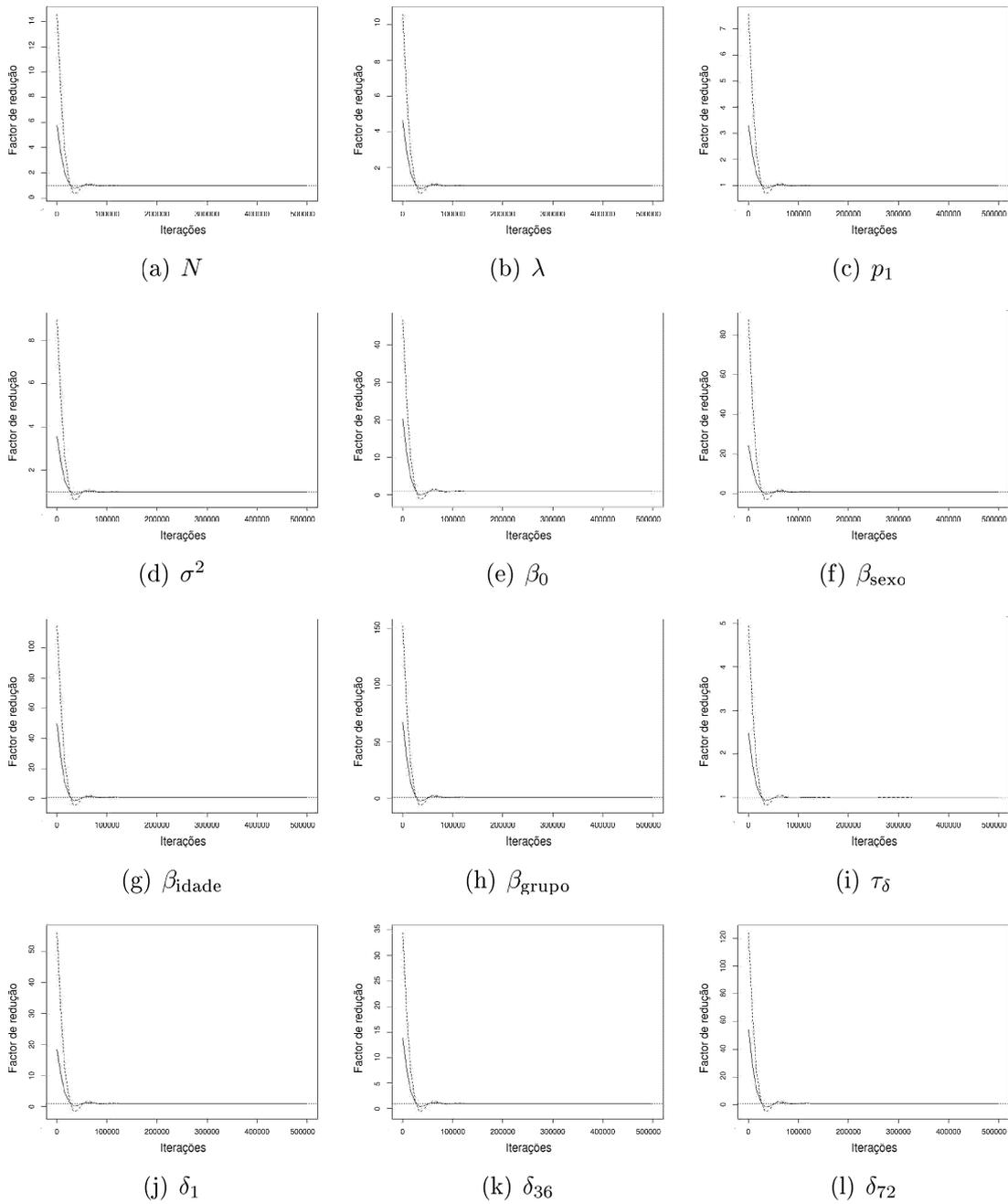


Figura 4.4: Teste de convergência Gelman e Rubin, utilizando duas cadeias de Markov com pontos iniciais diferentes, para os parâmetros  $N$ ,  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$ ,  $\tau_\delta$ ,  $\delta_1$ ,  $\delta_{36}$  e  $\delta_{72}$ , para o modelo  $MCH^*$ . - - - - 97.5%, ——— mediana.

Tabela 4.1: Estimativa do período de aquecimento e da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50), para as cadeias de Markov dos parâmetros  $N$ ,  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{peso}}$ ,  $\tau_\delta$ ,  $\delta_1$  e  $\delta_{72}$ , para o modelo  $MCH^*$ .

Parâmetros	Período de aquecimento	Função de autocorrelação			
		Espaçamento			
		5	10	20	50
$N$	12	0.0891	0.0075	0.0014	-0.0002
$\lambda$	10	0.0717	0.0063	0.0024	-0.0002
$p_1$	8	0.0423	0.0024	-0.0009	-0.0005
$\sigma^2$	8	0.0083	-0.0003	0.0027	0.0030
$\beta_0$	18	0.2667	0.0604	0.0113	-0.0021
$\beta_{\text{sexo}}$	18	0.1506	0.0248	0.0022	0.0019
$\beta_{\text{idade}}$	20	0.2953	0.1060	0.0179	-0.0002
$\beta_{\text{grupo}}$	26	0.3229	0.1192	0.0170	0.0006
$\tau_\delta$	64	0.5897	0.4341	0.2371	0.0337
$\delta_1$	24	0.1432	0.0263	0.0042	0.0020
$\delta_{72}$	20	0.1172	0.0200	-0.0007	-0.0004

Foram realizadas 500,000 iterações no WINBUGS, considerando que um período de aquecimento de tamanho igual a 5,000 iterações e um espaçamento de tamanho 50 entre iterações era adequado para assegurar convergência, remover a influência dos valores iniciais e diminuir a autocorrelação para valores próximos de zero. Assim, para uma amostra final de tamanho 9,900, foram seleccionados a mediana da distribuição *a posteriori* como estimativa pontual dos parâmetros e um intervalo de HPD de 95% para os parâmetros.

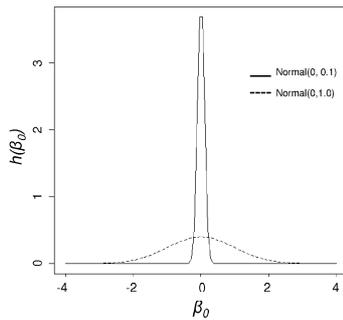
### 4.3.2 Análise de sensibilidade

Para a análise de sensibilidade, foram considerados quatro cenários, admitindo diferentes valores que definem as distribuições *a priori* dos parâmetros envolvidos no

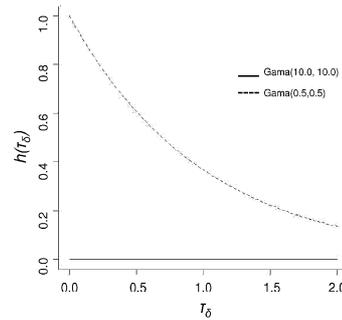
modelo logístico proposto para a probabilidade de detecção na distância zero:

- (1)  $\beta_0 \sim N(0, 1.0)$ ,  $\beta_{\text{sexo}} \sim N(0, 0.1)$ ,  $\beta_{\text{idade}} \sim N(0, 0.1)$ ,  $\beta_{\text{grupo}} \sim N(0, 0.1)$  e  $\delta_i | \tau_\delta \stackrel{i.i.d.}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim \text{Gama}(10.0, 10.0)$ , para  $i = 1, \dots, n_2$ ;
- (2)  $\beta_0 \sim N(0, 0.1)$ ,  $\beta_{\text{sexo}} \sim N(0, 0.1)$ ,  $\beta_{\text{idade}} \sim N(0, 0.1)$ ,  $\beta_{\text{grupo}} \sim N(0, 0.1)$  e  $\delta_i | \tau_\delta \stackrel{i.i.d.}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim \text{Gama}(10.0, 10.0)$ , para  $i = 1, \dots, n_2$ ;
- (3)  $\beta_0 \sim N(0, 1.0)$ ,  $\beta_{\text{sexo}} \sim N(0, 1.0)$ ,  $\beta_{\text{idade}} \sim N(0, 1.0)$ ,  $\beta_{\text{grupo}} \sim N(0, 1.0)$  e  $\delta_i | \tau_\delta \stackrel{i.i.d.}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim \text{Gama}(10.0, 10.0)$ , para  $i = 1, \dots, n_2$ ;
- (4)  $\beta_0 \sim N(0, 1.0)$ ,  $\beta_{\text{sexo}} \sim N(0, 0.1)$ ,  $\beta_{\text{idade}} \sim N(0, 0.1)$ ,  $\beta_{\text{grupo}} \sim N(0, 0.1)$  e  $\delta_i | \tau_\delta \stackrel{i.i.d.}{\sim} N(0, \tau_\delta)$  com  $\tau_\delta \sim \text{Gama}(0.5, 0.5)$ , para  $i = 1, \dots, n_2$ .

No primeiro cenário, admite-se maior precisão para a distribuição *a priori* do parâmetro geral na regressão logística ( $\beta_0$ ), embora que ainda seja pequena, e menor precisão nos restantes cenários.



(a)  $N(0, 0.1)$  e  $N(0, 1.0)$



(b)  $\text{Gama}(0.5, 0.5)$  e  $\text{Gama}(10.0, 10.0)$

Figura 4.5: Gráfico das distribuições *a priori* escolhidas para a análise de sensibilidade para os parâmetros: ( $\beta_0, \beta_{\text{sexo}}, \beta_{\text{idade}}, \beta_{\text{grupo}}, \delta$ ) e  $\tau_\delta$ , figuras a) e b) respectivamente.

Para os parâmetros  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$  e  $\beta_{\text{grupo}}$ , escolheu-se o terceiro cenário, como sendo aquele no qual todos apresentam uma distribuição *a priori* com muita precisão, e menor precisão nos outros cenários (Figura 4.5(a)). Os valores que definem a distribuição *a priori* hierárquica para o efeito aleatório,  $\delta$ , representam todos uma precisão igual a 1.0 para a distribuição Gaussiana, no entanto o parâmetro  $\tau_{\delta}$  tem maior precisão (=10.0) nos três primeiros cenários e menor precisão (=0.5) no último cenário (Figura 4.5(b)).

O desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) foi pequeno, indicando que as distribuições *a posteriori* foram adequadamente amostradas. As medianas *a posteriori* para os parâmetros de interesse são coerentes entre todas as análises (Tabelas 4.2 e 4.3 e Figuras 4.6 e 4.7). Nos quatro cenários, a mediana *a posteriori* (e o desvio padrão) para  $N$  foram, respectivamente, iguais a 230.0 (17.45), 230.0 (17.38), 232.0 (17.95), 230.0 (17.59). A variabilidade do parâmetro tamanho da população,  $N$ , não foi sensível à escolha dos valores que definem a distribuição *a priori* dos parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$  e  $\delta$ . Entretanto, a estimativa *a posteriori* para  $N$  foi sensível à escolha dos valores para as respectivas distribuições *a priori* dos parâmetros que definem  $g_0$ , apresentando maior mediana no terceiro cenário.

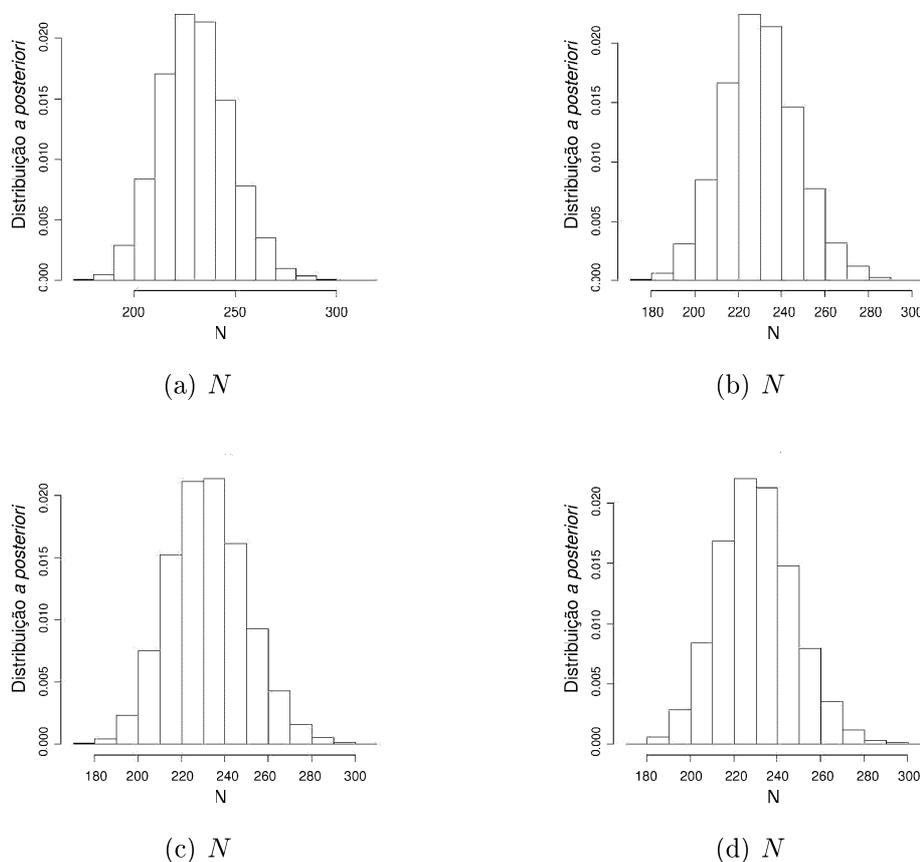


Figura 4.6: Gráfico das distribuições *a posteriori* para o parâmetros  $N$ , obtidas tendo em conta as distribuições *a priori* dos quatro cenários. Figuras a), b), c) e d), respectivamente.

A forma das distribuições *a posteriori* dos parâmetros  $\lambda$ ,  $p_1$ ,  $\sigma^2$  e  $\delta$  não foi sensível quando a escolha dos valores que definem as distribuições *a priori* apresentados nos quatro cenários, ver Figura 4.7. Os parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$  e  $\tau_\delta$  obtiveram diferentes formas para as respectivas distribuições *a posteriori*, sendo os quatro primeiros no cenário 2 e o parâmetro  $\tau_\delta$  no quarto cenário. As estimativas *a posteriori* desses parâmetros também foram diferente nos quatro cenários.

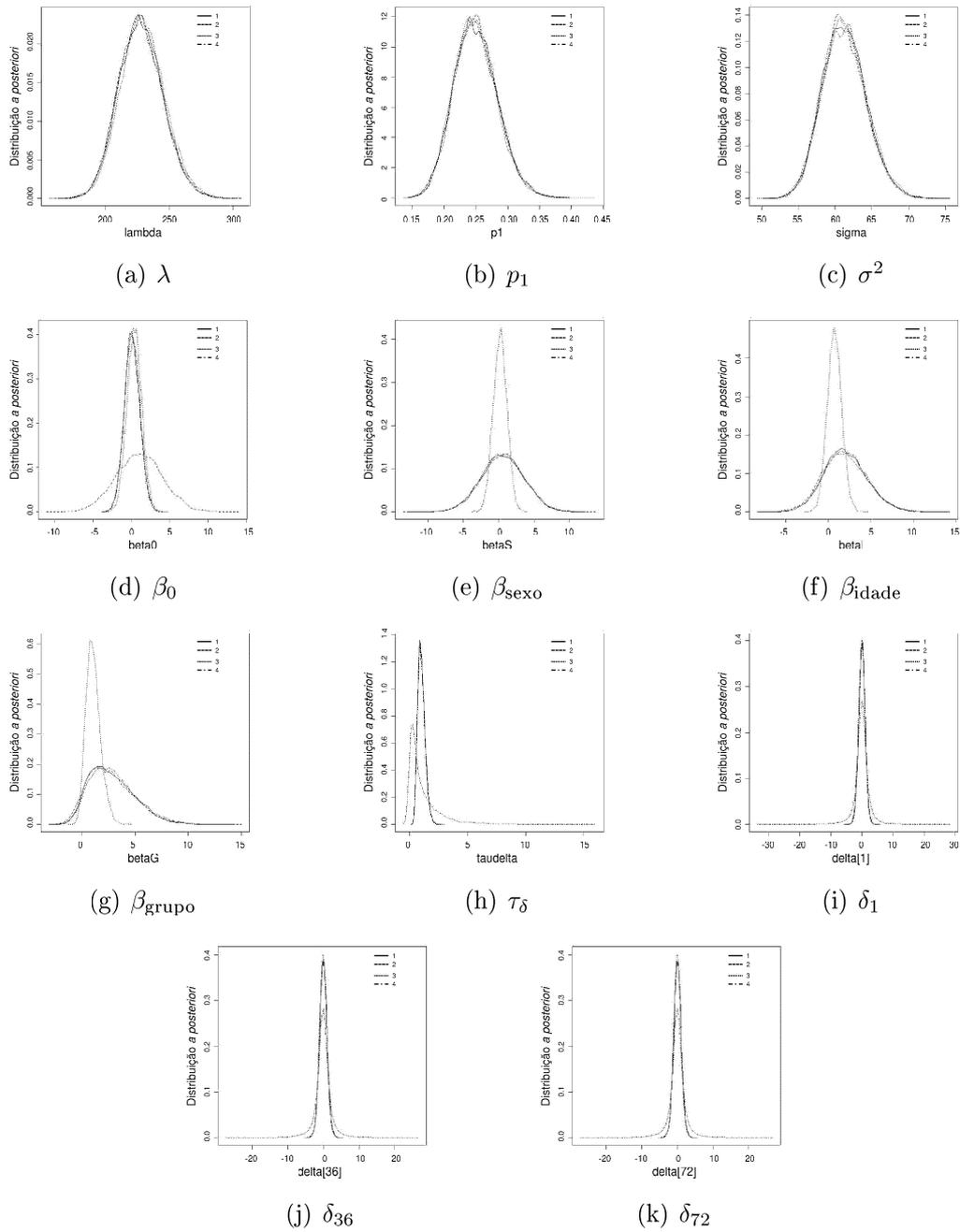


Figura 4.7: Gráfico das distribuições *a posteriori* para os parâmetros  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$ ,  $\tau_{\delta}$ ,  $\delta_1$ ,  $\delta_{36}$  e  $\delta_{72}$ , obtidas tendo em conta as distribuições *a priori* dos quatro cenários.

Tabela 4.2: Resultado da análise de sensibilidade para os parâmetros  $N$ ,  $\lambda$ ,  $p_1$ ,  $\sigma^2$ ,  $\beta_0$ ,  $\beta_{\text{sexo}}$ , tendo em conta as distribuições *a priori* nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição *a posteriori* é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD.

Distribuição <i>a priori</i>	Características da cadeia de Markov						Distribuição <i>a posteriori</i>		
	$se_b$	5	10	20	50	Mediana	$sd$	I.C./HPD	
$N$	1	0.0483	0.0775	0.0065	0.0002	-0.0005	230.00	17.45	[194.0,262.0]
	2	0.0495	0.0828	0.0067	0.0036	>-0.0001	230.00	17.38	[196.0,263.0]
	3	0.0515	0.0891	0.0075	0.0014	-0.0002	232.00	17.95	[197.0,267.0]
	4	0.0497	0.0875	0.0104	0.0034	0.0022	230.00	17.59	[195.0,263.0]
$\lambda$	1	0.0441	0.0608	0.0064	-0.0005	-0.0025	227.50	17.12	[194.6,261.4]
	2	0.0450	0.0653	0.0046	0.0003	-0.0003	227.50	17.44	[193.7,262.0]
	3	0.0464	0.0717	0.0063	0.0024	-0.0002	229.20	17.43	[196.2,264.2]
	4	0.0448	0.0674	0.0064	0.0002	0.0002	227.50	17.33	[194.9,261.4]
$p_1$	1	<0.0001	0.0386	0.0040	0.0003	-0.0024	0.2487	0.034	[0.181,0.315]
	2	<0.0001	0.0411	0.0022	0.0009	0.0016	0.2490	0.034	[0.186,0.318]
	3	<0.0001	0.0423	0.0024	-0.0009	-0.0005	0.2471	0.034	[0.185,0.317]
	4	<0.0001	0.0426	0.0050	0.0036	0.0002	0.2487	0.034	[0.182,0.315]
$\sigma^2$	1	0.0050	0.0081	0.0010	0.0023	0.0002	61.17	2.963	[55.82,67.32]
	2	0.0050	0.0067	0.0007	0.0018	0.0022	61.04	2.975	[55.54,67.09]
	3	0.0051	0.0083	-0.0003	0.0027	0.0030	61.23	3.008	[55.44,67.21]
	4	0.0050	0.0073	0.0019	0.0016	0.0022	61.17	2.959	[55.56,67.06]
$\beta_0$	1	0.0031	0.1295	0.0204	-0.0008	0.0031	0.0843	1.002	[-1.867,2.045]
	2	0.0111	0.2508	0.0729	0.0063	-0.0006	1.0320	3.011	[-4.788,6.873]
	3	0.0034	0.2667	0.0604	0.0113	-0.0021	0.3861	0.972	[-1.576,2.233]
	4	0.0033	0.1748	0.0330	0.0027	0.0.005	0.0990	0.996	[-1.888,2.044]
$\beta_{\text{sexo}}$	1	0.0100	0.1859	0.0406	0.0014	0.0023	0.5353	2.986	[-5.329,6.290]
	2	0.0095	0.1455	0.0245	0.0015	-0.0002	0.5917	3.002	[-5.290,6.309]
	3	0.0031	0.1506	0.0248	0.0022	0.0019	0.2084	0.973	[-1.762,2.036]
	4	0.0099	0.1753	0.0331	0.0022	0.0003	0.5862	3.000	[-5.377,6.289]

Tabela 4.3: Resultado da análise de sensibilidade para os parâmetros  $\beta_{idade}$ ,  $\beta_{grupo}$ ,  $\tau_{\delta}$ ,  $\delta_1$ ,  $\delta_{36}$  e  $\delta_{72}$ , tendo em conta as distribuições *a priori* nos cenários 1 a 4. As características das cadeias de Markov são expressas pelo desvio padrão entre as médias das réplicas das cadeias ( $se_b$ ) e a estimativa da função de autocorrelação (para espaçamento igual a 5, 10, 20 e 50). O resumo da distribuição *a posteriori* é dado pela mediana, pelo desvio padrão ( $sd$ ) e pelo intervalo de credibilidade/HPD.

Distribuição <i>a priori</i>	Características da cadeia de Markov					Distribuição <i>a posteriori</i>			
	$se_b$	Espaçamento				Mediana	$sd$	I.C./HPD	
		5	10	20	50				
$\beta_{idade}$	1	0.0097	0.2859	0.0995	0.0151	0.0005	1.897	2.495	-2.757,7.023
	2	0.0105	0.3130	0.1146	0.0178	-0.0014	1.785	2.608	-2.947,7.328
	3	0.0034	0.2953	0.1060	0.0179	-0.0002	0.722	0.857	-0.931,2.401
	4	0.0101	0.2962	0.1051	0.0200	0.0058	1.947	2.540	-2.735,7.246
$\beta_{grupo}$	1	0.0086	0.3273	0.1117	0.0139	-0.0008	2.518	2.122	-0.887,6.885
	2	0.0089	0.3268	0.1174	0.0212	0.0035	2.477	2.159	-1.140,6.892
	3	0.0028	0.3229	0.1192	0.0170	0.0006	1.002	0.685	-0.239,2.403
	4	0.0090	0.3421	0.1340	0.0298	0.0155	2.676	2.144	-0.937,7.121
$\tau_{\delta}$	1	0.0019	0.5866	0.4279	0.2284	0.0351	0.971	0.313	0.449,1.642
	2	0.0019	0.5921	0.4336	0.2354	0.0427	0.970	0.317	0.426,1.624
	3	0.0019	0.5897	0.4341	0.2371	0.0337	0.972	0.318	0.439,1.626
	4	0.0140	0.9440	0.9279	0.8975	0.8165	0.618	1.477	0.002,3.968
$\delta_1$	1	0.0033	0.1470	0.0237	0.0039	0.0023	-0.006	1.057	-2.117,2.077
	2	0.0033	0.1409	0.0172	0.0006	0.0003	0.016	1.055	-2.077,2.094
	3	0.0033	0.1432	0.0263	0.0042	0.0020	-0.017	1.029	-2.057,1.976
	4	0.0097	0.2423	0.0789	0.0013	-0.0118	-0.008	2.650	-5.284,5.197
$\delta_{36}$	1	0.0034	0.1508	0.0263	0.0037	-0.0021	0.016	1.050	-2.022,2.118
	2	0.0033	0.1442	0.0273	0.0019	-0.0016	0.001	1.060	-2.100,2.027
	3	0.0032	0.1253	0.0201	0.0008	-0.0002	0.004	1.056	-2.117,2.054
	4	0.0110	0.2866	0.1136	0.0254	-0.0030	-0.006	2.684	-5.615,4.999
$\delta_{72}$	1	0.0033	0.1373	0.0250	0.0042	-0.0029	-0.006	1.046	-2.076,2.032
	2	0.0034	0.1505	0.0277	0.0018	0.0012	-0.017	1.054	-2.194,1.985
	3	0.0032	0.1172	0.0200	-0.0007	-0.0004	0.013	1.040	-2.035,2.040
	4	0.0102	0.2622	0.0833	0.0013	-0.0158	0.017	2.676	-5.331,4.817

### 4.3.3 Selecção de modelos

A selecção do modelo apropriado para a inferência estatística é um aspecto muito importante, uma vez que os modelos representam um conjunto de hipóteses biológicas. Levanta-se então o problema de quais as covariáveis, que definem os modelos, devem ser utilizadas? Na teoria clássica, o rácio de verosimilhança ou o critério de informação de Akaike são utilizados frequentemente para a selecção de modelos. Na inferência bayesiana é comum utilizar o critério de informação da deviance (DIC) para a selecção de modelos, que está implementado no WINBUGS. Contudo, o DIC não é o critério mais adequado para modelos hierárquicos (e.g. modelos de efeitos aleatórios) e Celeux *et al.* (2006) mostraram como fazer as devidas alterações para utilizar o DIC em modelos hierárquicos.

Neste trabalho utilizou-se a metodologia proposta por Link e Barker (2009) para fazer a selecção bayesiana de modelos, que requer a definição de uma matriz, neste caso designada por `num.par`, que incorpora os modelos de acordo com o respectivo número de parâmetros utilizados.

Neste estudo, a matriz contém um conjunto de sete modelos para estimar  $g_0$ , como indicado em (4.4). As colunas e as linhas representam, respectivamente, os sete modelos,  $M_1$  a  $M_7$ , e as covariáveis utilizadas no modelo logístico definido para  $g_0$ , i.e.  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$ ,  $\beta_{\text{grupo}}$  e  $\delta_i$ . Todos os 7 modelos incluem o efeito aleatório individual  $\delta_i$ , além do efeito global  $\beta_0$ .

O número 1 na célula da matriz indica que a respectiva variável não pertence ao modelo, caso contrário o número corresponde ao total de parâmetros que constituem o modelo. Por exemplo, o modelo  $M_6$  apresenta um total de 4 tipo de parâmetros:

o efeito aleatório individual  $\delta_i$ ,  $\beta_0$  e mais 2 parâmetros fixos,  $\beta_{idade}$  e  $\beta_{grupo}$ :

$$\text{num.par} = \begin{matrix} & M_1 & M_2 & M_3 & M_4 & M_5 & M_6 & M_7 \\ \beta_0 & \left[ \begin{array}{ccccccc} 3 & 3 & 3 & 4 & 4 & 4 & 5 \\ 3 & 1 & 1 & 4 & 4 & 1 & 5 \\ 1 & 3 & 1 & 4 & 1 & 4 & 5 \\ 1 & 1 & 3 & 1 & 4 & 4 & 5 \\ 3 & 3 & 3 & 4 & 4 & 4 & 5 \end{array} \right] \\ \beta_{\text{sexo}} & \\ \beta_{\text{idade}} & \\ \beta_{\text{grupo}} & \\ \delta_i & \end{matrix} . \quad (4.4)$$

A seguir, define-se um índice, na linguagem WINBUGS, para cada parâmetro que se quer incluir ou excluir nos modelos:

`Ind.BetaSexo<-equals(modelo,1)+equals(modelo,4)+equals(modelo,5)+equals(modelo,7)`

`Ind.BetaIdade<-equals(modelo,2)+equals(modelo,4)+equals(modelo,6)+equals(modelo,7)`

`Ind.BetaGrupo<-equals(modelo,3)+equals(modelo,5)+equals(modelo,6)+equals(modelo,7).`

O índice associado ao parâmetro  $\beta_{\text{sexo}}$  implica que esse esteja incluído nos modelos 1, 4, 5 e 7. Subsequentemente, para cada parâmetro envolvido no modelo é associada uma precisão definida como o seguinte produto:

$$T^* \text{ num.par}[i,\text{modelo}],$$

onde

$$T \sim \text{dgamma}(3.29, 7.8),$$

e  $i$  é um inteiro que indica a linha em que o parâmetro está representado na matriz `num.par`, dada em (4.4). Link e Barker (2009) assumiram que os modelos têm distribuição *a priori* Categórica com igual probabilidade, para cada modelo:

$$\text{modelo} \sim \text{dcat}(\text{p.modelo}[1:7]),$$

ver mais detalhes no código de WINBUGS para a selecção de modelos no Apêndice A.4.

#### 4.3.4 Resultados e análise do modelo

O desempenho do estimador da probabilidade de detecção na distância zero, baseado no modelo combinado não heterogéneo (*MCNH*) e o modelo combinado heterogéneo foi analisado e comparado com os modelos onde foi modelada apenas a heterogeneidade observável (*MCH*) ou foram modeladas ambas as heterogeneidades, a observável e a não observável (*MCH\**). De notar que os Modelos 1-7 são Modelos Combinados que incorporam factores observáveis e não observáveis.

Para a selecção dos modelos foram considerados dois casos:

- a) o modelo tal como apresentado na matriz (4.4), que toma em conta todos os parâmetros que definem o modelo logístico para  $g_0$ ;
- b) não se considera os parâmetros  $\delta$  para efeito de selecção de modelos. Neste caso, suprime-se a última linha da matriz (4.4) e subtrai-se 1 aos números apresentado na matriz, com excepção do número 1.

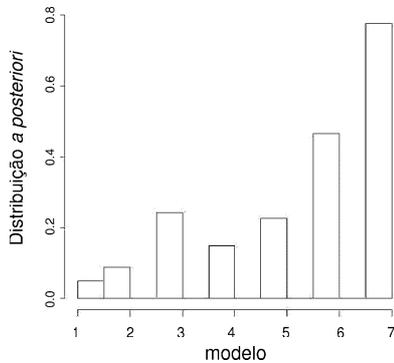
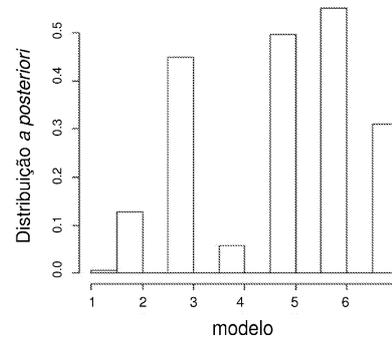
(a) Toma em conta  $\delta$ (b) Não toma em conta  $\delta$ 

Figura 4.8: Distribuição *a posteriori* para selecção de modelo. Os números e as barras designam os modelos e o seu respectivo valor de probabilidade *a posteriori*.

Na Figura 4.8, apresentam-se os gráficos da distribuição *a posteriori* para a selecção de modelos tendo em conta os dois casos referidos anteriormente, Figura 4.8(a) e 4.8(b), respectivamente. De acordo com a selecção de modelos bayesiana, no caso a), os modelos com maior probabilidade *a posteriori* de serem seleccionados foram  $M_6$  e  $M_7$ , sendo o último o seleccionado. Em geral, os modelos mais complexos apresentaram maior probabilidade de serem seleccionados e a variável grupo parece ser a mais importante na modelação da probabilidade de detecção na distância zero.

No caso b), os modelos  $M_3$ ,  $M_5$ ,  $M_6$  e  $M_7$  apresentaram maior probabilidade *a posteriori* de serem seleccionados. No entanto, o modelo que incorpora as covariáveis idade, tamanho do grupo e o efeito aleatório individual,  $M_6$ , foi o que apresentou maior probabilidade *a posteriori* de ser seleccionado. Todos os modelos com maior probabilidade *a posteriori* de serem seleccionados têm em comum a covariável tamanho de grupo. Por outro lado, modelos que não incluem a covariável tamanho de grupo apresentam menor probabilidade *a posteriori* de serem seleccionados.

Com base numa amostra de tamanho 9,900, adoptaram-se a mediana *a posteriori* e um intervalo de credibilidade de 95% para as estimativas. O sumário das estimativas *a posteriori* dos parâmetros para os modelos propostos para estimar  $g_0$  são apresentados na Tabela 4.4.

Considerando que 230 é o melhor valor de referência que se dispõe para o tamanho real da população de rebecos, a estimativa para  $N$  apresenta maior enviesamento, embora pequeno, no caso dos modelos  $M_1$  e  $M_4$ , com um enviesamento relativo negativo iguais a 5.22% e 2.61%, respectivamente. Os modelos  $M_6$  e  $M_7$  apresentam estimativas menos enviesadas.

Nos modelos com maior probabilidade *a posteriori* de serem rejeitados,  $g_0$  assumiu menor valor, quando comparado com as estimativas obtidas com  $MCH^*$  ( $M_6$ ), i.e. 0.875 para o modelo  $M_1$ , 0.95 para o modelo  $M_2$ , 0.936 para o modelo  $M_4$  e 0.989 para o modelo  $M_6$ . As estimativas para  $g_0$  foram também menores quando se considerou o modelo  $MCH$ . Não há grande diferença entre as estimativas para  $g_0$  dos modelos  $M_6$  e  $M_7$ , embora a estimativa do modelo  $M_6$  para  $g_0$  seja mais precisa que a estimativa obtida com o modelo 7. A distribuição *a posteriori* de  $g_0$  apresentou maior assimetria quando se considerou o modelo  $MCH^*$  que o modelo  $MCH$ , mas apresentaram a mesma forma (Figura 4.9).

Tabela 4.4: Estimativas e intervalo de credibilidade de  $N$  e estimativas para os parâmetros do modelo logístico para  $g_0$ , quando apenas a heterogeneidade observável foi modelada, e quando ambas, heterogeneidade observável e não observável, foram modeladas considerando diferente modelos,  $M_0 - M_7$ , Modelo Combinado Não Heterogêneo ( $MCNH$ ) e Modelo Combinado Heterogêneo ( $MCH$ ). Note-se que os Modelos 1-7 são Modelos Combinados Heterogêneo que incorporam factores observáveis e não observáveis ( $MCH^*$ ).

Modelos	$N$	$g_0$	$\beta_0$	$\beta_{\text{sexo}}$	$\beta_{\text{idade}}$	$\beta_{\text{grupo}}$
$M_1$	242.0 [206.0,282.0]	0.875 [0.751,0.978]	1.752 [0.759,3.859]	1.234 [-0.353,2.348]	-	-
$M_2$	235.0 [201.0,276.0]	0.950 [0.778,0.993]	1.001 [-0.425,2.067]	-	1.334 [0.343,2.739]	-
$M_3$	232.0 [200.0,271.0]	0.970 [0.856,0.995]	1.309 [-0.175,2.692]	-	-	0.927 [0.151,1.98]
$M_4$	236.0 [203.0,275.0]	0.936 [0.812,0.988]	0.691 [-0.871,1.736]	0.594 [-1.330,2.750]	1.362 [0.417,2.442]	-
$M_5$	235.0 [201.0,274.0]	0.951 [0.808,0.996]	0.546 [-1.367,2.236]	-0.0498 [-1.773,1.322]	-	1.208 [0.464,2.103]
$M_6$	231.0 [199.0,268.0]	0.989 [0.881,0.999]	0.871 [-0.865,2.177]	-	0.821 [-0.525,2.053]	1.198 [-0.136,2.428]
$M_7$	232.0 [197.0,267.0]	0.976 [0.868,1.0]	0.386 [-1.576,2.233]	0.208 [-1.762,2.036]	0.722 [-0.931,2.401]	1.002 [-0.239,2.403]
MCH	233.0 [200.0,272.0]	0.965 [0.816,0.999]	-0.183 [-1.843,2.04]	0.312 [-1.102,1.596]	0.505 [-0.857,1.679]	1.175 [0.004,2.285]
MCNH	235.0 [202.0,274.0]	0.948 [0.796,0.998]	-	-	-	-

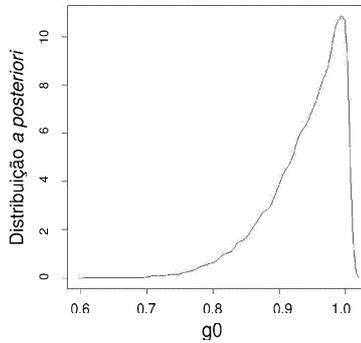
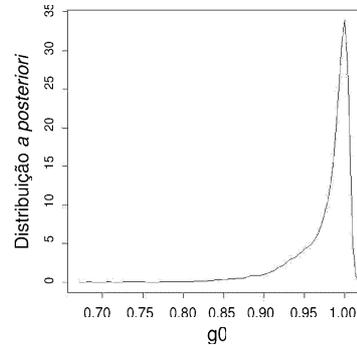
(a) Modelo *MCNH*(b) Modelo *MCH\**

Figura 4.9: Distribuição *a posteriori* do parâmetro  $g_0$  quando se considera o modelo não heterogêneo a) e o modelo heterogêneo que incorpora factores observáveis e não observáveis b).

## 4.4 Generalização do modelo de dupla plataforma e heterogeneidade

Nesta secção, apresenta-se um modelo combinado de captura-recaptura e transectos lineares heterogêneo generalizado, onde se considera a possibilidade de existirem mais de dois observadores na linha do transecto.

As detecções realizadas pelos observadores são considerados independentes entre si. A cada observador atribui-se uma função de detecção,  $g_o(y)$ , para  $o = 1, \dots, O$ , em que  $O$  corresponde ao número total de observadores. Aos indivíduos detectados pelo observador  $o$ , associa-se uma probabilidade de detecção de um indivíduo na

distância zero,  $g_0^o(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$ :

$$\begin{aligned} g_1(y|\sigma_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= g_0^1(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times g_1^*(y|\sigma_1) \\ g_2(y|\sigma_2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= g_0^2(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times g_2^*(y|\sigma_2) \\ &\vdots \\ g_o(y|\sigma_o, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= g_0^o(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times g_o^*(y|\sigma_o), \end{aligned}$$

$0 \leq y \leq w_o$ ,  $\sigma_o > 0$ , para  $o = 1, \dots, O$ , com  $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_o)$ . Os vectores parâmetros  $\boldsymbol{\beta}$ ,  $\boldsymbol{\delta}$  e  $\tau_\delta$  são os parâmetros que explicam o modelo logístico definido para cada um dos  $g_0^o(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$ .

Se a função de detecção for definida como sendo uma Semi-Normal, tem-se que  $g_1(y|\sigma_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) = g_0^1(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times e^{-\frac{1}{2}\left(\frac{y}{\sigma_1}\right)^2}$ ,  $g_2(y|\sigma_2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) = g_0^2(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times e^{-\frac{1}{2}\left(\frac{y}{\sigma_2}\right)^2}$ ,  $\dots$ ,  $g_o(y|\sigma_o, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) = g_0^o(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times e^{-\frac{1}{2}\left(\frac{y}{\sigma_o}\right)^2}$ .

Assim, para cada grupo de indivíduos detectados pelo observador  $o$ , faz-se corresponder a probabilidade de um indivíduo ter sido detectado por esse observador:

$$\begin{aligned} p_2^1(\sigma_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= g_0^1(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times p_2^{1*}(\sigma_1) \\ p_2^2(\sigma_2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= g_0^2(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times p_2^{2*}(\sigma_2) \\ &\vdots \\ p_2^o(\sigma_o, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= g_0^o(\boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \times p_2^{o*}(\sigma_o), \end{aligned}$$

onde cada  $\mu_o$  é aproximado por  $\mu_o = K_o \times \sigma_o$ , para  $K_o$  calculado de acordo com a informação sobre a variabilidade dos distâncias perpendiculares dos indivíduos detectados pelo observador  $o$ . Neste caso,  $p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$  passa a ser definida como sendo

a probabilidade de um indivíduo ter sido detectado por algum dos observadores, i.e:

$$\begin{aligned}
 p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= p_2^1(\sigma_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) + p_2^2(\sigma_2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) + \dots + p_2^O(\sigma_O, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \\
 &- p_2^1(\sigma_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)p_2^2(\sigma_2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) - p_2^1(\sigma_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)p_2^3(\sigma_2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \dots \\
 &- p_2^1(\sigma_1, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)p_2^2(\sigma_2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) \dots p_2^O(\sigma_O, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta).
 \end{aligned}$$

#### 4.4.1 Função de verosimilhança

Cada  $g_{0_i}^o$  associado aos respectivos indivíduos detectados pelo observador  $o$  é estimado tendo em conta o modelo  $MCH^*$ , anteriormente definido, com as devidas alterações:

$$\ln \left( \frac{g_{0_i}^o}{1 - g_{0_i}^o} \right) = \beta_0^o + \beta_{idade}^o \times idade_i^o + \beta_{sexo}^o \times sexo_i^o + \beta_{grupo}^o \times grupo_i^o + \delta_i^o,$$

e

$$g_{0_i}^o = \frac{e^{\beta_0^o + \beta_{sexo}^o \times sexo_i^o + \beta_{idade}^o \times idade_i^o + \beta_{grupo}^o \times grupo_i^o + \delta_i^o}}{1 + e^{\beta_0^o + \beta_{sexo}^o \times sexo_i^o + \beta_{idade}^o \times idade_i^o + \beta_{grupo}^o \times grupo_i^o + \delta_i^o}},$$

e

$$g_0^o = \sum_{i=1}^{n_2^o} \frac{g_{0_i}^o}{n_2^o},$$

para cada grupo de indivíduos  $n_2^o$  e  $o = 1, \dots, O$ .

Seja  $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_O)$ , onde cada  $\mathbf{z}_o$  corresponde ao vector que contém a informação sobre as características observáveis dos indivíduos, tais como o sexo,

a idade e o tamanho do grupo, que foram detectados pelo observador  $o$ . Sejam  $\boldsymbol{\beta} = (\boldsymbol{\beta}_0, \boldsymbol{\beta}_{\text{sexo}}, \boldsymbol{\beta}_{\text{idade}}, \boldsymbol{\beta}_{\text{grupo}})$  e  $\boldsymbol{\delta} = (\boldsymbol{\delta}^1, \dots, \boldsymbol{\delta}^O)$  os vectores com os parâmetros fixos e aleatórios, respectivamente, que definem os  $g_0^o$ . Note-se que  $\boldsymbol{\beta}_0 = (\boldsymbol{\beta}_0^1, \dots, \boldsymbol{\beta}_0^O)$ ,  $\boldsymbol{\beta}_{\text{sexo}} = (\boldsymbol{\beta}_{\text{sexo}}^1, \dots, \boldsymbol{\beta}_{\text{sexo}}^O)$ ,  $\boldsymbol{\beta}_{\text{idade}} = (\boldsymbol{\beta}_{\text{idade}}^1, \dots, \boldsymbol{\beta}_{\text{idade}}^O)$ ,  $\boldsymbol{\beta}_{\text{grupo}} = (\boldsymbol{\beta}_{\text{grupo}}^1, \dots, \boldsymbol{\beta}_{\text{grupo}}^O)$ ,  $\boldsymbol{\delta}^1 = (\boldsymbol{\delta}_1^1, \dots, \boldsymbol{\delta}_{n_2^1}^1), \dots, \boldsymbol{\delta}^O = (\boldsymbol{\delta}_1^O, \dots, \boldsymbol{\delta}_{n_2^O}^O)$ .

A função de verosimilhança para o modelo combinado é dada por:

$$\begin{aligned}
 f(\mathbf{n}, \mathbf{y}, \mathbf{z} | N, \lambda, p_1, \boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) &= \frac{N!}{n_{10}! n_{01}! n_{11}! (N-n)!} [p_1 (1 - p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))]^{n_{10}} \\
 &\times [(1 - p_1) p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)]^{n_{01}} [p_1 p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)]^{n_{11}} \\
 &\times [(1 - p_1) (1 - p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))]^{N-n} \\
 &\times \prod_{i=1}^{n_2^1} \frac{g^*(y_i | \sigma_1)}{\mu_1^*} \prod_{i=1}^{n_2^2} \frac{g^*(y_i | \sigma_2)}{\mu_2^*} \dots \prod_{i=1}^{n_2^O} \frac{g^*(y_i | \sigma_O)}{\mu_O^*}, \\
 0 &< p_1, p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta) < 1, \quad N \geq n. \tag{4.5}
 \end{aligned}$$

#### 4.4.2 Distribuição *a priori*

As distribuições *a priori* dos parâmetros desconhecidos são as definidas no capítulo 4. De forma análoga, utilizou-se a distribuição Gama Invertida (GI) para definir a distribuição *a priori* para os  $\sigma_o^2$ , i.e.:

$$\sigma_1^2 \sim GI(\alpha_1, \beta_1)$$

$$\sigma_2^2 \sim GI(\alpha_2, \beta_2)$$

⋮

$$\sigma_O^2 \sim GI(\alpha_O, \beta_O),$$

onde os hiperparâmetros  $\alpha_1, \alpha_2, \dots, \alpha_O$  e  $\beta_1, \beta_2, \dots, \beta_O$  são determinados a partir da informação disponível sobre variabilidade das distâncias.

### 4.4.3 Distribuição *a posteriori*

Seja  $\boldsymbol{\theta} = (N, \lambda, p_1, \boldsymbol{\sigma}^2, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)$  o vector dos parâmetros desconhecidos, e  $\mathbf{n}$ ,  $\mathbf{y}$  e  $\mathbf{z}$  os respectivos vectores com a informação observada (ver DAG do modelo combinado heterogéneo generalizado na Figura 4.10), sobre o pressuposto de independência entre os parâmetros, a distribuição *a posteriori* completa de  $\boldsymbol{\theta}$  é dada pelo produto entre a função de verosimilhança (4.5) e as distribuições *a priori* de cada um dos respectivos parâmetros desconhecidos:

$$\begin{aligned}
 h(\boldsymbol{\theta}|\mathbf{n}, \mathbf{y}, \mathbf{z}) &\propto f(\mathbf{n}, \mathbf{y}, \mathbf{z}|\boldsymbol{\theta})h(N|\lambda)h(\lambda)h(p_1)\prod_{o=1}^O h(\sigma_o^2) \\
 &\times \prod_{o=1}^O h(\beta_o^o)\prod_{o=1}^O h(\beta_{\text{sexo}}^o)\prod_{o=1}^O h(\beta_{\text{idade}}^o)\prod_{o=1}^O h(\beta_{\text{grupo}}^o)\prod_{o=1}^O h(\delta_i^o)h(\tau_\delta) \\
 &\propto \frac{N!}{n_{10}!n_{01}!n_{11}!(N-n)!} [p_1(1-p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))]^{n_{10}} \\
 &\times [(1-p_1)p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)]^{n_{01}} [p_1p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta)]^{n_{11}} \\
 &\times [(1-p_1)(1-p_2(\boldsymbol{\sigma}, \boldsymbol{\beta}, \boldsymbol{\delta}, \tau_\delta))]^{N-n} \frac{e^{-\sum_{o=1}^O \frac{\sum_{i=1}^{n_o^o} y_i^2}{2\sigma_o^2}}}{\prod_{o=1}^O (\sigma K_o)^{n_o^o}} \\
 &\times \frac{e^{-\lambda}\lambda^N \eta^\tau \lambda^{\tau-1} e^{-\eta\lambda}}{N! \Gamma(\tau)} \prod_{o=1}^O \frac{\beta_o^{\alpha_o} (\frac{1}{\sigma_o^2})^{\alpha_o-1} e^{-\frac{\beta_o}{\sigma_o^2}}}{\Gamma(\alpha_o)} \\
 &\times \frac{1}{4\pi^2} e^{-\sum_{o=1}^O \frac{(\beta_o^o)^2 + (\beta_{\text{sexo}}^o)^2 + (\beta_{\text{idade}}^o)^2 + (\beta_{\text{grupo}}^o)^2}{2}} \\
 &\times \prod_{o=1}^O \left( \left( \frac{\tau_\delta}{2\pi} \right)^{\frac{n_o}{2}} e^{-\frac{\tau_\delta \sum_{i=1}^{n_o} (\delta_i^o)^2}{2}} \frac{10^{10} \tau_\delta^9 e^{-10\tau_\delta}}{\Gamma(10)} \right).
 \end{aligned}$$

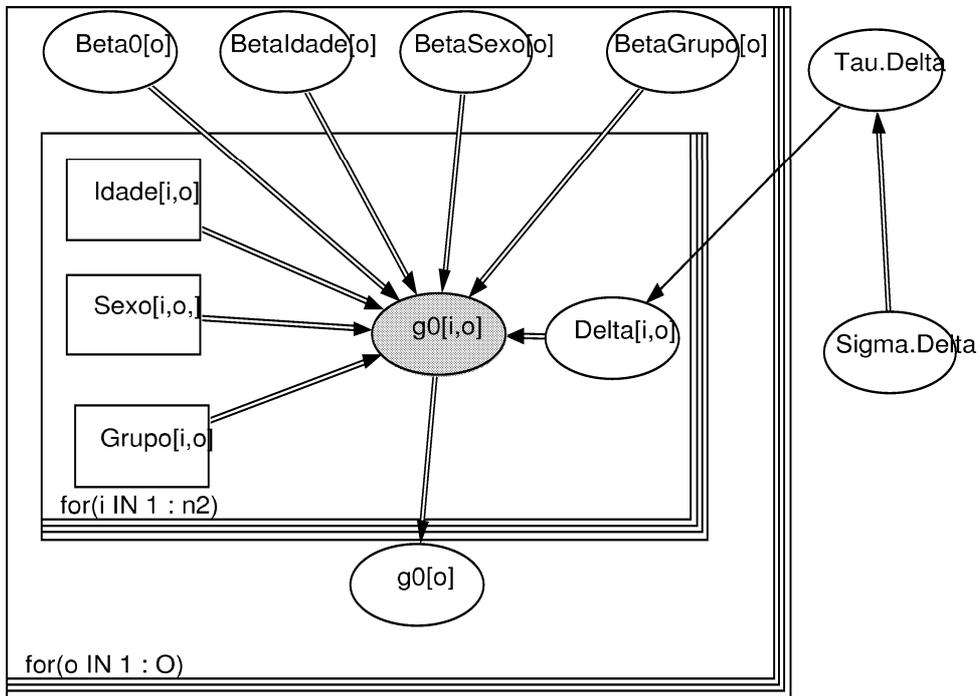


Figura 4.10: DAG do modelo combinado generalizado de captura-recaptura e transectos lineares, quando as probabilidades de captura dependem de características individuais e de outros factores não observáveis ( $MCHG^*$ ), para estimar a probabilidade de detectar um indivíduo na distância zero pelo observador  $o$ ,  $g_0^o$ . Os rectângulos maiores representam os ciclos em torno dos  $g_0^o$  individuais ( $g_{0_i}^o, i = 1, \dots, n_2^o, o = 1, \dots, O$ ).

# Capítulo 5

## Conclusão

## 5.1 Notas sobre MCMC

Uma das grandes vantagens do método MCMC é a possibilidade de se abordarem os modelos hierárquico bayesiano utilizando dados das amostragens realizadas pelos ecologistas. Desde dados com estruturas complexas, como é o caso de dados de captura-recaptura, as técnicas dos modelos hierárquicos bayesianos oferecem vantagens importantes em relação às análises tradicionais frequentistas. O uso dos métodos bayesianos, mostraram-se ainda eficazes quando há necessidade de incorporar covariáveis não observáveis nos modelos.

A implementação dos métodos de MCMC em programas como o WINBUGS tem sido muito importante em tornar possível a utilização dos modelos hierárquicos bayesianos. Contudo, ainda existem alguns problemas relacionados com o método de selecção para modelos hierárquicos bayesianos no WINBUGS. Recentemente, foi incorporado o método de Monte Carlo baseado em cadeias de Markov com saltos reversíveis (Green, 1995) no WINBUGS, através da rotina denominada JUMP, proposta em Lunn *et al.* (2006); Lunn (2008); Lunn *et al.* (2008). O método de Monte Carlo baseado em cadeias de Markov com saltos reversíveis permite construir uma cadeia de Markov reversível que tenha distribuição invariante desejada. Este método utiliza o algoritmo de M-H modificado, que inclui o Jacobiano da transformação no cálculo da probabilidade de aceitação do algoritmo de Metropolis-Hastings de modo a considerar a diferença existente na dimensão dos espaços paramétricos envolvidos no movimento. Deste modo, passa a ser possível fazer selecção de modelos, entre modelos com espaço paramétrico com diferentes dimensões.



Outra vantagem da abordagem bayesiana é poder garantir que a distribuição *a posteriori* de  $g_0$  e, conseqüentemente, qualquer estimativa correspondente se encontram no intervalo  $(0,1)$ , contrariamente ao que acontece com o método de MV não restringido.

Na Tabela 5.1, faz-se um breve resumo dos modelos de captura-recaptura, transectos lineares e combinados propostos por diferentes autores, utilizando diferentes abordagens para estimar os parâmetros associados. Indica-se para cada modelo proposto nesta tese o respectivo capítulo onde foi desenvolvido.

## 5.2 Principais desenvolvimentos e resultados

### 5.2.1 Modelação da heterogeneidade não observável em probabilidades de captura-recaptura

As análises feitas aos diferentes submodelos do modelo  $\mathcal{M}_{th}^*$  indicam que as estimativas *a posteriori* dos parâmetros desconhecidos,  $\theta$ , são muito sensíveis na escolha dos valores que definem as distribuições *a priori*. Entretanto, a estimativa *a posteriori* para o tamanho da população não se mostrou sensível na escolha dos valores que definem a distribuição *a priori*.

As estimativas *a posteriori* dos parâmetros desconhecidos obtidas com inferência bayesiana, têm maior precisão que as obtidas com o método de máxima verossimilhança.

As estimativas *a posteriori* dos parâmetros que definem as probabilidades de capturas são muito precisas, mas a estimativa *a posteriori* do desvio padrão do

parâmetro  $N$  é muito sensível em função da escolha dos valores para as distribuições *a priori* dos parâmetros  $\beta_0$ ,  $\beta_{\text{sexo}}$ ,  $\beta_{\text{idade}}$  e  $\beta_{\text{peso}}$ . A distribuição *a posteriori* de  $N$  apresenta uma assimetria positiva acentuada.

### 5.2.2 Modelo combinado sem heterogeneidade

Foi utilizado um modelo bayesiano hierárquico, onde é atribuída uma distribuição *a priori* Poisson com parâmetro  $\lambda$ , para o tamanho da população  $N$ , e para  $\lambda$  é atribuído uma distribuição *a priori* Gama com hiperparâmetros  $(\tau, \eta)$ .

As estimativas *a posteriori* do tamanho da população são diferentes das estimativas clássicas, que dependem exclusivamente da função de verossimilhança. A vantagem na utilização deste modelo bayesiano com a estrutura Poisson-Gama é que se pode incorporar informações *a priori* para o tamanho da população. Dessa forma, as estimativas incorporam as informações provenientes de históricos ou de pesquisas anteriores.

Portanto, o uso da estrutura Poisson-Gama para o modelo bayesiano de captura-recaptura pode ser uma alternativa eficaz para estimar  $N$ , onde se tem interesse em usar informações *a priori* do tamanho da população sem depender exclusivamente dos dados amostrais.

A análise dos dados simulados, nos quatro cenários, realçou o bom desempenho dos estimadores do modelo combinado de captura-recaptura e transectos lineares. As estimativas dos parâmetros associados ao modelo foram coerentes.

Nos dois exemplos apresentados, a estimativa obtida para o parâmetro  $g_0$ , utilizando a abordagem bayesiana, foi sempre menor que 1, ao contrário das estimativas

obtidas quando foi utilizada a metodologia clássica. No último caso, a probabilidade de detectar um animal na distância zero não tem significado do ponto de vista biológico. Da análise de sensibilidade realizada ao parâmetro  $g_0$ , no exemplo das latas de cerveja, concluiu-se que este é sensível a escolha da distribuição *a priori*, mas no entanto essa sensibilidade não afecta as estimativas para o tamanho da população. As distribuições *a posteriori* para  $g_0$ , resultantes da análise de sensibilidade, apresentaram todas uma assimetria negativa, a mesma forma, mas com escalas diferentes.

### 5.2.3 Modelo combinado que incorpora heterogeneidade observável e não observável em $g_0$

Neste capítulo,  $g_0$  foi estimado considerando eventuais factores de heterogeneidade. Nesse sentido, foi avaliado o efeito da heterogeneidade nas estimativas de  $g_0$  e nas probabilidades de detecção, relativa às características inerentes aos indivíduos e também modelando co-variáveis.

Concluiu-se que os modelos  $M_7$  (dependente das covariáveis sexo, idade, tamanho de grupo e factores não observáveis) e  $M_6$  (dependente das covariáveis idade, tamanho de grupo e factores não observáveis), considerados para  $g_0$ , têm maior probabilidade *a posteriori* de serem seleccionados que os outros modelos, respectivamente, para os dois casos de selecção de modelos apresentados. A variável tamanho de grupo é muito importante na estimação de  $g_0$ , uma vez que todos os modelos que consideraram a covariável tamanho de grupo tiveram maior probabilidade *a posteriori* de serem seleccionados, i.e.,  $M_6$ ,  $M_5$ ,  $M_3$  e  $M_7$ .

As estimativas dos modelos  $M_7$  e  $M_6$  são muito similares, e as estimativas para

$g_0$  são próximos de 1, com mediana *a posteriori*, respectivamente, iguais a 0.976 e 0.986. As estimativas obtidas utilizando os modelos *MCNH* (Modelo Combinado Não Heterogéneo) e *MCH* (Modelo Combinado Heterogéneo) são menores que as anteriormente apresentadas. Quando além da heterogeneidade observável se tenta modelar a heterogeneidade não observável (*MCH* versus *MCH\**, Modelo Combinado Heterogéneo - heterogeneidade observável e não observável), as estimativas da probabilidade de detectar um indivíduo na distância zero parecem aumentar, o que pode ser explicado pela presença de outras variáveis desconhecidas.

Da abordagem bayesiana resultaram estimativas para o tamanho da população menos enviesadas, se se assumir o facto de que o tamanho da população real ser 230 rebecos. Uma vez que nesta metodologia foi considerada mais informação, os parâmetros são mais precisos.

# Bibliografia

- Agresti, A. (1994). Simple capture-recapture models permitting unequal catchability and variable sampling effort. *Biometrics* **50**, 494–500.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* **19**, 716–723.
- Alho, J. M. (1990). Logistic regression in capture-recapture models. *Biometrics* **46**, 623–635.
- Alpizar-Jara, R. e Pollock, K. H. (1996). A combination line transect and capture-recapture sampling model for multiple observers in aerial surveys. *Journal Environmental and Ecological Statistics* **3**, 311–327.
- Alpizar-Jara, R. e Pollock, K. H. (1999). Combining line transect capture-recapture for mark-resighting studies. Em *Marine Mammal Survey and Assessment Methods*. (Garner, G. W, Amstrup, S. C., Laake, J. L., Manly, B. F. J., McDonald, L. L. e Robertson, D. G. eds). A. A. Balkema, Rotterdam, 99–114 pp.
- Ananda, M. M. A. (1997). Bayesian methods for mark-resighting surveys. *Communications in Statistics - Theory and Methods* **26**, 685–697.
- Andriolo, A., Piovezan, U., Costa, M. J. R. P, Laake, J. e Duarte, J. M. B. (2005). Aerial line transect survey to estimate abundance of marsh deer (*Blastocerus dichotomus*) (Illiger, 1815). *Brazilian Archives of Biology and Technology* **48**, 807–814.
- Bartolucci, F. e Forcina, A. (2001). Analysis of capture-recapture data with a Rasch-type model allowing for conditional dependence and multidimensionality. *Biometrics* **57**, 714–719.

- Basu, S. e Ebrahimi, N. (2001). Bayesian capture-recapture methods for error detection and estimation of population size: Heterogeneity and dependence. *Biometrika* **88**, 269–279.
- Bayes, T. R. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* **53**, 370–418.
- Becker, N. G. (1984). Estimating population size from capture-recapture experiments in continuous time. *Australian Journal of Statistics* **26**, 1–7.
- Bernillon, P., Lievre L., Pillonel, J., Laporte, A. e Costagliola, D. (2000). Record-linkage between two anonymous databases for a capture-recapture estimation of underreporting of AIDS cases: France 1990-1993. *International Journal of Epidemiology* **29**, 168–174.
- Bishop, Y. M. M., Fienberg, S. E. e Holland, P. W. (1975). *Discrete multivariate analysis: Theory and Practice*. MIT Press, Cambridge, 568 pp.
- Borchers, D. L. (1996). *Line transect estimation with uncertain detection on the trackline*. Tese de doutoramento, Universidade de Cape Town, Cape Town.
- Borchers, D. L., Buckland, S. T., Goedhart, P., Clarke, E. e Hedley, S. (1998a). Horvitz-Thompson estimators for double-platform line transect surveys. *Biometrics* **54**, 1221–1237.
- Borchers, D. L., Zucchini, W. e Fewster, R. (1998b). Mark-recapture models for line transect surveys. *Biometrics* **54**, 1207–1220.
- Borchers, D. L., Buckland, S. T. e Zucchini, E. (2002). *Estimating animal abundance: closed populations*. Springer-Verlag, Londres, 330 pp.
- Borchers, D. L., Laake, J. L., Southwell, C. e Paxton, C. G. M. (2006). Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. *Biometrics* **62**, 372–378.
- Brooks, S. P., Catchpole, E. A. e Morgan, B. J. T. (2000). Bayesian animal survival estimation. *Statistical Science* **15**, 357–376.

- Brooks, S. P., Catchpole, E. A. Morgan, B. J. T. e Harris, M. P. (2002). Bayesian methods for analysing ringing data. *Journal of Applied Statistics* **29**, 187–206.
- Buckland, S. T. e Turnock, B. J. (1992). A robust line transect method. *Biometrics* **48**, 901–909.
- Buckland, S. T., Anderson, D. R., Burnham, K. P. e Laake, J. L. (1993). Distance sampling: estimating abundance of biological populations. Chapman e Hall, Londres, 446 pp.
- Buckland, S. T., Goudie, I. B. J. e Borchers, D. L. (2000). Wildlife population assessment: past developments and future directions. *Biometrics* **56**, 1–12.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. e Thomas, L. (2001). *Introduction to distance sampling-estimating abundance of biological populations*. Oxford University Press, Oxford, 448 pp.
- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. e Thomas, L. (2004). *Advanced distance sampling - estimating abundance of biological populations*. Oxford University Press, Oxford, 434 pp.
- Burnham, K. P. (1972). *Estimation of population size in multinomial capture-recapture studies when capture probabilities vary among animals*. Tese de doutoramento, Oregon State University, Oregon.
- Burnham, K. P. e Overton, W. S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* **65**, 625–633.
- Burnham, K. P. e Overton, W. S. (1979). Robust estimation of population size when capture probabilities vary among animals. *Ecology* **60**, 927–936.
- Burnham, K. P., Anderson, D. R. e Laake, J. L. (1980). Estimation of density from line transect sampling of biological populations. *Wildlife Monographs* **72**, 3–202.
- Burnham, K. P. e Anderson, D. R. (2002). Model selection and multi-model inference: A practical information-theoretic approach. Springer-Verlag, Nova York, 496 pp.

- Burnham, K. P. e White, G. C. (2002). Evaluation of some random effects methodology applicable to bird ringing data. *Journal of Applied Statistics* **29**, 245–264.
- Burnham, K. P. e Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods Research* **33**, 261–304.
- Carretta, J. V., Taylor, B. L. e Chivers, S. J. (2001). Abundance and depth distribution of harbor porpoise (*Phocoena phocoena*) in northern California determined from a 1995 ship survey. *Fishery Bulletin* **99**, 29–39.
- Castledine, B. J. (1981). A Bayesian analysis of multiple-recapture sampling for a closed population. *Biometrika* **68**, 197–210.
- Celeux, G. Forbes, F. Robert, C. P. e Titterington, D. M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis* **1**, 651–674.
- Chao, A. (1989). Estimating population size for sparse data in capture-recapture experiments. *Biometrics* **45**, 427–438.
- Chao, A., Lee, S.-M. e Jeng, S. L. (1992). Estimating population size for capture-recapture data when capture probabilities vary by time and individual animal. *Biometrics* **48**, 201–216.
- Chao, A., Chu, W. e Hsu, C.-H. (2000). Capture-recapture when time and behavioral response affect capture probabilities. *Biometrics* **56**, 427–433.
- Chao A. (2001). An overview of closed capture-recapture models. *Journal of Agricultural, Biological and Environmental Statistics* **6**, 158–175.
- Chao, A., Yip, P., Lee, S.-M. e Chu, W. (2001). Population size estimation based on estimating functions for closed capture-recapture models. *Journal of Statistical Planning and Inference* **92**, 213–232.
- Chapman, D. G. (1951). Some properties of the hypergeometric distribution with applications to zoological sample censuses. *University of California Public Statistics* **1**, 131-160.

- Char, B. W., Geddes, K. O., Gonnet, G. H., Leong, B. L., Monagan, M. B. e Watt, S. M. (1991). *Maple V library reference manual*. Springer-Verlag, Berlin, 686 pp.
- Chavez-Demoulin, V. (1999). Bayesian inference for small-sample capture-recapture data. *Biometrics* **55**, 727–731.
- Chen, S. X. (2000). Animal abundance estimation for independent observer line transect surveys. *Environmental and Ecological Statistics* **7**, 285–299.
- Chen, Z. e Dunson, D. B. (2003). Random effects selection in linear mixed models. *Biometrics* **59**, 762–769.
- Clark, J. S., LaDeau, S. e Ibanez, I. (2004). Fecundity of trees and the colonization-competition hypothesis. *Ecological Monographs* **74**, 415–442.
- Clark, J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters* **8**, 2–14.
- Cochran, W. G. (1978). Laplace's ratio estimators. Em *Contributions to survey sampling and applied statistics*. (David, H. A. eds). Academic Press, Nova York, 3–10 pp.
- Cormack, R. M. (1988). Statistical challenges in the environmental sciences: a personal view. *Journal of the Royal Statistical Society* **151**, 201–210.
- Cormack, R. M. (1989). Log-linear models for capture-recapture. *Biometrics* **45**, 395–413.
- Coull, B. A. e Agresti, A. (1999). The use of mixed logit models to reflect heterogeneity in capture-recapture studies. *Biometrics* **55**, 294–301.
- Cowan, C. D. e Malec, D. (1986). Capture-recapture models when both sources have clustered observations. *Journal of the American Statistical Association* **81**, 347–353.

- Crampe, J. P. (1996). Caractéristiques bio-démographiques d'une population d'isards (*Rupicapra pyrenaica*) dans le Parc National des Pyrénées. Mémoire de Diplôme d'Études Supérieures Universitaires, Université Paul Sabatier, Toulouse.
- Crampe, J. P., Gaillard, J. M. e Loison, A. (2002). L'enneigement hivernal: un facteur de variation du recrutement chez l'isard (*Rupicapra pyrenaica pyrenaica*). *Canadian Journal of Zoology* **80**, 1306–1312.
- Darroch, J. N. (1958). The multiple-recapture census: I. Estimation of a closed population. *Biometrika* **45**, 343–359.
- Darroch, J. N. (1959). The multiple recapture census: II. Estimation when there is immigration or death. *Biometrika* **46**, 336–351.
- Dempster, A. P. (1974). The direct use of likelihood for significance testing. *Journal Statistics and Computing* **7**, 247–252.
- Dorazio, R. M. e Royle, J. A. (2003). Mixture models for estimating the size a closed population when capture rates vary among individuals. *Biometrics* **59**, 351–364.
- Durban, J. W. e Elston, D. A. (2005). Mark-recapture with occasion and individual effects: abundance estimation through bayesian model selection in a fixed dimensional parameter space. *Journal of Agricultural, Biological and Environmental Statistics* **10**, 291–305.
- Eguchi, T. e Gerrodette, T. (2009). A Bayesian approach to line-transect analysis for estimating abundance. *Ecological Modelling* **220**, 1620–1630.
- Ellison, A. M. (2004). Bayesian inference in ecology. *Ecology Letters* **7**, 509–520.
- Evans, M. A., Bonett, D. G. e McDonald, L. L. (1994). A general theory for modeling capture-recapture data from a closed population. *Biometrics* **50**, 396–405.
- Fienberg, S. E. (1972). The multiple-recapture census for closed populations and incomplete contingency tables. *Biometrika* **59**, 591–603.

- Fienberg, S. E., Johnson, M. S. e Junker, B. W. (1999). Classical multilevel and bayesian approaches to population size estimation using multiple lists. *Journal of the Royal Statistical Society: Series A* **162**, 383–405.
- Forbes, S. A. (1907). An ornithological cross-section of Illinois in autumn. *The Auk* **24**, 358–359.
- Forbes, S. A. e Gross, A. O. (1921). The orchard birds of an Illinois summer. *Illinois Natural History Survey Bulletin* **14**, 1–8.
- Gates, C. E. (1979). Line transect and related issues. Em *Sampling biological populations*. (Cormack, R. M, Patil, G. P., e Robson, D. S.). International Co-operative Publishing House, Fairland, 71–154 pp.
- Gelfand, A. E. e Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409.
- Gelfand, A. E. e Ghosh, S. K. (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika* **85**, 1–11.
- Gelman, A. e Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science* **7**, 457–511.
- Gelman, A., Carlin, J. B., Stern, H. S. e Rubin, D. B. (2004). *Bayesian data analysis*. Chapman e Hall/CRC Press, Boca Raton, 668 pp.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Analysis* **1**, 515–534.
- Genkin, A. Lewis, D. D. e Madigan, D. (2007). Large-scale Bayesian logistic regression for text categorization. *Technometrics* **49**, 291–304.
- George, E. I. e Robert, C.P. (1992). Capture-recapture estimation via Gibbs sampling. *Biometrika* **79**, 677–683.
- George, E. I. e McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association* **88**, 881–889.

- Ghosh, K. S. e Norris, J. L. (2005). Bayesian capture-recapture analysis and model selection allowing for heterogeneity and behavioral effects. *Journal of Agricultural, Biological and Environmental Statistics* **10**, 35–49.
- Gilks, W. R., Richardson, S. e Spiegelhalter, D. J. (1996). *Markov chain Monte Carlo in practice*. Chapman e Hall, Londres, 512 pp.
- Gimenez, O., Bonner, S., King, R., Parker, R. A., Brooks, S. P., Jamieson, L. E., Grosbois, V., Morgan, B. J. T. e Thomas, L. (2008). WINBUGS for population ecologists: Bayesian modeling using Markov chain Monte Carlo methods. Em *Modelling demographic processes in marked populations. Environmental and Ecological Statistics*. (Thomson, D. L., Cooch, E. G. e Conroy, M. J. eds). Springer, Nova York, 883–916 pp.
- Gould, W. R. e Nichols, J. D. (1998). Estimation of temporal variability of survival in animal populations. *Ecology* **79**, 2531–2538.
- Green, J. P. (1995). Reversible JUMP Markov chain monte Carlo computation and bayesian model determination. *Biometrika* **82**, 711–732.
- Hald, A. (1990). John Graunt and the Observations upon the Bills of Mortality, 1662. Em *A history of probability and statistics and their applications before 1750*. Wiley, Nova York, 586 pp.
- Harley, S. J., Myers, R. A. e Field, C. A. (2004). Hierarchical models improve abundance estimates: spawning biomass of hoki in Cook Strait, New Zealand. *Ecological Applications* **14**, 1479–1494.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109.
- Hoeting, J. A., Madigan, D., Raftery, A. E. e Volinsky, C. T. (1999). Bayesian model averaging: A tutorial (with Discussion). *Statistical Science* **14**, 382–401.
- Holzmann, H. Munk, A. e Zucchini, W. (2006). On identifiability in capture-recapture models. *Biometrics* **62**, 934–936.

- Hook, E. B. e Regal, R. R. (1995). Capture-recapture methods in epidemiology: methods and limitations. *Epidemiology Reviews* **17**, 243–264.
- Horvitz, D. G. e Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**, 663–685.
- Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika* **76**, 133–140.
- Huggins, R. M. (1991). Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* **47**, 725–732.
- Huggins, R. M. (2002). A parametrical empirical bayes approach to the analysis of capture-recapture experiments. *Australian and New Zealand Journal of Statistics* **44**, 55–56.
- Hunter, A. J. e Griffiths, H. J. (1978). Bayesian approach to estimation of insect population size. *Technometrics* **20**, 231–234.
- International Working Group for Disease Monitoring and Forecasting (1995a). Capture-recapture and multiplo-record system estimation, I: History and theoretical development. *American Journal of Epidemiology* **141**, 1047–1058.
- International Working Group for Disease Monitoring and Forecasting (1995b). Capture-recapture and multiplo-record system estimation, I: Application in human diseases. *American Journal of Epidemiology* **141**, 1059–1088.
- Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **186**, 453–461.
- Jeffreys, H. (1961). *Theory of probability*. Claredon Press, Oxford, 470 pp.
- Johnson, D. S. e Hoeting, J. A. (2003). Autoregressive models for capture-recapture data: A Bayesian approach. *Biometrics* **59**, 341–350.

- Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration - stochastic model. *Biometrika* **52**, 225-47.
- Jolly, G. M. (1982). Mark-recapture models with parameters constant in time. *Biometrics* **38**, 301-321.
- Karunamuni, R. J. e Quinn II, T. J. (1995). Bayesian estimation of animal abundance for line transect sampling. *Biometrics* **51**, 1325-37.
- Kenneth B. H. (2001). Temporal variation in fall bird density at high elevation habitats in southern British Columbia: an analysis using distance sampling. Forest Resources Management, Faculty of Forestry, University of British Columbia. Vancouver, 50 pp.
- King, R. e Brooks, S. P. (2001). On the Bayesian analysis of population size. *Biometrika* **88**, 317-336.
- King, R. e Brooks, S. P. (2002). Bayesian model discrimination for multiple strata capture-recapture data. *Biometrika* **89**, 785-806.
- Laake, J. L. e Borchers, D. L. (2004). Methods for incomplete detection at distance zero. Em *Advanced Distance Sampling*. (S. T. Buckland, D. R. Anderson, K. P. Burnham, J.L. Laake, D. L. Borchers, e L.Thomas, eds).Oxford University Press, Oxford, 108-189 pp.
- Laake, J. L., Calambokdis, J., Osmeck, S. D. e Rugh, D. J. (1997). Probability of detecting harbour porpoise from aerial surveys: Estimating  $g(0)$ . *Journal of Wildlife Management* **61**, 63-75.
- Laird, N. M. e Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* **38**, 963-974.
- Lambert, P. C., Sutton, A. J., Burton, P. R., Abrams, K. R. e Jones, D. R. (2005). How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WINBUGS. *Statistics in Medicine* **24**, 2401-2428.

- Laplace, P. S. (1812). *Théorie analytique des probabilités*. Courcier, Paris, 846 pp.
- Lee, S.-M. e Chao, A. (1994). Estimating population size via sample coverage for closed capture-recapture models. *Biometrics* **50**, 88–97.
- Lee, S.-M. e Chen, C. W. S. (1998). Bayesian Inference of population size for behavioural response models. *Statistica Sinica* **8**, 33–1247.
- Lehmann, E. L. e Casella, G. (1998). *Theory of point estimation*. Springer, Nova York, 589 pp.
- Leite, J. G., Rodrigues, J. e Milan, L. A. (2000). A Bayesian analysis for estimating the number of species in a population using nonhomogeneous Poisson process. *Statistics and Probability Letters* **48**, 153–161.
- Leopold, A. (1933). *Game management*. Charles Scribner's Sons, Nova York, 481 pp.
- Lincoln, F. C. (1930). *Calculating waterfowl abundance on the basis of banding returns*, U.S. G.P.O., Washington D.C., 4 pp.
- Link, W. A. (2003). Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics* **59**, 1123–1130.
- Link, W. A. e Barker, R. J. (2005). Modeling association among demographic parameters in analysis of open population capture-recapture data. *Biometrics* **61**, 46–54.
- Link, W. A. e Barker, R. J. (2006). Model weights and the foundations of multimodel inference. *Ecology* **87**, 2626–2635.
- Link, W. A. e Barker, R. J. (2009). *Bayesian inference: with ecological applications*. Academic Press Inc., 400 pp.
- Lloyd, C. J. e Yip, P. (1991). *A unification of inference for capture-recapture studies through Martingale estimating functions*. Em *Estimating equations*. (Godambe, V. P. eds). Clarendon Press, Oxford, 65–88 pp.

- Lloyd, C. J. (1994). Efficiency of martingale methods in recapture studies. *Biometrika* **81**, 305–315.
- Longford, N. T. (1993). *Random Coefficient Models*. Oxford, Clarendon, 288 pp.
- Lunn, D. J. (2003). WINBUGS development interface (WBDev). *ISBA Bulletin* **10**, 10–11.
- Lunn, D. J., Whittaker, J. C. e Best, N. (2006). A Bayesian toolkit for genetic association studies. *Genetic Epidemiology* **30**, 231–247.
- Lunn, D. J. (2008). WINBUGS JUMP interface: User manual. MRC Biostatistics Unit. Disponível em <http://www.winbugs-development.org.uk/rjmc.html>. Consultado a 5 de Outubro de 2010.
- Lunn, D. J., Best, N. e Whittaker, J. (2008). Generic reversible jump MCMC using graphical models. *Statistics and Computing*. DOI: 10.1007/s11222-008-9100-0.
- Madigan, D. e York, J. C. (1997). Bayesian methods for estimation of the size of a close population. *Biometrika* **84**, 19–31.
- Manly, B. F. J., McDonald, L. L. e Garner, G. W. (1996). Maximum likelihood estimation for the double-count method with independent observers. *Journal of Agricultural, Biological, and Environmental Statistics* **1**, 170–189.
- Marsh, H. e Sinclair, D. F. (1989). Correcting for visibility bias in strip transect aerial surveys for aquatic fauna. *Journal of Wildlife and Management* **53**, 1017–1024.
- McCarthy, M. A. (2007). *Bayesian Methods for Ecology*. Cambridge University Press, Cambridge, 310 pp.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. e Teller, H. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics* **21**, 1087–1091.

- Mingoti, S. A. (2000). A stepwise Bayesian estimator for the total number of distinct species infinite populations: sampling by elements. *Journal of Applied Statistics* **275**, 651–670.
- Monteiro, J. F. G. (2001). *Estimadores de máxima verosimilhança para a combinação captura-recaptura e trajectos lineares e as suas propriedades*. Trabalho Fim de Curso, Departamento de Matemática, Universidade de Évora, Évora, 114 pp.
- Monteiro, J. F. G. e Alpizar-Jara, R. (2006). Estimación Bayesiana de  $g_0$  em amostragem por distâncias. Em *Ciência Estatística - Actas do XIII Congresso Anual da SPE*. (Canto e Castro, L., Martins, E. G., Rocha, C., Oliveira, M. F., Leal, M. M. e Rosado, F. eds.). Sociedade Portuguesa de Estatística, Ericeira, 501–510 pp.
- Monteiro, J. F. G. e Alpizar-Jara, R. (2008). Modelando heterogeneidade não observável em probabilidades de captura-recaptura. Em *Estatística da Teoria à Prática - Actas do XV Congresso Anual da SPE*. (Hill, M. M., Ferreira, M. A., Dias, J. G., Salgueiro, M. F., Carvalho, H., Vicente, P. e Braumann, C. eds.). Sociedade Portuguesa de Estatística, Lisboa, 365-374 pp.
- Monteiro, J. F. G., Alpizar-Jara, R., Serrano, E., Crampe, J. P. e Pérez, J. M. (2008). Estimación Bayesiana de  $g_0$  usando el muestreo por distancias y su aplicación en las estimas de densidad de ungulados de montaña. Em *Tendencias Actuales en el Estudio y Conservación de los Caprinos Europeos*. (Torres, J. E. G., León, J. C-M., Paris, P. F. e de Llano Aguilar, R. C. eds). II Congreso Internacional del género CAPRA en Europa, Jaén, 207-217 pp.
- Monteiro, J. F. G., Alpizar-Jara, R., Serrano, E., Campre, J.-P. e Pérez, J. M. (2010). A Bayesian estimation approach for  $g_0$  estimation in line transect mark-resighting framework. (*em revisão*)
- Monteiro, J. F. G. e Alpizar-Jara, R. (2010). Bayesian estimation of probability at distance zero using a double platform. (*em preparação*)
- Morgan, B. J. T. e Ridout, M. S. (2007). A new mixture model for recapture heterogeneity. Technical Report.

- Norris, J. L., e Pollock, K. H. (1995). Capture-recapture model with heterogeneity and behavioural response. *Environmental and Ecological Statistics* **2**, 305–313.
- Norris, J. L. e Pollock, K. H. (1996). Nonparametric MLE under two closed capture-recapture models with heterogeneity. *Biometrics* **52**, 639–649.
- Okamura, H., Kitakado, T., Hiramatsu, K. e Mori, M. (2003). Abundance estimation of diving animals by the double-platform line transect method. *Biometrics* **59**, 512–520.
- Otis, D. L., Burnham, K. P., White, G. C. e Anderson, D. R. (1978). *Statistical inference from capture data on closed animal populations*. Wildlife Monographs **62**, 3–135.
- Otto, M. C. e Pollock, K. H. (1990). Size bias in line transect sampling: a field test. *Biometrics* **46**, 239–245.
- Palka, D. (1995). Abundance estimate of the gulf of Maine harbor porpoise. *Reports of the International Whaling Commission* **46**, 575–582.
- Parker, R. A. (1963). On the estimation of population size, mortality and recruitment. *Biometrics* **19**, 318–323.
- Paulino, C. D., Amaral Turkman, M. A. e Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa, 446 pp.
- Peréz, J. M., Serrano, E., Crampe, J. P. e Alpizar-Jara, R. (em preparação). Assessment of line transect sampling when estimating density of mountain ungulates.
- Petersen, G. G. J. (1896). The yearly immigration of young plaice into the Limfjord from the German Sea. *Report of Danish Biology Statistics* **6**, 1–48.
- Pledger, S. (2000). Unified maximum likelihood estimates for closed capture-recapture models using mixtures. *Biometrics* **56**, 434–442.
- Plummer, M., N. G. Best, K. Cowles e K. Vines (2006). CODA: Output analysis and diagnostics for MCMC. *R News* **6**, 7–11.

- Pollock, K. H. (1974). *The assumption of equal catchability of animals in tag-recapture experiments*. Tese de doutoramento, Cornell University, Ithaca.
- Pollock, K. H. e Otto, M. C. (1983). Robust estimation of population size in closed animal populations from capture-recapture experiments. *Biometrics* **39**, 1035–1049.
- Pollock, K. H., Hines, J. E. e Nichols, J. D. (1984). The use of auxiliary variables in capture-recapture and removal experiments. *Biometrics* **40**, 329–340.
- Pollock, K. H. e Kendall, W. L. (1987). Visibility bias in aerial surveys: a review of estimation procedures. *Journal of Wildlife Management* **51**, 502–510.
- Pollock, K. H. (1991). Modeling capture, recapture, and removal statistics for estimation of demographic parameters for fish and wildlife populations: past, present, and future. *Journal of the American Statistical Association* **86**, 225–238.
- Pollock, K. H. (2000). Capture-recapture models. *Journal of the American Statistical Association* **95**, 293–296.
- Pollock, K. H. (2002). The use of auxiliary variables in capture-recapture modelling: an overview. *Journal of Applied Statistics* **29**, 85–102.
- Pollock, K. H. e Alpizar-Jara, R. (2005). *Classical open-population capture-recapture models*. Em *Handbook of capture-recapture methods*. (Manly, B., McDonald T. e Amstrup, S. eds). Princeton University Press, 36–57 pp.
- Quang, P. X. e Becker, E. F. (1997). Combining line transect and double count sampling techniques for aerial surveys. *Journal of Agricultural, Biological, and Environmental Statistics* **2**, 230–242.
- Quang, P. X. e Becker, E. F. (1999). Aerial survey sampling of contour transects using double-count and covariate data. Em *Marine Mammal Survey and Assessment Methods*. (Garner, G. W, Amstrup, S. C., Laake, J. L., Manly, B. F. J., McDonald, L. L. e Robertson, D.G. eds). A. A. Balkema, Rotterdam, 87–97 pp.

- R Development Core Team (2006). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Disponível em <http://www.R-project.org>. Consultado a 5 de Outubro de 2010.
- Raftery, A. E. e Lewis, S. M. (1995). Implementing MCMC. Em *Markov chain monte Carlo in practice*. (Gilks, W. R., Richardson, S. e Spiegelhalter, D. J. eds). Chapman e Hall, Londres, 115–130 pp.
- Reay, N. (2005). Estimation of  $g(0)$  for bottlenose dolphin, grey seal and harbour porpoise in the Cardigan Bay SAC. Tese de mestrado, University of Wales Bangor, 106 pp.
- Rexstad, E. e Burnham, K.P. (1991). User's Guide for Interactive Program CAPTURE. Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, Colorado. Disponível em <http://www.mbr-pwrc.usgs.gov/software.html>. Consultado a 5 de Outubro de 2010.
- Rivest, L.-P. e Daigle, G. (2004). Loglinear models for the robust design in mark-recapture experiments. *Biometrics* **60**, 100–107.
- Royle, J. A. e Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities*. Elsevier, Academic Press, Amsterdão. Os códigos do livro estão disponíveis em <http://www.mbr-pwrc.usgs.gov/pubanalysis/roylebook/chapters.htm>.
- Royle, J. A., Dorazio, R. M. e Link, W. A. (2007). Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics* **16**, 67–85.
- Sanathanan, L. (1972). Models and estimation methods in visual scanning experiments. *Technometrics* **14**, 813–829.
- Schnabel, Z. E. (1938). The estimation of the total fish population of a lake. *American Mathematical Monthly* **45**, 348–352.

- Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464.
- Schwarz, C. J. e Seber, G. A. F. (1999). Estimating animal abundance: review III. *Statistical Science* **14**, 427–456.
- Schweder, T., Skaug, H. J., Langaas, M. e Dimakos. X. K. (1999). Simulated likelihood methods for complex double platform line transect surveys. *Biometrics* **55**, 678–687.
- Seber, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika* **52**, 249–259.
- Seber, G. A. F. (1982). The estimation of animal abundance and related parameters. 2nd ed. Charles Griffin and Company Ltd., Londres, 506 pp.
- Seber, G. A. F. (1986). A review of estimating animal abundance. *Biometrics* **42**, 267–292.
- Seber, G. A. F. (1992). A review of estimating animal abundance II. *International statistical Review* **60**, 129–166.
- Seber, G. A. F. e Schwarz, C. J. (2002). Capture-recapture: before and after EU-RING 2000. *Journal of Applied Statistics* **29**, 5–18.
- Skaug, H. J. e Fournier, D. A. (2006). Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. *Computational Statistics and Data Analysis* **51**, 699–709.
- Smith, B. J. (2005). *Bayesian output analysis program (BOA), Version 1.1.5*. The University of Iowa. Disponível em <http://www.public-health.uiowa.edu/boa>. Consultado a 5 de Outubro de 2010.
- Smith, E. P. e van Belle, G. (1984). Nonparametric estimation of species richness. *Biometrics* **40**, 119–129.

- Smith, P. J. (1988). Bayesian methods for multiple capture-recapture Surveys. *Biometrics* **44**, 1177–1189.
- Smith, P. J. (1991). Bayesian analyses for a multiple capture-recapture model. *Biometrika* **78**, 399–407.
- Sociedade Portuguesa de Estatística e Associação Brasileira de Estatística (2007). *Glossário estatístico inglês-português*. Disponível em [http://www.spestatistica.pt/?q=ver\\_documento/english-portuguesespe-abe.pdf](http://www.spestatistica.pt/?q=ver_documento/english-portuguesespe-abe.pdf). Consultado a 5 de Outubro de 2010.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. e Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society* **64**, 583–616.
- Spiegelhalter, D., Thomas, A., Best, N. e Lunn, D. (2004). *WINBUGS user manual, version 1.4*. Cambridge: MRC Biostatistics Unit. Disponível em <http://www.mrc-bsu.cam.ac.uk/bugs>. Consultado a 5 de Outubro de 2010.
- Stanghellini, E. e van der Heijden, P. G. M. (2004). A multiple-record systems estimation method that takes observed and unobserved heterogeneity into account. *Biometrics* **60**, 510–516.
- Stigler, S. M. (1986). *The history of statistics: the measurement of uncertainty before 1900*. The Belknap Press of Harvard University Press. Cambridge.
- Sturtz, S., U. Ligges e A. Gelman (2005). R2WINBUGS: A package for running WINBUGS from R. *Journal of Statistical Software* **12**, 1–16.
- Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics-Theory and Methods* **A7**, 13–26.
- Tardella, L. (2002). A new Bayesian method for nonparametric capture-recapture models in presence of heterogeneity. *Biometrika* **89**, 807–817.

- Thomas, L., Laake, J. L., Strindberg, S., Marques, F. F. C., Buckland, S. T., Borchers, D. L., Anderson, D. R., Burnham, K. P., Hedley, S. L., Pollard, J. H., Bishop, J. R. B. e Marques, T. A. (2006). *Distance 5.0. Release 2*. Research Unit for Wildlife Population Assessment. University of St. Andrews, UK. Disponível em <http://www.ruwpa.st-and.ac.uk/distance/>. Consultado a 5 de Outubro de 2010.
- van Dongen, S. (2006). Prior specification in Bayesian statistics: three cautionary tales. *Journal of Theoretical Biology* **242**, 90–100.
- Wang, X. (2002). *Bayesian analysis Of capture-recapture models*. Tese de doutoramento, Division of Mathematics, Towson University, Maryland.
- White, G. C., Anderson, D. R., Burnham, K. P. e Otis, D. L. (1982). Capture-recapture and removal methods for sampling closed populations. Los Alamos National Laboratory LA-8787-NERP. 235 pp.
- White, G. C e Burnham, K. P. (1999). Program MARK: survival estimation from populations of marked animals. *Bird Study* **46**, 120–138. Disponível em <http://warnercnr.colostate.edu/~gwhite/software.html>. Consultado a 5 de Outubro de 2010.
- Williams, B. K., Nichols, J. D. e Conroy, M. J. (2002). *Analysis and Management of Animal Populations*. Academic Press, San Diego, 1040 pp.
- Wintrebert, C., Zwinderman, A. H., Cam, E., Pradel, R. e van Houwelingen, J. C. (2005). Joint modeling of breeding and survival of *Rissa tridactyla* using frailty models. *Ecological Modelling* **181**, 203–213.
- Yip, P. S. F. (1989). An inference procedure for a capture and recapture experiment with time-dependent capture probabilities. *Biometrics* **45**, 471–479.
- Yip, P. S. F. (1991). A martingale estimating equation for a capture-recapture experiment in discrete time. *Biometrics* **47**, 1081–1088.

- Yip, P. S. F. e Fong, D. Y. T. (1993). Estimating population size by recapture sampling via estimating function. *Communications in Statistics–Stochastic Models* **9**, 179–193.
- Yip, P. S. F., Wan, E. C. Y. e Chan, K. S. (2001). A unified approach for estimating population size in capture-recapture studies with arbitrary removals. *Journal of Agricultural, Biological, and Environmental Statistics* **6**, 183–194.
- Yoshida, O. S., Leite, J. G. e Bolfarine, H. (1999). Stochastic monotonicity properties of Bayes estimation of the population size for capture-recapture data. *Statistics and Probability Letters* **42**, 257–266.

# Apêndice A

## Código WINBUGS

Lunn (2003) desenvolveu um algoritmo para a função Semi-Normal, a partir da distribuição Normal, de tal forma que `djl.dnorm.trunc( $\mu$ ,  $\tau$ , esquerda, direita)`, onde “`esquerda`” e “`direita`” representam os limites inferior e superior, respectivamente. Os hiperparâmetros  $\mu$  e  $\tau$  representam a média e a precisão. No WINBUGS, o parâmetro escala é dado pela precisão, em vez da variância, como é usual na definição de algumas distribuições.

## A.1 Modelo de captura-recaptura tipo $\mathcal{M}_{th}^*$

```

Model;
{
  #Informação por amostragem#
  for( i in 1 : n ) {
    A[i,1]<-(1-pa[i,1])*(1-pa[i,2])*(1-pa[i,3])*(1-pa[i,4])*(1-pa[i,5])*(1-pa[i,6])
    A[i,2]<-(1-pa[i,2])*(1-pa[i,3])*(1-pa[i,4])*(1-pa[i,5])*(1-pa[i,6])
    A[i,3]<-(1-pa[i,3])*(1-pa[i,4])*(1-pa[i,5])*(1-pa[i,6])
    A[i,4]<-(1-pa[i,4])*(1-pa[i,5])*(1-pa[i,6])
    A[i,5]<-(1-pa[i,5])*(1-pa[i,6])
    A[i,6]<-(1-pa[i,6])
    for( j in 1 : t ) {
      pa[i,j]<-p[i,j]
      g[i,j]<-p[i,j]/(1-(1-c[i,j])*A[i,j])
      x[i,j]~ dbern(g[i,j])          #Informação sobre o histórico das capturas
      .
      #Logaritmo da função de verosimilhança
      llike[i,j]<-x[i,j]*log(g[i,j])+(1-x[i,j])*log(1-g[i,j]) }
      . #Modelo logístico para a probabilidade de captura de um indivíduo i no momento j
      for( y in 1 : t-1 ) {
        logit(p[i,y])<-Beta0+BetaSexo*Sexo[i]+BetaIdade*Idade[i]+BetaPeso*Peso[i]
        +BetaTempo[y]+Delta[i] }
        logit(p[i,6])<-Beta0+BetaSexo*Sexo[i]+BetaIdade*Idade[i]+BetaPeso*Peso[i]+Delta[i]
        #Cálculo dos p[i]
        pest[i]<-1-(1-p[i,1])*(1-p[i,2])*(1-p[i,3])*(1-p[i,4])*(1-p[i,5])*(1-p[i,6])
        pinv[i]<-1/pest[i] }
        .
        #Distribuições a priori
        Beta0~dnorm(0,1.0)          #Explica o efeito global
        BetaSexo~dnorm(0,0.1)      #Explica o efeito sexo
        BetaIdade~dnorm(0,0.1)    #Explica o efeito idade
        BetaPeso~dnorm(0,0.1)     #Explica o efeito peso
        for( h in 1 : t-1 ) {BetaTempo[h] dnorm(0,0.1) } #Explica o efeito tempo
        for( f in 1 : n ) {Delta[f] dnorm(0,Tau.Delta) } #Explica o efeito aleatório
        Tau.Delta~dgamma(10.0, 10.0)
        Sigma.Delta<-1/sqrt(Tau.delta)
        .
        #Estimador do tipo Horvitz-Thompson
        Nest<-sum(pinv[])
      }
    }
  }

```

## A.2 Modelo *MCNH*

```

Model;
{
    #Algumas funções#
    p2<-g0*p2t1
    p2t1<-mu/w          #Probabilidade de um indivíduo ter sido avistado
    pT<-p1+p2-p1*p2    #Probabilidade de um indivíduo ter sido marcado ou avistado
    p[1]<-(p1*(1-p2))/pT
    p[2]<-((1-p1)*p2)/pT
    p[3]<-(p1*p2)/pT
    mu<-K*sigma
    sigma<-sqrt(1/s)
    .
    #Informação por amostragem#
    n<-sum(ni[1:3])      #Número de indivíduos marcados ou avistados
    n~dbin(pT,N)        #Informação da captura
    ni[1:3]~dmulti(p[],n) #Informação da captura
    for(j in 1:n2) { YTL1[j]<-YTL[j]*mu } #Informação das distâncias
    for(k in 1:n2) { YTL1[k]~djl.dnorm.trunc(5,s,0,w)}
    .
    #Informação a priori#
    p1~dunif(0,1)       #Probabilidade de um indivíduo ter sido marcado
    N~dpois(lambda)     #Tamanho da população
    lambda~dgamma(tau,eta)
    s~dgamma(alpha,beta) #Parâmetro que define a função de detecção
    g0~dunif(0,1)      #Probabilidade de um indivíduo na linha ter sido avistado
}

```

### A.3 Modelo *MCH*\*

```

Model;
{
    #Algumas funções#
    p2<-g0*mu/w          #Probabilidade de um indivíduo ter sido avistado
    pT<-p1+p2-p1*p2     # Probabilidade de um indivíduo ter sido marcado ou avistado
    p[1]<-(p1*(1-p2))/pT
    p[2]<-((1-p1)*p2)/pT
    p[3]<-(p1*p2)/pT
    mu<-K*sigma
    sigma<-sqrt(1/s)
    .
    #Informação por amostragem#
    n<-sum(ni[1:3])      #Número de indivíduos marcados ou avistados
    n~dbin(pT,N)        #Informação da captura
    ni[1:3]~dmulti(p[],n) #Informação da captura
    for(j in 1:n2) { YTL1[j] <-YTL[j]*mu } #Informação das distâncias
    for(k in 1:n2) { YTL1[k]~dj1.dnorm.trunc(196.57,s,0,w)} #Informação das distâncias
    .
    #Informação a priori#
    p1~dunif(0,1)       #Probabilidade de um indivíduo ter sido marcado
    N~dpois(lambda)     # Tamanho da população
    lambda~dgamma(tau,eta)
    s~dgamma(alpha,beta) #Parâmetro que define a função de detecção
    #Informação de g0
    #Probabilidade de um indivíduo ter sido avistado na distância zero
    for(l in 1:n2){logit(g0i[l]) <- Beta0+BetaIdade*Idade[l]+BetaSexo*Sexo[l]
    .
    +BetaGrupo*Grupo[l]+Delta[l]}
    Beta0~dnorm(0.0,1.0) #Explica o efeito global
    BetaIdade~dnorm(0.0,1.0) #Explica o efeito idade
    BetaSexo~dnorm(0.0,1.0) #Explica o efeito sexo
    BetaGrupo~dnorm(0.0,1.0) #Explica o efeito tamanho de grupo
    for( i in 1 : n2 ) { Delta[i]~dnorm(0,Tau.Delta)} #Explica o efeito aleatório
    Tau.Delta~dgamma(10.0, 10.0)
    g0<-sum(g0i[])/n2 #g0 médio
}

```

## A.4 Seleção de modelos para o modelo *MCH*\*

```

#Probabilidade de um indivíduo ter sido avistado na distância zero
for(1 in 1:n2){logit(g0i[1]) <- Beta0+Ind.BetaIdade*BetaIdade*Idade[1]
+Ind.BetaSexo*BetaSexo*Sexo[1]+Ind.BetaGrupo*BetaGrupo*Grupo[1]+Delta[1]}
Beta0~dnorm(0.0,Tau.Beta0) #Explica o efeito global
BetaIdade~dnorm(0.0,Tau.BetaIdade) #Explica o efeito idade
BetaSexo~dnorm(0.0,Tau.BetaSexo) #Explica o efeito sexo
BetaGrupo~dnorm(0.0,Tau.BetaGrupo) #Explica o efeito tamanho de grupo
for( i in 1 : n2 ) { Delta[i]~dnorm(0,Tau.Delta)} # Explica o efeito aleatório
#g0 médio
g0<-sum(g0i[])/n2
T~dgamma(3.29,7.8)
Tau.Beta0<-T*num.par[1,modelo]
Tau.BetaSexo<-T*num.par[2,modelo]
Tau.BetaIdade<-T*num.par[3,modelo]
Tau.BetaGrupo<-T*num.par[4,modelo]
Tau.Delta<-T*num.par[5,model]
modelo~dcat(p.modelo[1:7])
for (i in 1:7){Ind.Modelo[i]<-equals(modelo,i)}
for (i in 1:7){p.modelo[i]<-1/7}
Ind.BetaSexo<-equals(modelo,1)+equals(modelo,4)+equals(modelo,5)+equals(modelo,7)
Ind.BetaIdade<-equals(modelo,2)+equals(modelo,4)+equals(modelo,6)+equals(modelo,7)
Ind.BetaGrupo<-equals(modelo,3 )+equals(modelo,5)+equals(modelo,6)+equals(modelo,7)

```

## A.5 Modelo *MCHG*\*

```

Model;
{
#Algumas funções#
p[1]<-(p1*(1-p2))/pT
p[2]<-((1-p1)*p2)/pT
p[3]<-(p1*p2)/pT
for(o in 1:O) {p20[o]<-g0[o]*mu[o]/w[o]
p2<-sum(p20[])-prod(p20[])-C #Probabilidade de um indivíduo ter sido avistado por
algum observador. Onde C corresponde-se as outras possíveis intersecções.
pT<-p1+p2-p1*p2 # Probabilidade de um indivíduo ter sido marcado ou avistado
por algum observador
for(o in 1:O) {mu[o]<-K[o]*sigma[o] }
for(o in 1:O) {sigma[o]<-sqrt(1/s[o])}
. #Informação por amostragem#
n<-sum(ni[1:3]) #Número de indivíduos marcados ou avistados
n~dbin(pT,N) #Informação da captura
ni[1:3]~dmulti(p[],n) #Informação da captura
for( o in 1 : O ) {
for(j in 1:n2o[o]) { YTL1[j,o] <-YTL[j,o]*mu[o] } #Informação das distâncias
for(k in 1:n2o[o]) { YTL1[k,o]~djl.dnorm(media[o],s[o],0,w[o])}
s[o]~dgamma(alpha[o],beta[o]) } #Parâmetro que define cada uma das
funções de detecção
. #Informação a priori#
p1~dunif(0,1) #Probabilidade de um indivíduo ter sido marcado
N~dpois(lambda) # Tamanho da população
lambda~dgamma(tau,eta)
#Informação dos g0[o] para cada grupo de indivíduos detectado pelo observador o
#Probabilidade de um indivíduo ter sido avistado na distância zero
for( o in 1 : O ) {
for(l in 1:n2o[o]){logit(g0i[l,o]) <- Beta0[o]+BetaIdade[o]*Idade[l,o]
. +BetaSexo[o]*Sexo[l,o]+BetaGrupo[o]*Grupo[l,o]+Delta[l,o]}
Beta0[o]~dnorm(0.0,1.0) #Explica o efeito global
BetaIdade[o]~dnorm(0.0,1.0) #Explica o efeito idade
BetaSexo[o]~dnorm(0.0,1.0) #Explica o efeito sexo
BetaGrupo[o]~dnorm(0.0,1.0) #Explica o efeito tamanho de grupo
for( i in 1 : n2 ) { Delta[i,o]~dnorm(0,Tau.Delta)} #Explica o efeito aleatório
Tau.Delta~dgamma(10.0, 10.0)
g0[o]<-sum(g0i[o])/n2[o] } #g0 médio
}

```

# Apêndice B

## Dados

## B.1 Dados de capturas de ratos

Tabela B.1: Dados de capturas de ratos (*Peromyscus maniculatus*) realizadas por V. Reid, na zona Este de Stuart Gulch em Colorado, EUA. Foram realizados 6 momentos de amostragem, de onde foram obtidas características individuais de 38 ratos, tais como o sexo (f=fêmea e m=macho), a idade (j=juvenil, sa=semi-adulto e a=adulto) e o peso (em gramas).

Sexo	Idade	Peso(g)	Capturas	Sexo	Idade	Peso(g)	Capturas
m	j	12	1 1 1 1 1 1	m	j	13	0 1 1 0 1 0
f	j	15	1 0 0 1 1 1	f	j	5	0 1 0 1 0 1
m	j	15	1 1 0 0 1 1	f	a	20	0 1 0 0 0 1
m	j	15	1 1 0 1 1 1	m	j	12	0 1 0 0 1 1
m	j	13	1 1 1 1 1 1	f	j	6	0 0 1 0 0 0
m	a	21	1 1 0 1 1 1	f	a	22	0 0 1 1 1 1
m	j	11	1 1 1 1 1 0	f	j	10	0 0 1 0 1 1
m	sa	15	1 1 1 0 0 1	f	j	14	0 0 1 1 1 1
m	j	14	1 1 1 1 1 1	f	a	19	0 0 1 0 0 0
m*	j	14	1 1 0 1 1 1	f*	a	19	0 0 1 0 1 0
m	j	13	1 1 0 1 1 1	f	a	20	0 0 0 1 0 0
f	a	22	1 1 1 0 1 1	m	sa	16	0 0 0 1 1 1
m	j	14	1 1 1 1 1 1	f	j	11	0 0 0 1 1 0
m	j	11	1 0 1 1 1 0	m	j	14	0 0 0 0 1 0
f	j	10	1 0 0 1 0 0	f	j	11	0 0 0 0 1 0
f	a	23	0 1 0 0 1 0	m	a	24	0 0 0 0 1 0
f	j	7	0 1 1 0 0 1	m	j	9	0 0 0 0 0 1
m	j	8	0 1 0 0 0 1	m	sa	16	0 0 0 0 0 1
m	a	19	0 1 0 1 0 1	f	a	16	0 0 0 0 0 1

Nota: \*Esses dois indivíduos estão em falta no Apêndice I de Huggins (1991), adaptado de Yip *et al.* (2001).

## B.2 Experiência de Otto e Pollock (1990)

Tabela B.2: Dados da experiência feita por Otto, em 1982 (Otto e Pollock, 1990), com latas castanhas de cerveja, num campo perto de Raleigh, nos Estados Unidos. Cada grupo individual de distâncias perpendiculares das latas em relação à linha central do transecto é apresentado juntamente com quantos dos 9 observadores que o avistaram. O número 0 significa que a lata de cerveja não foi avistada.

Tamanho dos grupos							
1		2		3		4	
Distância	Visto por	Distância	Visto por	Distância	Visto por	Distância	Visto por
0.11	8	0.93	9	0.88	9	0.14	9
2.18	9	1.11	9	1.77	9	0.26	9
2.19	9	1.33	9	3.35	9	0.30	9
3.51	1	1.69	9	4.35	9	0.74	9
3.56	0	2.31	9	5.98	0	1.12	9
3.75	9	4.15	8	7.24	0	2.68	9
4.05	8	4.26	1	8.56	0	2.91	9
4.09	8	4.35	9	8.84	0	4.13	8
4.10	2	5.24	1	8.97	8	5.58	7
5.16	1	5.28	8	9.58	7	6.03	6
7.26	0	5.48	2	9.61	8	7.27	4
7.34	0	5.87	0	10.34	0	7.63	9
7.80	0	7.20	6	10.88	0	7.73	0
8.00	0	7.31	7	10.98	0	7.90	9
8.42	1	9.52	0	11.40	0	8.35	1
8.70	0	9.53	0	12.19	0	8.51	9
8.82	0	11.59	0	12.57	0	9.20	0
9.28	0	11.69	0	12.91	0	9.69	2
9.75	0	11.78	0	13.69	0	11.79	1
10.46	0	12.80	0	13.95	0	11.88	0
10.54	0	13.03	0	15.99	0	13.04	0
12.17	0	13.63	0	16.20	0	13.49	0
12.74	0	14.54	0	16.54	0	14.60	0
13.85	0	15.59	0	16.63	0	14.80	0
14.57	0	15.86	0	16.72	0	14.86	0
14.75	0	16.20	0	16.94	0	17.37	0
15.27	0	16.44	0	17.13	0	17.68	0
15.28	0	17.01	4	17.15	0	17.82	0
15.75	0	17.44	0	17.90	0	18.84	0
17.94	0	18.57	0	18.00	0	18.94	0
18.89	0	18.94	0	18.04	0	18.96	0
18.91	0	19.59	0	18.25	0	19.26	0
19.75	0	19.63	0	18.71	0	19.64	0

### B.3 Dados de rebecos no ano 2002

Tabela B.3: Dados da amostragem, do ano 2002, das distâncias perpendiculares da população de rebecos (*Rupicapra p. pyrenaica*), em *Clot-Cayan*, Parque Nacional de Pirenéus.

635	635	663	694	694	464	1025	1025	1025	1025
1025	1025	1025	1025	1025	1025	1025	1025	1025	1025
800	800	800	800	775	750	750	652	652	652
800	875	875	800	800	800	800	600	600	650
650	650	650	650	650	650	600	600	720	593
593	530	530	620	593	593	593	593	593	593
593	593	1570	1570	1570	700	1580	610	1228	1228
1228	1021	1021	1021	1021	1021	1021	1021	1021	1021
1021	1021	1021	1021	1021	1021	1021	1021	1021	1021
1021	1021	1021	584	584	650	650	550	550	860

## B.4 Dados de rebecos no ano 2007

Tabela B.4: Dados da amostragem da população de rebecos (*Rupicapra p. pyrenaica*), em *Clot-Cayan*, do ano 2007, Parque Nacional de Pirenéus, França. Foram avistados 72 rebecos, dos quais foram registadas características individuais, tais como a idade (1=cria/juvenil e 2=adulto), o sexo (0=fêmea e 1=macho) e o tamanho do grupo do qual o animal pertencia.

Distâncias	Idade	Sexo	Grupo	Distâncias	Idade	Sexo	Grupo
200	2	1	5	200	2	1	1
200	2	0	5	70	2	0	1
200	2	0	5	170	2	1	1
200	2	0	5	230	2	1	5
200	2	0	5	230	2	1	5
230	2	1	1	230	2	1	5
230	2	0	2	230	2	1	5
230	2	0	2	230	2	1	5
200	2	1	2	250	2	1	4
200	2	0	2	250	2	1	4
180	2	0	2	250	2	1	4
180	1	1	2	250	2	1	4
78	2	1	1	60	2	1	2
163	2	0	5	60	2	0	2
163	1	1	5	80	2	1	1
163	2	0	5	64	2	1	1
163	1	1	5	265	2	1	3
163	1	1	5	265	2	1	3
70	2	0	2	265	2	1	3
70	2	0	2	252	2	0	2
50	2	1	1	252	1	0	2
225	2	0	4	247	2	0	6
225	2	0	4	247	2	0	6
225	1	0	4	247	2	0	6
225	1	0	4	247	2	0	6
225	2	1	3	247	2	0	6
225	2	1	3	247	2	0	6
225	2	1	3	200	2	1	6
170	2	0	1	200	2	0	6
240	2	1	7	200	2	0	6
240	2	0	7	200	2	0	6
240	2	0	7	200	2	0	6
240	1	1	7	200	2	0	6
240	1	1	7	120	2	1	3
240	1	0	7	120	1	0	3
240	1	0	7	120	1	0	3

