

Universidade Federal de Minas Gerais

Departamento de Estatística

Programa de Pós-Graduação

Dissertação de Mestrado:

**Estimação Bayesiana
de probabilidades de captura para populações abertas
via modelo logístico.**

Mestranda: Pollyanna Vieira Gomes da Silva

Orientadora: Cibele Queiroz da Silva, Ph.D.

Belo Horizonte, fevereiro de 2006

Sumário

1	Introdução	1
1.1	Objetivos da dissertação	4
2	Métodos de captura-recaptura	8
2.1	Populações fechadas - Duas ocasiões amostrais	8
2.2	Populações fechadas - Mais de duas ocasiões amostrais	9
2.3	Populações abertas - O Modelo de Cormack	13
2.3.1	Notação	14
2.3.2	O Modelo de Cormack	15
2.4	O modelo de Jolly-Seber	15
2.5	A função de verossimilhança de Cormack com probabilidades de captura expressas em escala logística	16
3	Descrição do Modelo Bayesiano de Cormack	18
3.1	A distribuição condicional completa <i>a posteriori</i> de ϕ_{jg}	19
3.2	A distribuição condicional completa <i>a posteriori</i> de β	20
3.3	Descrição da função geradora de candidatos de β	21
3.3.1	Especificação da distribuição geradora de candidatos de β	22
3.3.2	Geração de dados a partir de uma Normal Multivariada	24
3.4	Um exemplo utilizando o modelo logístico	25
3.4.1	O modelo logístico	25
3.4.2	Verossimilhança associada ao modelo logístico	25
3.4.3	Cálculo da Matriz de Informação de Fisher	27

3.4.4	Análise Bayesiana para a regressão logística	28
3.4.5	Análise de dados utilizando o modelo logístico Bayesiano . . .	28
3.4.6	Comparação dos Resultados	30
4	Especificação de Modelos	33
4.1	Modelo 1	35
4.2	Modelo 2	36
4.3	Modelo 3	37
4.4	Modelo 4	38
4.5	Modelo 5	40
4.6	Modelo 6	40
4.7	Modelo 7	41
4.8	Modelo 8	42
4.9	Modelo 9	42
4.10	Modelo 10	43
4.11	Modelo 11	44
4.12	Modelo 12	44
5	Obtenção das derivadas parciais da log- condicional completa <i>a posteriori</i> de β para a obtenção de \hat{H}	46
5.1	Estudando χ_{tg}	46
5.1.1	Modelos 1, 5 e 9	46
5.1.2	Modelos 2, 6 e 10	47
5.1.3	Modelos 3, 7 e 11	47
5.1.4	Modelos 4, 8 e 12	48
5.2	Obtenção das derivadas parciais da log- condicional completa <i>a posteriori</i> de β para a obtenção de \hat{H} para os modelos 1, 2, 3 e 4	48
5.3	Obtenção das derivadas parciais da log- condicional completa <i>a posteriori</i> de β para obtenção de \hat{H} para os modelos 5, 6, 7, e 8.	53
5.4	Obtenção das derivadas parciais da log- condicional a posteriori de β para obtenção de \hat{H} para os modelos 9, 10, 11 e 12.	55

6	Funções geradoras de candidatos para ϕ	60
6.1	Caso 1 - Modelos 1, 5 e 9:	60
6.2	Caso 2 - Modelos 2, 6 e 10	61
6.3	Caso 3 - Modelos 3, 7 e 11:	62
6.4	Caso 4 - Modelos 4, 8 e 12:	62
7	Aplicação a dados reais	64
7.1	Descrição dos dados	64
7.2	Diagnósticos de Convergência	66
7.3	Sensibilidade das inferências à escolha da <i>priori</i>	69
7.4	Estimativas dos parâmetros e seleção dos modelos	71
7.4.1	Modelo 1	71
7.4.2	Modelo 2	72
7.4.3	Modelo 3	73
7.4.4	Modelo 4	73
7.4.5	Modelo 5	74
7.4.6	Modelo 6	74
7.4.7	Modelo 7	74
7.4.8	Modelo 8	78
7.4.9	Modelo 9	78
7.4.10	Modelo 10	78
7.4.11	Modelo 11	78
7.4.12	Modelo 12	80
7.4.13	Seleção de modelos e conclusões	83
8	Conclusões e considerações futuras	89
A	Descrição do Algoritmo de Metropolis-Hastings	96
A.1	Amostragem por Aceitação-Rejeição	97
A.2	Princípios do Algoritmo M-H	98

B	A escolha de $q(x, y)$	99
B.1	Cadeia do tipo Passeio Aleatório	99
B.2	Cadeia Independente	100
B.3	Cadeia do tipo Amostragem por Rejeição	100
C	A perspectiva Bayesiana e o fator de Bayes na seleção de modelos	101
C.1	O Fator de Bayes	101
C.1.1	Interpretação	102
C.1.2	Avaliação do <i>Fator de Bayes</i> e a aproximação BIC	103
C.1.3	O critério AIC	105
C.1.4	Alguns problemas com o Fator de Bayes na seleção de modelos	106
C.2	Distribuições Preditivas	107
C.2.1	Densidade Preditiva via validação cruzada e Pseudo Fator de Bayes	107
C.2.2	Aspectos computacionais do cálculo da densidade preditiva via validação cruzada	108
C.2.3	Densidade preditiva a posteriori	109
C.3	Estimador de Média harmônica de $P(Y)$	110

Lista de Figuras

2.1	Relações entre modelos de captura-recaptura.	10
3.1	Diagrama de dispersão - Frequentista X Bayesiano	30
3.2	Densidades estimadas de $\beta_0, \beta_1, \beta_2, \beta_7$	31
3.3	Gráficos de dispersão dos betas <i>versus</i> número da observação	32
7.1	Trajectoria da cadeia e densidade a posteriori - β_{01}	66
7.2	Trajectoria da cadeia e densidade a posteriori - β_{11}	67
7.3	Densidade a posteriori da taxa de sobrevivência	88
A.1	Princípios do algoritmo de Metropolis-Hastings.	97

Lista de Tabelas

2.1	História hipotética de captura de 4 animais em três ocasiões amostrais.	9
2.2	Modelos de captura-recaptura para populações fechadas	12
3.1	Código das variáveis associadas ao baixo peso de bebês ao nascer . . .	29
3.2	Estimações para o modelo Logístico Multivariado	30
4.1	Modelos	34
4.2	Parâmetros a estimar em cada um dos modelos	35
7.1	Tabela de capturas e recapturas - Baleias bowhead	64
7.2	Teste de Gelman e Rubin (CODA) - Modelo 5	68
7.3	Estimativas Modelos 5 - <i>Priori Beta</i> (0, 5; 0, 5)	69
7.4	Estimativas Bayesianas dos parâmetros do Modelo 5 - <i>Priori Beta</i> (1, 0; 1, 0)	70
7.5	Probabilidades de captura Modelo 5 - <i>Priori Beta</i> (0, 5; 0, 5)	70
7.6	Probabilidades de captura para o modelo 5 - <i>Priori Beta</i> (1, 0; 1, 0)) .	71
7.7	Estimativas Bayesianas dos parâmetros - Modelo 1	72
7.8	Estimativas Bayesianas dos parâmetros do Modelo 2	72
7.9	Estimativas Bayesianas dos parâmetros do Modelo 3	73
7.10	Estimativas Bayesianas dos parâmetros do Modelo 4	75
7.11	Estimativas Bayesianas dos parâmetros do Modelo 5	76
7.12	Estimativas Bayesianas dos parâmetros do Modelo 6	76
7.13	Estimativas Bayesianas dos parâmetros do Modelo 7	77
7.14	Estimativas Bayesianas dos parâmetros do Modelo 8	79
7.15	Estimativas Bayesianas dos parâmetros do Modelo 9	80
7.16	Estimativas Bayesianas dos parâmetros do Modelo 10	80

7.17	Estimativas Bayesianas dos parâmetros do Modelo 11	81
7.18	Estimativas Bayesianas dos parâmetros do Modelo 12	82
7.19	Estimador de média harmônica	83
7.20	Fator de Bayes obtidos por meio dos estimadores de média harmônica	84
7.21	Modelos ajustados através do <i>software</i> Mark	85
7.22	Probabilidades de captura e sobrevivência para o modelo 5 (Bayesiano)	86
7.23	Estimativas frequentistas para os parâmetros referentes ao modelo 5 .	87
C.1	Regras de interpretação de Jeffreys para o fator de Bayes	102

Capítulo 1

Introdução

No monitoramento de populações animais é comum o uso de técnicas destinadas à estimação de abundância e taxa de sobrevivência. Os valores estimados de tais parâmetros podem proporcionar subsídios para a tomada de decisões e estabelecimento de políticas de proteção ambiental. Neste sentido, os métodos de captura-recaptura são amplamente utilizados na prática.

Nosso interesse, nesse trabalho, reside na modelagem de dados de captura-recaptura de uma população de mamíferos, a população das baleias bowhead, *Balaena mysticetus*, conhecida como baleia da Groenlândia. No entanto, as técnicas a serem apresentadas podem ser usadas para qualquer espécie que esteja sendo monitorada através de estudos de foto-identificação.

Passaremos a descrever, sumariamente, aspectos históricos e do processo de reconhecimento dos animais.

Segundo Zeh *et al.* (2002), o estoque das baleias bowhead (*Balaena mysticetus*), nos mares Bering-Chukchi-Beaufort (Alaska) foi severamente prejudicado pela caça comercial na última metade do século XIX. A caça vem sendo regulamentada pela Comissão Baleeira Internacional (International Whaling Commission - IWC) que estipula cotas destinadas à subsistência dos esquimós. Recentemente, essas cotas vêm sendo estipuladas de acordo com cálculos Bayesianos, em que se utiliza um modelo dinâmico de populações para sintetizar informações a respeito do tamanho da

população e a taxa de crescimento. No modelo utiliza-se dados biológicos, tais como taxas de sobrevivência e de fecundidade e dados de captura. Uma das variáveis de entrada para o modelo é a taxa de sobrevivência anual para bowheads adultas. Neste contexto, são consideradas adultas, todas as baleias que atingiram a maturidade sexual. O Comitê Científico da Comissão Baleeira Internacional (1999) propôs uma distribuição a priori para a taxa de sobrevivência que faz uso da taxa de mortalidade para uma outra espécie de baleia e de alguns poucos dados de longevidade das bowheads. Entretanto, muitos membros deste comitê acreditavam que seria preferível uma priori baseada em mais dados (Punt e Butterworth,1999).

A taxa de sobrevivência para baleias pode ser estimada através de dados de foto-identificação, aplicando-se técnicas de captura-recaptura. Experimentos de captura-recaptura consistem em se retirar amostras de animais da população em estudo, em k tempos distintos, e marcá-los artificialmente, sendo que cada animal deve possuir marca única. Há várias maneiras de se fazer isso, e a escolha depende da espécie que está sendo estudada. Em experimentos de captura-recaptura envolvendo pássaros, por exemplo, é comum marcar cada indivíduo colocando em sua perna um anel numerado. Entretanto, no caso das baleias bowheads, a marcação artificial é impraticável. Embora sejam predominantemente pretas, as bowheads podem apresentar pigmentação branca na cabeça e na cauda, além de cicatrizes. São essas marcas naturais que tornam possível a re-identificação das baleias, por meio da análise comparativa de fotos aéreas.

As baleias bowhead têm sido fotografadas para estudos de foto-identificação desde a metade da década de 70. A maioria das fotografias apropriadas para identificação individual das baleias, utilizando-se suas marcas naturais, estão guardadas em acervos fotográficos da LGL Limited e do Laboratório Nacional de Mamíferos Marinhos (National Marine Mammal Laboratory - NMML), em Seattle, Washington. Rugh *et al.* (1992) descreveu como as fotografias são tiradas, as coleções organizadas e os indivíduos são identificados.

A primeira tentativa de se estimar a taxa de sobrevivência das bowheads, através de foto-identificação e técnicas de captura-recaptura, foi apresentada ao Comitê

Científico IWC por Whicher *et al.* (1996). Entretanto, suas análises apresentaram alguns problemas, sendo um deles provocado pelo sistema de classificação das fotografias, que não permite a classificação única dos animais. Rugh *et al.* (1998) desenvolveram um novo sistema para classificar as fotos dos animais de acordo com sua qualidade e identificabilidade. Na classificação quanto a qualidade da imagem definiu-se as seguintes categorias:

- 1+: as áreas de identificação são totalmente visíveis; todas as marcas existentes são claramente visíveis; a visibilidade não é comprometida por qualquer fator (como por exemplo, nebulosidade, profundidade da água, quantidade de lama na baleia, iluminação, esguicho da baleia);
- 1-: as áreas de identificação são totalmente visíveis, mas as marcas pequenas podem ser perdidas; o comprometimento da visibilidade é desprezível, e todas as partes significativas das áreas de identificação são visíveis;
- 2+: muitas áreas de identificação são visíveis; as marcas grandes e médias existentes podem ser vistas, mas podem não ser muito claras; a visibilidade é comprometida, e partes pequenas das áreas de identificação (até 5%) podem não ser vistas;
- 2-: as regiões do corpo próximas às áreas de identificação são raramente visíveis; marcas médias e grandes podem ser perdidas; a visibilidade é significativamente comprometida, e partes pequenas não podem ser vistas;
- 3: as áreas de identificação são dificilmente vistas; até grandes marcas existentes podem ser perdidas.

Assim como a qualidade da imagem, a identificabilidade das baleias também foi categorizada. Cada uma das áreas de identificação é classificada como a seguir:

- H+: altamente marcada (facilmente reidentificável); deve ter duas marcas grandes, ou pelo menos uma marca grande com três ou mais marcas no dorso,

e duas ou mais na cabeça ou na cauda; as marcas grandes devem ocupar mais do que 10% da área de identificação;

- H-: bem marcada; deve ter uma marca grande e muito diferente, ou uma marca grande com duas ou mais marcas médias no dorso, ou uma ou mais na cabeça ou na cauda, ou duas marcas grandes (não necessariamente diferentes);
- M+: moderadamente marcada (as marcas devem ser vistas em fotografias de boa qualidade); deve haver uma marca grande (não necessariamente diferente), com uma marca média, ou quatro ou mais marcas médias no dorso e três ou mais na cabeça e na cauda;
- M-: marcada (as marcas podem ser vistas em fotografias de boa qualidade); deve ter uma marca grande e diferente, ou pelo menos três marcas médias no dorso, na cabeça ou na cauda;
- U+: pouco marcada; deve ter uma marca grande, não diferente, ou uma ou duas marcas médias no dorso, ou uma marca média na cabeça ou na cauda;
- U-: não marcada (as marcas disponíveis são inadequadas para reidentificação em fotografias aéreas); não existem marcas grandes ou médias.

Na próxima seção descreveremos os objetivos desta dissertação.

1.1 Objetivos da dissertação

Trabalhando com dados de foto-identificação das baleias bowheads, *Balaena mysticetus*, no período de 1981 a 1998, Zeh *et al.* (2002) utilizaram uma abordagem Bayesiana na estimação de parâmetros relativos aos modelos de Jolly-Seber e Cormack-Jolly-Seber. Os autores consideraram dados de 11 anos de estudos de captura-recaptura e o modelo $\{\phi, p_t\}$, que considera taxas de sobrevivência constantes ao longo do tempo e probabilidades de captura dependentes da ocasião amostral.

Numa tentativa de reduzir o número de parâmetros a estimar, Zeh *et al.* (2002) modelou as probabilidades de captura utilizando um modelo logístico frequentista. As análises foram implementadas com o auxílio do *software* MARK (White and Burnham, 1999).

No modelo em questão, os autores consideraram taxas de sobrevivência constantes, isto é, $\phi = \phi_1 = \dots = \phi_{s-1}$ e probabilidades de captura p_t , $t = 2, \dots, s$, descritas como funções do esforço amostral. Além disso, levou-se em conta, no modelo, heterogeneidades nas probabilidades de captura devidas à extensão das marcas naturais das bowheads, uma vez que baleias maduras podem ser muito mais marcadas (cicatrizes, despigmentação, etc) do que baleias imaturas, e, portanto, possivelmente mais capturáveis.

Mais especificamente, Zeh *et al.* (2002) consideraram o modelo logístico, descrito como a seguir:

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}, \quad (1.1)$$

onde

- p_{tg} é a probabilidade de captura no tempo t para o grupo g , $g = 1$ para baleias imaturas e $g = 2$ para baleias maduras;
- $x_{tg_1} = g - 1$ é uma variável que designa grupo, onde $g = 2$ indica indivíduos maduros, e $g = 1$ indivíduos imaturos;
- x_{tg_2} = dias de esforço gastos fotografando baleias no tempo t e grupo g .

Apesar de parecer promissor, o modelo em questão não foi implementado por Zeh *et al.* (2002), em sua versão Bayesiana. Os autores justificaram que os resultados obtidos através do *software* MARK para o modelo (1.1), foram semelhantes aos obtidos utilizando o modelo $\{\phi, p_t\}$ (implementado em sua versão Bayesiana), apesar deste último não levar em conta uma possível heterogeneidade nas probabilidades de captura causada pela extensão das marcas dos animais. Desta forma, os autores deixaram implícito que, devido a sua complexidade, não valeria a pena implementar

o modelo logístico Bayesiano, uma vez que, possivelmente, os resultados com as duas abordagens seriam semelhantes.

No entanto, apesar de sua complexidade, a implementação Bayesiana do modelo de Cormack empreendida nesta dissertação, considerando as probabilidades de captura modeladas através de (1.1), nos ajudou a esclarecer aspectos importantes no estudo da estimação de parâmetros em modelos de captura-recaptura Bayesianos para dados de foto-identificação e populações abertas em geral. É fato conhecido que os estimadores de máxima verossimilhança dos parâmetros dos modelos de Cormack (1964) e Jolly-Seber (1965) não apresentam boa performance para pequenos tamanhos amostrais, como é o caso de muitos experimentos de captura-recaptura empreendidos na prática, inclusive no caso das baleias bowhead.

Em resumo, nesta dissertação implementamos a versão Bayesiana do modelo de Cormack considerando probabilidades de captura modeladas através de (1.1), bem como os modelos reduzidos obtidos a partir do mesmo. Além disso, procedemos à seleção Bayesiana para a escolha do modelo que melhor descreve os dados.

Esta dissertação está organizada como a seguir:

No Capítulo 2 faremos uma breve revisão dos modelos de captura-recaptura mais disseminados, tanto para populações fechadas (o tamanho populacional não varia ao longo do tempo) quanto para populações abertas (o tamanho populacional é variável ao longo do tempo). Além disso, descreveremos a função de verossimilhança de Cormack considerando a existência de grupos na população e as probabilidades de captura expressas em função das covariáveis.

No Capítulo 3 descrevemos o modelo Bayesiano de Cormack com probabilidades de captura expressas em escala logística bem como os procedimentos utilizados na descrição de funções geradoras de candidatos das taxas de sobrevivência e do vetor de coeficientes, β , de um modelo logístico.

No Capítulo 4 sugerimos um total de 12 modelos Bayesianos, a serem comparados, para estimar taxas de sobrevivência e probabilidades de captura. Esses modelos permitem a existência de grupos de indivíduos na população, através da incorporação de covariáveis, de modo a permitir que taxas de sobrevivência e proba-

bilidades de captura sejam estimadas para cada estrato de indivíduos. A modelagem das probabilidades de captura feita através de um modelo logístico permite flexibilidade adicional ao modelo.

No Capítulo 5 apresentamos cálculos necessários à especificação das funções geradoras de candidatos β (coeficientes do modelo logístico para as probabilidades de captura) de cada um dos 12 modelos. No Capítulo 6 apresentamos detalhes das funções geradoras de candidatos das taxas de sobrevivência de cada um dos doze modelos. No Capítulo 7 apresentamos uma aplicação a dados reais das baleias bowhead, seleção de modelos, diagnóstico de convergência do processo Bayesiano de simulação, análise de sensibilidade à escolha de *prioris* e conclusões.

Capítulo 2

Métodos de captura-recaptura

2.1 Populações fechadas - Duas ocasiões amostrais

Segundo Pollock *et al.* (1990), experimentos de captura-recaptura foram utilizados, pela primeira vez, por Laplace em 1786, na estimação do tamanho da população da França. Posteriormente, esse método foi aplicado a dados ecológicos por Petersen (1896).

Inicialmente, consideremos a versão mais simples dos experimentos de captura-recaptura, em que temos apenas 2 ocasiões amostrais. Considere uma população de tamanho N onde, na ocasião amostral 1, n_1 animais são capturados, marcados e devolvidos à população. Posteriormente, n_2 indivíduos são capturados, dos quais m são marcados. Supondo que a proporção de animais marcados na população e a de animais marcados na segunda amostra são iguais, temos:

$$\frac{m}{n_2} = \frac{n_1}{N}. \quad (2.1)$$

Desta forma, é possível obter um estimador para o tamanho populacional dado por:

$$\hat{N} = \frac{n_1 n_2}{m} \quad (2.2)$$

No entanto, a expressão (2.2) representa um estimador viciado de N . Segundo Seber (1970,1982) uma versão modificada, \hat{N}_c , com um vício menor, foi originalmente proposta por Chapman (1951) e é dada a seguir:

Tabela 2.1: História hipotética de captura de 4 animais em três ocasiões amostrais.

Animal	Ocasião amostral 1	Ocasião amostral 2	Ocasião amostral 3
1	1	1	0
2	1	0	0
3	0	1	0
4	0	0	1

$$\hat{N}_c = \frac{(n_1 + 1)(n_2 + 1)}{m + 1} - 1. \quad (2.3)$$

Um estimador não viciado aproximado para variância de \hat{N}_c é dado por:

$$\widehat{Var}(\hat{N}_c) = \frac{(n_1 + 1)(n_2 + 1)(n_1 - m)(n_2 - m)}{(m + 1)^2(m + 2)}. \quad (2.4)$$

A seguir, apresentaremos uma descrição sumária de métodos de captura-recaptura para populações fechadas envolvendo mais de duas ocasiões amostrais.

2.2 Populações fechadas - Mais de duas ocasiões amostrais

No caso mais geral, com $t > 2$ ocasiões amostrais, podemos organizar a história de captura de cada animal em uma tabela em que, cada célula, recebe ou 1 ou 0. Estes números indicam captura e não-captura, respectivamente. Vejamos um exemplo para três ocasiões amostrais.

Na Tabela 2.1, o animal 1 foi capturado na primeira amostra, recapturado na segunda e não foi capturado na ocasião amostral 3. O animal 2 foi capturado na primeira ocasião mas não foi recapturado em nenhuma outra. O animal 3 foi capturado, pela primeira vez, na segunda ocasião amostral e não foi recapturado. O animal 4 foi capturado apenas na terceira ocasião amostral. No total, há $2^3 - 1$ histórias de captura possíveis: $\{(1, 0, 0), (0, 1, 0), (0, 0, 1), (1, 1, 0), (1, 0, 1), (0, 0, 1), (1, 1, 1)\}$.

As probabilidades de captura podem ser modeladas de diversas maneiras, gerando

classes de modelos que passaremos a relatar. Em todos os casos que descreveremos os seguintes pressupostos devem ser observados:

1. A população é fechada, ou seja, não ocorrem nascimentos, mortes ou migrações o que faz com que o tamanho populacional seja constante durante o tempo de realização dos estudos;
2. As capturas são independentes nos diversos tempos;
3. Os animais não perdem suas marcas entre as ocasiões amostrais

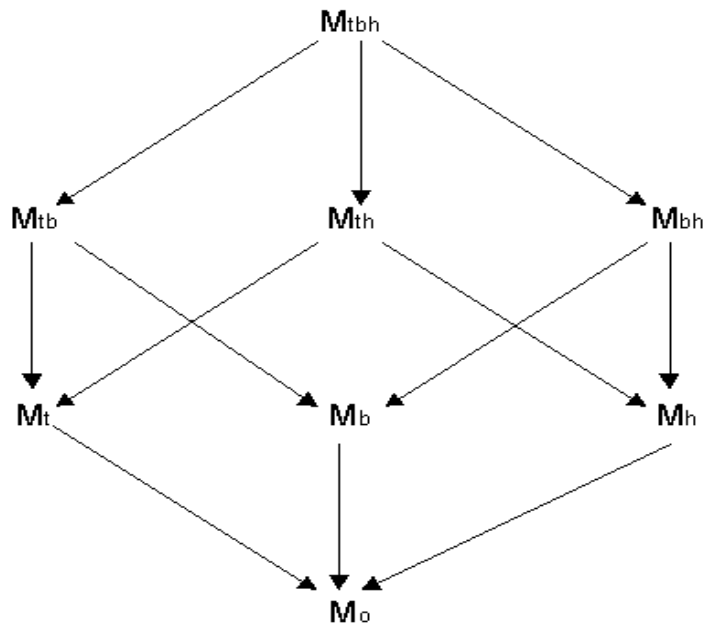


Figura 2.1: Relações entre modelos de captura-recaptura.

- **Modelo M_0 - Modelo capturabilidade uniforme.**

Neste modelo assume-se que todo animal da população possui a mesma probabilidade de captura p para cada ocasião amostral. Sob este modelo, o estimador de máxima verossimilhança (EMV) só pode ser encontrado numericamente, ou seja, não há uma forma explícita para o EMV.

- **Modelo M_h - Modelo de heterogeneidade.**

Este modelo admite heterogeneidade nas probabilidades de captura dos

animais, ou seja, assume que cada um dos indivíduos possui sua própria probabilidade de captura p_j , $j = 1, \dots, N$, que se mantém constante em todas as ocasiões amostrais. O vetor de frequências de captura, (f_1, \dots, f_k) , que é formado pelo número de animais capturados $1, 2, \dots, k$ vezes, contém toda a informação necessária para se estimar N . Sob este modelo, um estimador para N seria o número de animais distintos que foram capturados durante todo o experimento. Porém, este não é um bom estimador, visto que apresenta grande vício quando as probabilidades de captura são baixas ou o número de ocasiões amostrais é pequeno.

- **Modelo M_b - Modelo resposta à captura.**

Neste modelo considera-se a resposta do animal à captura, mas não a heterogeneidade ou variação das probabilidades de captura no tempo. O modelo é válido sob as seguintes suposições:

- (1) Todo animal não marcado na população tem a mesma probabilidade de captura p para todas as ocasiões amostrais.
- (2) Uma vez que foi capturado, todo animal marcado na população tem a mesma probabilidade de recaptura c para todas as ocasiões amostrais.

Este modelo é útil quando a probabilidade de recaptura de um animal muda após a primeira captura. Quando $p < c$, diz-se que o animal tornou-se “trap happy”, isto é, mais propenso a ser recapturado, possivelmente pela chance de obter alimento. Quando $p > c$, diz-se que o animal tornou-se “trap shy”, sendo menos propenso a ser capturado, possivelmente devido a algum trauma.

- **Modelo M_{bh} - Modelo de heterogeneidade e resposta à captura.**

O modelo M_{bh} leva em conta tanto a heterogeneidade quanto a resposta do animal à captura, e assume que cada animal tem seu próprio e único par de probabilidades de captura e recaptura $(p_j, c_j; j = 1, \dots, N)$, com p_j e c_j referindo-se às probabilidades de captura de animais não marcados e marcados respectivamente.

Tabela 2.2: Modelos de captura-recaptura para populações fechadas

Modelo	Fonte de variação na probabilidade de captura			
	Heterogeneidade	Resposta à captura	Tempo	Estimador disponível
M_0				x
M_h	x			x
M_b		x		x
M_{bh}	x	x		x
M_t			x	x
M_{th}	x		x	
M_{tb}		x	x	
M_{tbh}	x	x	x	

- **Modelo M_t - O Modelo de Schnabel** (Variação temporal das probabilidades de captura).

Este modelo assume que todo animal na população tem a mesma probabilidade de captura a cada ocasião amostral, $p_i; i = 1, 2, \dots, k$. Este é um modelo de captura-recaptura para populações fechadas muito utilizado na prática, e é válido sob as seguintes suposições adicionais:

1. No tempo j as probabilidades de captura, p_j , são iguais para todos os indivíduos;
2. As probabilidades de captura sofrem apenas efeito temporal.

- **Outros modelos dependentes do tempo.**

Um dos pressupostos dos modelos M_h, M_b e M_{bh} é o de que as probabilidades de captura não variam no tempo. As generalizações desses modelos, incorporando variação temporal, são, respectivamente, M_{th}, M_{tb} e M_{tbh} .

A Tabela 2.2 sumariza as informações sobre os modelos descritos. O símbolo “x” representa dada característica associada ao modelo. Note que os modelos M_{th}, M_{tb} e M_{tbh} , embora mais realísticos, não têm um estimador disponível, ainda que de forma não fechada. Na verdade, tais modelos incorporam mais parâmetros a estimar do que é possível com os dados que estariam disponíveis, fazendo-se necessário o uso de simplificações adicionais para que se possa estimar os parâmetros.

No caso específico do modelo M_t , a verossimilhança associada é dada pela seguinte expressão:

$$P(\{a_w\}) = \frac{N!}{(N-r)! \prod_w a_w!} \prod_{j=1}^t p_j^{n_j} (1-p_j)^{N-n_j}, \quad (2.5)$$

onde

- r =número de animais distintos capturados no experimento;
- a_w =número de animais marcados com história de captura w ;
- p_j =probabilidade de captura no tempo j ;
- t =número de ocasiões amostrais;
- n_j =número de indivíduos capturados no tempo j .

Na próxima seção descreveremos o modelo de Cormack (Cormack,1964) para populações abertas. Tal modelo é amplamente utilizado na prática, e é de interesse central nesta dissertação.

2.3 Populações abertas - O Modelo de Cormack

No caso de populações abertas, isto é, populações em que são permitidas as ocorrências de mortes e nascimentos, bem como imigrações e emigrações, Cormack (1964) sugeriu um modelo que permite estimar taxas de sobrevivência e probabilidades de captura. Os pressupostos do modelo são dados a seguir:

1. Todo animal presente na população no tempo t ($t = 1, \dots, k$) tem a mesma probabilidade de captura p_t .
2. Todo animal marcado, presente na população imediatamente após a amostra i , tem a mesma probabilidade de sobrevivência (ϕ_i) até a ocasião da amostra $(i+1)$, ($i = 1, \dots, k$).
3. As marcas não são perdidas ou despercebidas.

4. Todas as amostras são instantâneas e cada liberação dos indivíduos é feita imediatamente após a amostragem.
5. Todo animal capturado na amostra i tem probabilidade ν_i de ser devolvido à população e, portanto, $(1 - \nu_i)$ é a probabilidade de morte acidental.

2.3.1 Notação

Considerando as notações de Cormack (1964), Seber (1982) e Zeh *et al.* (2002), sejam:

- u_t =número de animais vistos pela primeira vez no tempo t ;
- a_w =número de animais marcados com história de captura w ;
- s =número de ocasiões amostrais;
- ϕ_t =probabilidade de que um animal capturado no tempo t sobreviva ao tempo $t + 1$;
- p_t =probabilidade de recaptura no tempo t ;
- χ_t =probabilidade de que um animal, capturado no tempo t , não seja capturado em ocasiões futuras, ou seja

$$\chi_t = 1 - \sum_{k=t}^{s-1} \left\{ \left(\frac{p_{k+1}}{1 - p_{k+1}} \right) \prod_{i=t}^k \phi_i (1 - p_{k+1}) \right\}. \quad (2.6)$$

- c_t =número de animais capturados no tempo t , que não foram capturados em ocasiões futuras;
- z_t =número de animais capturados antes e depois de t mas não em t ;
- m_t =número de animais recapturados no tempo t .

2.3.2 O Modelo de Cormack

Considerando os pressupostos acima, Cormack (1964) desenvolveu uma verossimilhança condicional ao vetor u_t . O modelo a seguir também é conhecido como modelo $\{\phi_t, p_t\}$, uma vez que tanto as taxas de sobrevivência, ϕ_t , quanto as probabilidades de captura, p_t , podem variar com o tempo.

$$f(\{a_w\}|\{u_t\}) = \frac{\prod_{t=1}^s u_t!}{\prod_w a_w!} \prod_{t=1}^{s-1} \chi_t^{c_t} [\phi_t(1 - p_{t+1})]^{z_{t+1}} (\phi_t p_{t+1})^{m_{t+1}}. \quad (2.7)$$

Cormack (1989) descreveu os seguintes modelos, como simplificações de seu modelo geral $\{\phi_t, p_t\}$, descrito em Cormack (1964):

- $\{\phi, p\}$ - probabilidades de sobrevivência e de captura constantes;
- $\{\phi_t, p\}$ - probabilidades de sobrevivência variáveis e probabilidades de captura constantes;
- $\{\phi, p_t\}$ - probabilidades de sobrevivência constantes e probabilidades de captura variáveis.

Embora flexível, o modelo de Cormack não permite a estimação de abundância. Na seção seguinte apresentaremos outro modelo para populações abertas, o famoso modelo de Jolly-Seber, que permite a estimação de abundância.

2.4 O modelo de Jolly-Seber

Jolly (1965) e Seber (1965) desenvolveram, independentemente, a seguinte verossimilhança conjunta, que permite, a cada ocasião amostral, a estimação de probabilidades de captura, taxas de sobrevivência e abundância:

$$f(\{a_w\}, \{u_t\}) = f(\{a_w\}|\{u_t\})f(\{u_t\}).$$

Tal modelo é baseado na suposição de que as probabilidades de captura e recaptura, em uma dada ocasião amostral, são as mesmas. Assumindo amostragem Binomial, temos:

$$f(\{u_t\}) = \prod_{u_t=1}^{u_s} \left\{ \binom{U_t}{u_t} p_t^{u_t} (1-p_t)^{U_t-u_t} \right\}, \quad (2.8)$$

onde U_t = número de animais marcados que permanecem vivos imediatamente antes da ocasião amostral t , e que ainda não foram capturados.

Com base no modelo de Cormack (1964), é possível descrever um modelo em que a população é estratificada em grupos independentes de indivíduos. Tal modelo é apresentado na próxima seção, em uma versão em que se considera as probabilidades de captura expressas através de um modelo logístico.

2.5 A função de verossimilhança de Cormack com probabilidades de captura expressas em escala logística

Considere o modelo de Cormack (1964) e a presença de G grupos distintos e independentes de indivíduos na população. A função de verossimilhança associada é dada por:

$$L(\phi_{tg}, p_{tg}) \propto \prod_{g=1}^G \left[\prod_{t=1}^{s-1} \chi_{tg}^{c_{tg}} \prod_{t=1}^{s-1} [\phi_{tg}(1-p_{t+1,g})]^{z_{t+1,g}} (\phi_{tg} p_{t+1,g})^{m_{t+1,g}} \right]. \quad (2.9)$$

Observe que L pode ser fatorada nos seguintes elementos:

- (i) $V_{\phi p} = \prod_{g=1}^G \prod_{t=1}^{s-1} \chi_{tg}^{c_{tg}}$.
- (ii) $L_{\phi} = \prod_{g=1}^G \prod_{t=1}^{s-1} \phi_{tg}^{z_{t+1,g} + m_{t+1,g}}$.
- (iii) $L_p = \prod_{g=1}^G \prod_{t=1}^{s-1} (1-p_{t+1,g})^{z_{t+1,g}} (p_{t+1,g})^{m_{t+1,g}}$.

Logo,

$$L \propto V_{\phi p} \times L_{\phi} \times L_p.$$

Consideremos que, no grupo g , k covariáveis são utilizadas para descrever a probabilidade de captura no tempo t , isto é,

$$p_{tg} = \frac{\exp(\beta' X_{t+1,g})}{1 + \exp(\beta' X_{t+1,g})},$$

onde $\beta' = (\beta_0, \dots, \beta_k)$ e $X_{t+1,g} = (1, x_{t+1,g_1}, \dots, x_{t+1,g_k})'$. Desta forma, L (vide expressão 2.9) fica parametrizada em função de β . Portanto, considere a verossimilhança parcial L_{β} dada a seguir.

$$\begin{aligned} L_{\beta} &= \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta' X_{t+1,g})} \right]^{z_{t+1,g} + m_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} [\exp(\beta' X_{t+1,g})]^{m_{t+1,g}}. \end{aligned} \quad (2.10)$$

Logo,

$$L \propto V_{\phi\beta} \times L_{\beta} \times L_{\phi},$$

ou seja,

$$\begin{aligned} L &\propto \prod_{g=1}^G \prod_{t=1}^{s-1} [\chi_{tg}^{c_{tg}}] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta' X_{t+1,g})} \right]^{z_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta' X_{t+1,g})}{1 + \exp(\beta' X_{t+1,g})} \right]^{m_{t+1,g}} \times \prod_{g=1}^G \prod_{t=1}^{s-1} \phi_{tg}^{(z_{t+1,g} + m_{t+1,g})}. \end{aligned} \quad (2.11)$$

No próximo capítulo descreveremos a metodologia Bayesiana que adotamos, nesta dissertação, na estimação dos parâmetros de interesse, no modelo (2.11).

Capítulo 3

Descrição do Modelo Bayesiano de Cormack

Tanto para o modelo de Cormack quanto para o de Jolly-Seber os estimadores de máxima verossimilhança dos parâmetros de interesse têm boa performance para grandes tamanhos amostrais, que propiciam a observação de um número razoável de recapturas (parte do conjunto de estatísticas suficientes para os parâmetros). Os respectivos estimadores da variância dos estimadores dos parâmetros são expressões assintóticas, de modo que, para pequenas populações e/ou pequenos tamanhos amostrais, os intervalos de confiança podem ser inválidos, por exemplo, fornecendo limites inferiores menores do que zero, ou superiores maiores do que 1, para um parâmetro que designa uma probabilidade.

A vantagem dos estimadores Bayesianos sobre os frequentistas é que eles são mais úteis para populações finitas e amostras pequenas.

Considere o modelo de Cormack quando se leva em conta a presença de G grupos de indivíduos na população e probabilidades de captura expressas em função de covariáveis, vide expressão (2.11). Seja $\theta = (\phi, \beta)$ o vetor de parâmetros de interesse, onde $\beta = (\beta_0, \dots, \beta_k)$ e $\phi = (\phi_{11}, \dots, \phi_{t1}, \dots, \phi_{1G}, \dots, \phi_{tG})$. Denota-se por D os dados observados. Considerando-se que ϕ e β sejam independentes *a priori*, a distribuição condicional completa *a posteriori* de θ , representando nossa *distribuição*

alvo, é dada por:

$$\pi(\theta|D) = \pi(\phi, \beta|D) \propto L(\phi, \beta, D)\pi(\phi, \beta) = L(\phi, \beta, D)\pi(\beta)\pi(\phi). \quad (3.1)$$

Amostras geradas através da distribuição alvo, $\pi(\phi, \beta|D)$, podem ser obtidas através da geração de valores a partir de distribuições condicionais completas *a posteriori* de cada parâmetro (vide Gilks *et al.*, 1996). Desta forma, nas seções seguintes, passaremos a descrever, em linhas gerais, as distribuições condicionais completas *a posteriori* de cada um dos parâmetros em estudo.

3.1 A distribuição condicional completa *a posteriori* de ϕ_{jg}

Seja

$$\pi(\phi_{jg}|D) \propto V_{\phi\beta} \times L_{\phi} \times \pi(\phi_{jg}).$$

Como em Poole (2002), considere uma priori $Beta(\alpha, \gamma)$ para ϕ_{jg} , $j = 1, \dots, s-1$ e $g = 1, \dots, G$. A distribuição condicional completa *a posteriori* de ϕ_{jg} é dada por:

$$\begin{aligned} & \pi(\phi_{jg}|\phi_{1g}, \dots, \phi_{j-1g}, \dots, \phi_{s-1g}, \dots, \phi_{1G}, \dots, \phi_{tG}; \beta, D) \equiv \pi(\phi_{jg}|\phi_{-jg}; \beta, D) \\ & \propto \left[\prod_{i=1}^j (\chi_{ig})^{c_{ig}} \right] \phi_{jg}^{(z_{j+1,g} + m_{j+1,g})} \times \phi_{jg}^{\alpha-1} (1 - \phi_{jg})^{\gamma-1} \\ & = \left[\prod_{i=1}^j (\chi_{ig})^{c_{ig}} \right] \phi_{jg}^{(z_{j+1,g} + m_{j+1,g} + \alpha - 1)} \times (1 - \phi_{jg})^{\gamma-1}. \end{aligned}$$

Como $\pi(\phi_{jg}|\phi_{-jg})$ não apresenta uma forma conhecida, Poole (2002) considerou o uso do algoritmo de Metropolis-Hastings (vide Anexo B) para obter uma aproximação de $\pi(\phi_{jg}|\phi_{-jg})$. Deste modo, o autor considerou a função geradora de candidatos ϕ_{jg}^* de ϕ_{jg} dada por:

$$q(\phi_{jg}^*|\phi_{jg}) = \begin{cases} \frac{1}{2k\phi_{jg}^*(1-\phi_{jg}^*)}; & -k < \log\left(\frac{\phi_{jg}^*}{1-\phi_{jg}^*}\right) - \log\left(\frac{\phi_{jg}}{1-\phi_{jg}}\right) < k. \\ 0 & \text{c.c.} \end{cases} \quad (3.2)$$

Desta forma, a probabilidade de aceitação é

$$\alpha(\phi_{jg}, \phi_{jg}^*) = \min \left\{ 1, \frac{q(\phi_{jg}|\phi_{jg}^*)}{q(\phi_{jg}^*|\phi_{jg})} \times \frac{\pi(\phi_{jg}|\phi_{jg}^*)}{\pi(\phi_{jg}^*|\phi_{jg})} \right\}.$$

Por simplicidade computacional considere

$$\begin{aligned} B_{jg} &= \frac{q(\phi_{jg}|\phi_{jg}^*)}{q(\phi_{jg}^*|\phi_{jg})} \times \frac{\pi(\phi_{jg}|\phi_{jg}^*)}{\pi(\phi_{jg}^*|\phi_{jg})} \\ &= \left[\frac{\phi_{jg}^*}{\phi_{jg}} \right]^{(z_{j+1,g} + m_{j+1,g} + \alpha)} \left[\frac{1 - \phi_{jg}^*}{1 - \phi_{jg}} \right]^\gamma \frac{\prod_{i=1}^j (\chi_{ig}^*)^{c_{ig}}}{\prod_{i=1}^j (\chi_{ig})^{c_{ig}}}, \end{aligned}$$

e a probabilidade de aceitação dada por

$$\alpha(\phi_{jg}, \phi_{jg}^*) = \min \{1, \exp(\log B_{jg})\}.$$

3.2 A distribuição condicional completa *a posteriori* de β

Considere a expressão (2.11), que representa a verossimilhança do modelo de Cormack com as probabilidades de captura expressas em escala logística. Além disso, considere a verossimilhança reduzida

$$L_{\phi\beta} = V_{\phi\beta} \times L_{\beta}$$

que relaciona os termos em (2.11) que envolvem o parâmetro β .

Desta forma, $L_{\phi\beta}$, pode ser expressa por:

$$\begin{aligned} L_{\phi\beta} &\propto \prod_{g=1}^G \prod_{t=1}^{s-1} [\chi_{tg}^{c_{tg}}] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta' X_{t+1,g})} \right]^{z_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta' X_{t+1,g})}{1 + \exp(\beta' X_{t+1,g})} \right]^{m_{t+1,g}}. \end{aligned}$$

Portanto,

$$\begin{aligned} L_{\phi\beta} &\propto \prod_{g=1}^G \prod_{t=1}^{s-1} [\chi_{tg}^{c_{tg}}] \prod_{g=1}^G \prod_{t=1}^{s-1} [\exp(\beta' X_{t+1,g})]^{m_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} [1 + \exp(\beta' X_{t+1,g})]^{-(z_{t+1,g} + m_{t+1,g})}. \end{aligned} \quad (3.3)$$

Seja a distribuição condicional completa *a posteriori* de β dada por

$$\pi(\beta|\phi, D) \propto V_{\phi\beta} \times L_{\beta} \times \pi(\beta) = L_{\phi\beta} \times \pi(\beta), \quad (3.4)$$

e considere a distribuição *a priori* de β , $\pi(\beta)$, dada por

$$\beta \sim N_p(\mathbf{0}, \sigma^2 \mathbf{I}) \propto \exp \left[-\frac{\mathbf{1}}{2\sigma^2} \beta^t \beta \right].$$

onde $\mathbf{0}$ indica um vetor nulo e \mathbf{I} uma matriz identidade.

A utilização de prioris normais para os coeficientes de modelos lineares ou lineares generalizados é bem conhecida na literatura, e é discutida em detalhes nos capítulos 15 e 16 de Gelman *et al.* (2003).

Neste estudo σ^2 é considerado fixo e grande, de modo a se obter uma priori vaga. A expressão (3.4) não possui forma fechada, de modo que faz-se necessário obter uma aproximação de $\pi(\beta|\phi, D)$ através do algoritmo de Metropolis-Hastings.

3.3 Descrição da função geradora de candidatos de β

Considerando a distribuição condicional completa *a posteriori* de β , descrita na expressão (3.4), e a necessidade do estabelecimento de uma função geradora de candidatos de β através do algoritmo de Metropolis-Hastings, nossa função geradora de candidatos de β consiste em tomar uma densidade normal, q , tal que

$$\beta^* | \beta^s \sim N(\hat{\beta}, \hat{H}^{-1}); \quad (3.5)$$

onde $\hat{\beta}$ é o valor modal de $\pi(\beta|\phi, D)$, obtido ao maximizarmos a distribuição condicional completa *a posteriori* de β . Na expressão acima, β^s é o valor atual de β e β^* é um valor candidato de β , enquanto,

$$\hat{H} = -\frac{\partial^2}{\partial\beta\partial\beta'} \log [\pi(\beta | \phi, D)] |_{\beta=\hat{\beta}}, \quad (3.6)$$

representa a informação de Fisher observada e relativa *a posteriori* de β , avaliada no ponto observado $\hat{\beta}$. Tal procedimento é discutido no capítulo 12 de Gelman *et al.* (2003).

Desta forma, a função geradora de candidatos, q , de β , consiste em uma distribuição Normal multivariada com vetor de médias $\hat{\beta}$ e matriz de covariâncias expressa pelo inverso da informação observada de Fisher associada à densidade condicional completa *a posteriori* de β . Mais detalhes na obtenção de \hat{H} , para cada modelo, serão dados no Capítulo 5.

Em nosso contexto, utilizamos tal função geradora de candidatos, normal multivariada, por ela nos fornecer uma aproximação para a distribuição condicional completa *a posteriori* de β , que é de fácil amostragem. Segundo Gelman *et al.* (2003) capítulo 11, no algoritmo de Metropolis-Hastings a função geradora de candidatos ideal iguala a distribuição a partir da qual se deseja amostrar, devendo ser de fácil amostragem. Além disso, tal função geradora de candidatos deve ser capaz de fazer uma boa varredura no espaço paramétrico e os candidatos gerados não serem rejeitados muito frequentemente.

Nas próximas seções faremos um estudo mais detalhado da especificação da distribuição geradora de candidatos de β e da amostragem a partir de uma distribuição normal multivariada.

3.3.1 Especificação da distribuição geradora de candidatos de β

Suponha que a distribuição *a priori* dos parâmetros desconhecidos, β , especifique que cada coordenada do vetor siga distribuição $N(0, \sigma^2)$, com os β_i 's independentes. Desta forma, a densidade *a priori* conjunta e a densidade *a posteriori* são dadas, respectivamente, por:

$$\pi(\beta) \propto \exp\left(-\frac{\beta' \beta}{2\sigma^2}\right)$$

e

$$\pi(\beta | D) \propto L(\beta)\pi(\beta).$$

A densidade *a posteriori*, $\pi(\beta | D)$, não pode ser amostrada, diretamente, de forma trivial. Neste caso, utilizaremos o algoritmo de Metropolis-Hastings, com cadeia independente (vide apêndices A e B), para gerar os valores candidatos de β . Deste modo, seja a *função geradora de candidatos* de β dada por:

$$q(\beta^{(s)}, \beta) = g(\beta),$$

descrevendo uma densidade que

- (1) representa uma boa aproximação para $\pi(\beta | D)$, e
- (2) torna possível a geração de valores-candidatos $\beta^{(s)}$ a partir de $q(\beta^{(s)}, \beta)$.

Considere

$$g(\beta) \propto \exp\left(-\frac{1}{2}(\beta - \hat{\beta})'H(\hat{\beta})(\beta - \hat{\beta})\right),$$

isto é, a função geradora de candidatos $\beta^{(s)}$ segue o modelo $N_{p+1}(\hat{\beta}, H^{-1}(\hat{\beta}))$, onde $\hat{\beta}$ é o vetor modal de $\pi(\beta | D)$, e

$$H(\hat{\beta}) = -\frac{\partial^2}{\partial\beta\partial\beta} \log(\pi(\beta | D)) |_{\beta=\hat{\beta}}.$$

Portanto, a função geradora de candidatos de β , $g(\beta)$, é descrita através de uma distribuição normal multivariada cujo vetor de médias é dado pela moda da distribuição condicional completa *a posteriori*, $\hat{\beta}$, e a matriz de covariâncias é expressa através da inversa da matriz de informação observada da distribuição condicional completa *a posteriori*, $H^{-1}(\beta)$, avaliada no ponto $\beta = \hat{\beta}$, representando uma estimativa de β .

Em suma, em nosso procedimento Metropolis-Hastings para gerar vetores β a partir da distribuição condicional completa *a posteriori* $\pi(\beta | \phi, D)$, utilizaremos:

- (a) Função geradora de candidatos β^* de β :

$$\beta^* \sim N(\hat{\beta}, \hat{H}^{-1}).$$

- (b) Priori de β :

$$\beta^* \sim N(0, \sigma^2 I).$$

(c) Probabilidade de aceitação de β^* :

$$\alpha(\beta, \beta^*) = \min\{1, B\},$$

onde

$$B = \frac{\pi(\beta \mid \phi, D)}{\pi(\beta^* \mid \phi, D)} \times \frac{g(\beta^*)}{g(\beta)}.$$

Na próxima seção descreveremos o processo de geração de dados a partir de uma distribuição normal multivariada.

3.3.2 Geração de dados a partir de uma Normal Multivariada

Considere a seguinte propriedade da distribuição normal:

Transformação Linear de uma variável aleatória normal multivariada

Sejam $x \sim N_p(\mu, \Sigma)$, A uma matriz $r \times p$ e D um vetor r -dimensional. Segundo Mardia (1979), uma combinação linear de x segue a propriedade,

$$y = Ax + D \sim N_r(A\mu + D, A\Sigma A^t). \quad (3.7)$$

A propriedade acima será utilizada adiante na descrição da geração dos dados a partir de uma distribuição $N_{p+1}(\hat{\beta}, H^{-1}(\hat{\beta}))$.

Considere o vetor aleatório $z^{(s)} \in \mathbf{R}^{p+1}$ formado por $p + 1$ coordenadas, isto é, $z^{(s)} = (z_1^{(s)}, \dots, z_{p+1}^{(s)})$, tais que os $z_j^{(s)}$'s sejam independentes e identicamente distribuídos com $z_k^{(s)} \sim N(0, 1)$. Desta forma,

$$z^{(s)} \sim N_{p+1}(0, I).$$

Na geração do vetor candidato $\beta^{(s)}$, considere

$$\beta^{(s)} = Lz^{(s)} + \hat{\beta}_{obs}, \quad (3.8)$$

onde L é tal que $LL^t = H^{-1}(\hat{\beta})$, representando uma matriz triangular superior obtida por meio da decomposição de Cholesky de H^{-1} , e $\hat{\beta}_{obs}$ representa a moda

observada da distribuição condicional completa *a posteriori* $\pi(\beta | D)$. Na geração de $\beta^{(s)}$, $\hat{\beta}$ é considerado um vetor fixo, isto é, não aleatório, de modo que

$$E(\beta^{(s)}) = E(Lz^{(s)}) + E(\hat{\beta}_{obs}) = LE(z^{(s)}) + E(\hat{\beta}_{obs}) = \hat{\beta}_{obs}, \text{ e}$$

$$Cov(\beta^{(s)}) = Var(Lz^{(s)}) = LIL^t = H^{-1}(\hat{\beta}_{obs}).$$

Utilizando a propriedade de combinação linear de distribuições normais multivariadas descrita acima, tem-se que

$$\beta^{(s)} = Lz^{(s)} + \hat{\beta}_{obs} \sim N_{p+1}(\hat{\beta}_{obs}, H^{-1}(\hat{\beta}_{obs})). \quad (3.9)$$

Portanto, na geração de valores candidatos $\beta^{(s)}$, basta gerarmos um vetor $z^{(s)}$ como descrito acima, e aplicarmos (3.8).

3.4 Um exemplo utilizando o modelo logístico

Para ilustrar as técnicas discutidas na seção (3.3), será apresentado nesta seção um exemplo utilizando-se uma análise logística.

3.4.1 O modelo logístico

Métodos de regressão são essenciais em análises de dados em que se pretende descrever a relação entre a variável resposta e as variáveis explanatórias. No caso em que a variável resposta é discreta, assumindo dois ou mais valores possíveis, o modelo de regressão logístico é amplamente utilizado. Nas subseções que se seguem descreveremos algumas propriedades deste modelo.

3.4.2 Verossimilhança associada ao modelo logístico

Considere uma amostra aleatória de n indivíduos de uma dada população em que são registrados:

- y_i : resposta binária do indivíduo i ;

- x_i : vetor $(p + 1)$ -dimensional de covariáveis associadas ao indivíduo i , $i = 1, \dots, n$.

A verossimilhança associada ao modelo logístico é dada por:

$$L(\beta) = \prod_{i=1}^n [\pi(x_i)]^{y_i} \cdot [1 - \pi(x_i)]^{1-y_i},$$

A probabilidade de “sucesso”, isto é, de $y = 1$, quando se observa um vetor de covariáveis x é dada por:

$$\pi(x) = \frac{e^{x'\beta}}{1 + e^{x'\beta}},$$

onde β é um vetor de parâmetros de dimensão $(p + 1)$, isto é, $\beta' = (\beta_0, \beta_1, \dots, \beta_p)$.

Logo, a verossimilhança relativa ao modelo logístico é dada por:

$$\begin{aligned} L(\beta) &= \prod_{i=1}^n \frac{e^{y_i x_i' \beta}}{(1 + e^{x_i \beta})^{y_i}} \cdot \left[\frac{1}{(1 + e^{x_i \beta})} \right]^{1-y_i} \\ &= \prod_{i=1}^n e^{y_i x_i' \beta} \left(\frac{1}{1 + e^{x_i' \beta}} \right) \\ &= \prod_{i=1}^n \frac{\exp(y_i x_i' \beta)}{1 + \exp(x_i' \beta)}. \end{aligned}$$

Seja $l(\beta) = \log L(\beta)$, a log-verossimilhança do modelo em questão. Então,

$$l(\beta) = \sum_{i=1}^n y_i x_i' \beta - \sum_{i=1}^n \log(1 + e^{x_i' \beta}). \quad (3.10)$$

Os estimadores de máxima verossimilhança dos parâmetros do modelo são obtidos efetuando-se a diferenciação de $l(\beta)$ com respeito a cada uma das $(p + 1)$ coordenadas de β , e, então resolvendo-se o sistema de equações de verossimilhança.

Portanto, como

$$\begin{aligned} \frac{\partial l(\beta)}{\partial \beta_0} &= \sum_{i=1}^n y_i - \sum_{i=1}^n \frac{1}{1 + \exp(x_i' \beta)} \cdot \exp(x_i' \beta) \\ &= \sum_{i=1}^n y_i - \sum_{i=1}^n \pi(x_i), \end{aligned}$$

e

$$\begin{aligned}\frac{\partial l(\beta)}{\partial \beta_j} &= \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n \frac{1}{1 + \exp(x_i' \beta)} \cdot \exp(x_i' \beta) \cdot x_{ij} \\ &= \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n \pi(x_i) \cdot x_{ij},\end{aligned}$$

onde $j = 1, \dots, p$. O sistema de equações de verossimilhança é dado por:

$$\begin{cases} \sum_{i=1}^n (y_i - \pi(x_i)) = 0 \\ \sum_{i=1}^n x_{ij} (y_i - \pi(x_i)) = 0; \quad j = 1, 2, \dots, p. \end{cases} \quad (3.11)$$

Entretanto, essas equações não são lineares em β , e requerem o uso de métodos numéricos para a sua solução.

3.4.3 Cálculo da Matriz de Informação de Fisher

Seja $\hat{\beta}$ o vetor solução do sistema de equações dado pela expressão (3.11). A matriz de variâncias e covariâncias dos coeficientes estimados, $\Sigma(\beta)$, é obtida através da inversa da matriz de informação de Fisher, isto é,

$$\Sigma(\beta) = I^{-1}(\beta).$$

Uma formulação da matriz de informação que é útil (vide Hosmer e Lemeshow, 1989), tanto quando se discute o ajuste do modelo, quanto na avaliação do mesmo, é dada por

$$\hat{I}(\hat{\beta}) = X' V X,$$

onde, X é a matriz de covariáveis independentes dada por:

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{np} \end{bmatrix} = \begin{bmatrix} 1 & x^{(1)} \\ 1 & x^{(2)} \\ \vdots & \vdots \\ 1 & x^{(n)} \end{bmatrix}$$

e V é uma matriz diagonal $n \times n$ com elementos expressos por $\hat{\pi}_i(1 - \hat{\pi}_i)$, onde $\hat{\pi}_i$ denota $\pi(x_i)$ avaliada em $\hat{\beta}$. Desta forma,

$$V = \begin{bmatrix} \hat{\pi}_1(1 - \hat{\pi}_1) & 0 & \dots & 0 \\ 0 & \hat{\pi}_2(1 - \hat{\pi}_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\pi}_n(1 - \hat{\pi}_n) \end{bmatrix}.$$

A inversa da matriz de informação, I^{-1} , pode ser expressa por meio da transformação de Cholesky, isto é,

$$I^{-1}(\hat{\beta}) = L^{-1}(L^{-1})^t = \Sigma(\beta). \quad (3.12)$$

onde L é uma matriz simétrica $(p + 1) \times (p + 1)$ de valores reais.

A seguir apresentaremos a formulação de uma solução Bayesiana para a estimação do vetor β de parâmetros para o modelo logístico.

3.4.4 Análise Bayesiana para a regressão logística

Considere o problema da Inferência Bayesiana do modelo de regressão logística padrão com verossimilhança

$$L(\beta) = \prod_{i=1}^n \frac{\exp(y_i x_i' \beta)}{1 + \exp(x_i' \beta)}, \quad (3.13)$$

onde y_i são respostas binárias e x_i é um vetor $(p + 1)$ -dimensional de covariáveis.

De modo a ilustrar a técnica Bayesiana de estimação do modelo logístico, na próxima seção a aplicaremos a um conjunto de dados estudado por Hosmer e Lemeshow (1989).

3.4.5 Análise de dados utilizando o modelo logístico Bayesiano

Considere o conjunto de dados estudado por Hosmer and Lemeshow (1989), que relaciona informações sobre 189 nascimentos de filhos de mulheres atendidas pelo Centro Médico de Baystate em Springfield, Massachusetts, em 1986. O objetivo do estudo foi identificar os principais fatores relacionados ao baixo peso ao nascer.

Tabela 3.1: Código das variáveis associadas ao baixo peso de bebês ao nascer

Variável	Descrição	Valores
Y	Bebê com baixo peso ao nascer	0 se $\geq 2500g$ 1 se $< 2500g$
IDADE	Idade da mãe	Anos
PESO	Peso da mãe no último período menstrual	Libras
RAÇA1	Raça 1	1 = negro 0 = não negro
RAÇA2	Raça 2	1 = outra raça 0 = branco ou negro
TABACO	Presença de tabagismo durante a gestação	0 = Não 1 = Sim
PARTOP	História de trabalho de parto prematura	0 = Nenhum 1 = Um 2 = Dois ou mais
HIPER	História de hipertensão	0 = Não 1 = Sim
IRRITU	Presença de irritabilidade uterina	0 = Não 1 = Sim

As variáveis do estudo estão descritas na Tabela 3.1. Dentre elas encontram-se algumas relacionadas às condições físicas específicas no período, comportamento durante a gestação, como o hábito de fumar, e características pessoais como raça.

O modelo frequentista de Hosmer e Lemeshow (1989) utilizado para descrever a probabilidade de um recém-nascido ter baixo peso ao nascer, isto é, possuir peso inferior a $2.500g$, envolve oito covariáveis (vide Tabela 3.1).

Na próxima seção apresentaremos uma análise comparativa dos valores estimados dos parâmetros obtidos por Hosmer e Lemeshow (1989), obtidos com base em uma análise frequentista, com os valores estimados que obtivemos com a técnica Bayesiana.

De modo a obter as estimativas Bayesianas desenvolvemos um código na linguagem Fortran. Os 4.000 resultados foram obtidos considerando-se uma cadeia com 20.000 observações das quais as 4.000 primeiras foram descartadas uma vez que correspondem ao período de *Burn-in*. Efetuou-se uma amostragem sistemática com período 4, de modo a se obter uma amostra pseudo-aleatória. Ao final considerou-se, na análise Bayesiana, uma amostra gerada de valores de β de tamanho 4.000.

Tabela 3.2: Estimativas para o modelo Logístico Multivariado

Variável	Estimativas encontradas em Hosmer & Lemeshow		Estimativas obtidas através modelo Bayesiano Logístico	
	Coefficientes	Desvio Padrão	Coefficientes	Desvio Padrão
INTERCEPTO	0,464	1,201	0,48067	1,36618
IDADE	-0,027	0,036	-0,02955	0,03678
PESO	-0,015	0,007	-0,01542	0,00619
RAÇA1	1,263	0,525	1,27206	0,57809
RAÇA2	0,862	0,438	0,88024	0,50315
TABACO	0,923	0,400	0,93868	0,44086
PARTOP	0,542	0,345	0,54333	0,37044
HIPER	1,834	0,690	1,86228	0,96050
IRRITU	0,759	0,459	0,76749	0,46591

3.4.6 Comparação dos Resultados

Na Tabela 3.2 apresentamos as estimativas pontuais e o desvio padrão, dos coeficientes do modelo logístico proposto por Hosmer e Lemeshow (1989), para prever a probabilidade de baixo peso ao nascer.

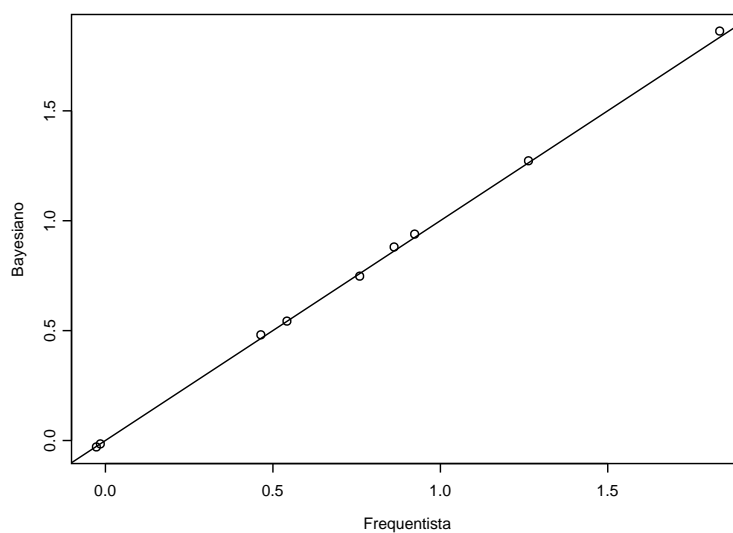


Figura 3.1: Diagrama de dispersão - Frequentista X Bayesiano

Na Tabela 3.2 são apresentadas tanto as estimativas frequentistas obtidas por Hosmer e Lemeshow (1989) quanto as estimativas Bayesianas que obtivemos através da análise logística Bayesiana. Pode-se perceber que há uma grande similaridade entre os valores estimados utilizando ambos os métodos. Na figura 3.1 apresenta-se o diagrama de dispersão entre os valores estimados dos coeficientes obtidos por meio dos métodos Bayesiano e frequentista. Observa-se um ajuste quase perfeito.

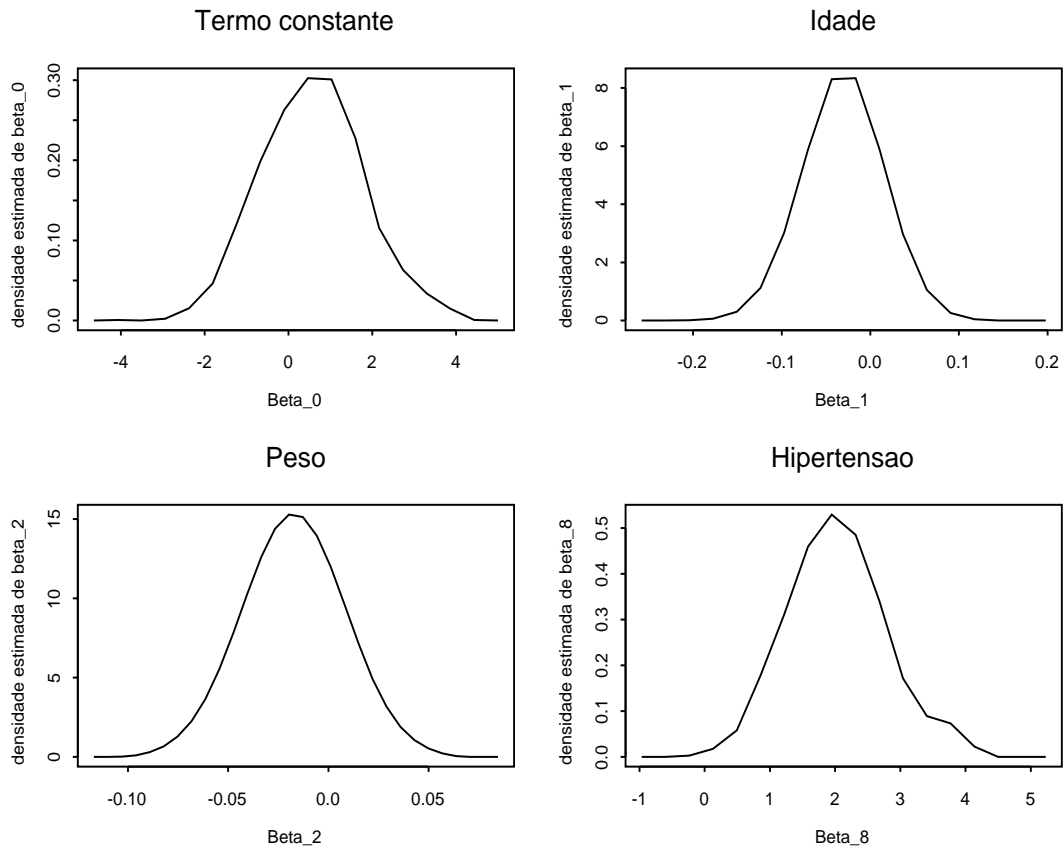


Figura 3.2: Densidades estimadas de $\beta_0, \beta_1, \beta_2, \beta_7$

A fim de ilustrar o processo de estimação dos parâmetros, na Figura 3.2 apresentamos os gráficos correspondentes às densidades *a posteriori* estimadas do intercepto e daqueles parâmetros que determinam as covariáveis idade, peso e história de hipertensão.

Na Figura 3.3 apresentamos os gráficos de dispersão daqueles parâmetros para as 20.000 observações. Embora a convergência da cadeia ainda não tenha sido avaliada

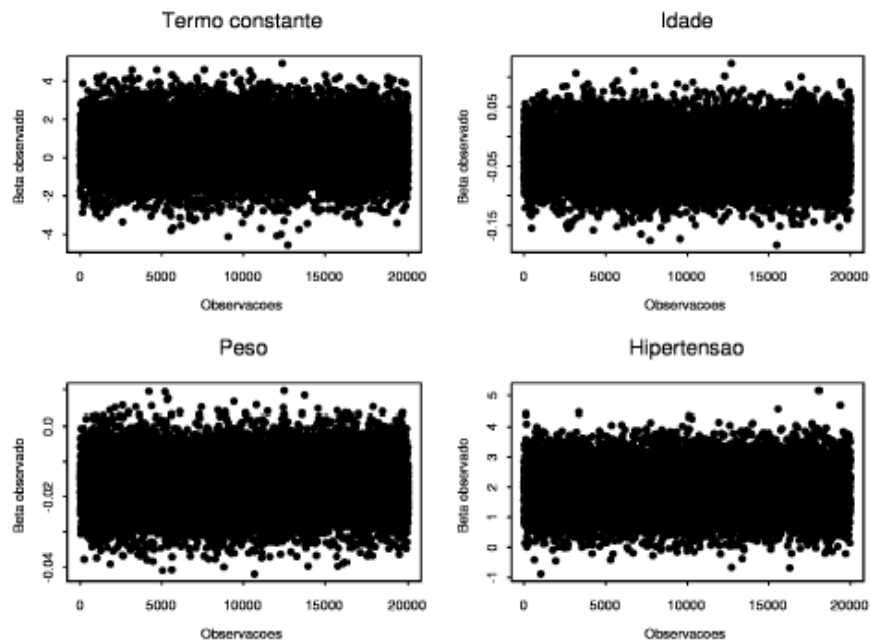


Figura 3.3: Gráficos de dispersão dos betas *versus* número da observação

formalmente a partir de diversos procedimentos já estabelecidos na literatura, a partir desta figura podemos notar que as nuvens de pontos são bastante densas em torno das estimativas dos parâmetros, o que sugere que a convergência, para cada um dos parâmetros analisados, foi atingida.

No próximo capítulo voltaremos a discutir a modelagem Bayesiana dos dados relativos às bowheads.

Capítulo 4

Especificação de Modelos

Neste capítulo descreveremos um total de doze modelos, a serem comparados, em que o logito das probabilidades de captura é escrito como uma função linear de covariáveis, e as taxas de sobrevivência são ora consideradas constantes, para cada grupo, e ora consideradas variando com as ocasiões amostrais.

Como no modelo frequentista de Zeh *et al.* (2002) consideremos, inicialmente, o modelo logístico

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}, \quad (4.1)$$

onde

- p_{tg} é a probabilidade de captura no tempo t para o grupo g ;
- $x_{tg_1} = g - 1$ é uma variável que designa grupo, e
- x_{tg_2} é uma variável que designa o número de dias empregados fotografando os indivíduos do grupo g no tempo t .

Os modelos a serem estudados, que representam casos particulares do modelo de Zeh *et al.* (2002), são sumarizados na Tabela 4.1 a seguir. A Tabela 4.2 apresenta os parâmetros a serem estimados para cada um dos modelos.

Nos modelos 1 a 4 considera-se que o logito da probabilidade de captura no tempo t e grupo g , $\text{logito}(p_{tg})$, é expresso como uma função linear que depende

Tabela 4.1: Modelos

Modelo	Probabilidade de captura	Taxas de sobrevivência
1	$\text{logito}(p_{tg}) = \beta_0 + \beta_2 x_{tg_2}$	$\phi_{tg} = \phi; t = 1, \dots, s-1 \text{ e } g = 1, 2$
2	$\text{logito}(p_{tg}) = \beta_0 + \beta_2 x_{tg_2}$	$\phi_{tg} = \phi_g; t = 1, \dots, s-1 \text{ e } g = 1, 2$
3	$\text{logito}(p_{tg}) = \beta_0 + \beta_2 x_{tg_2}$	$\phi_{tg} = \phi_t; t = 1, \dots, s-1 \text{ e } g = 1, 2$
4	$\text{logito}(p_{tg}) = \beta_0 + \beta_2 x_{tg_2}$	$\phi_{tg}; t = 1, \dots, s-1 \text{ e } g = 1, 2$
5	$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g} x_{tg_2}$	$\phi_{tg} = \phi; t = 1, \dots, s-1 \text{ e } g = 1, 2$
6	$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g} x_{tg_2}$	$\phi_{tg} = \phi_g; t = 1, \dots, s-1 \text{ e } g = 1, 2$
7	$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g} x_{tg_2}$	$\phi_{tg} = \phi_t; t = 1, \dots, s-1 \text{ e } g = 1, 2$
8	$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g} x_{tg_2}$	$\phi_{tg}; t = 1, \dots, s-1 \text{ e } g = 1, 2$
9	$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$	$\phi_{tg} = \phi; t = 1, \dots, s-1 \text{ e } g = 1, 2$
10	$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$	$\phi_{tg} = \phi_g; t = 1, \dots, s-1 \text{ e } g = 1, 2$
11	$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$	$\phi_{tg} = \phi_t; t = 1, \dots, s-1 \text{ e } g = 1, 2$
12	$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$	$\phi_{tg}; t = 1, \dots, s-1 \text{ e } g = 1, 2$

apenas do tempo t , isto é, sem variações devidas ao efeito de grupo. Além disso, as taxas de sobrevivência são expressas de 4 formas distintas, a saber: no modelo 1 - taxas de sobrevivência constantes para todos os tempos e grupos; no modelo 2 - taxas de sobrevivência variando apenas com o grupo; no modelo 3 - taxas de sobrevivência variando apenas com o tempo; no modelo 4 - taxas de sobrevivência variando com tempo e grupo.

Nos modelos 5 a 8 considera-se que o logito da probabilidade de captura, $\text{logito}(p_{tg})$, também é expresso como uma função linear do tempo, mas admite-se funções lineares distintas para cada grupo. Quanto às taxas de sobrevivência, considera-se os mesmos 4 modelos, descritos na Tabela 4.1.

Os modelos 9 a 12 representam extensões dos modelos 1 a 4, respectivamente, onde se considera que as probabilidades de captura variam com o tempo e grupo. Porém, um mesmo modelo linear é utilizado para descrever o $\text{logito}(p_{tg})$. Novamente, no caso da modelagem das taxas de sobrevivência, prevê-se 4 tipos de modelos.

A seguir descreveremos as verossimilhanças associadas a cada um dos modelos.

Tabela 4.2: Parâmetros a estimar em cada um dos modelos

Modelo	Parâmetros a estimar
1	β_0, β_1, ϕ
2	$\beta_0, \beta_1, \phi_1, \phi_2$
3	$\beta_0, \beta_1, \phi_1, \dots, \phi_{s-1}$
4	$\beta_0, \beta_1, \phi_{11}, \dots, \phi_{s-1,1}, \phi_{12}, \dots, \phi_{s-1,2}$
5	$\beta_{01}, \beta_{02}, \beta_{11}, \beta_{12}, \phi$
6	$\beta_{01}, \beta_{02}, \beta_{11}, \beta_{12}, \phi_1, \phi_2$
7	$\beta_{01}, \beta_{02}, \beta_{11}, \beta_{12}, \phi_1, \dots, \phi_{s-1}$
8	$\beta_{01}, \beta_{02}, \beta_{11}, \beta_{12}, \phi_{11}, \dots, \phi_{s-1,1}, \phi_{12}, \dots, \phi_{s-1,2}$
9	$\beta_0, \beta_1, \beta_2, \phi$
10	$\beta_0, \beta_1, \beta_2, \phi_1, \phi_2$
11	$\beta_0, \beta_1, \beta_2, \phi_1, \dots, \phi_{s-1}$
12	$\beta_0, \beta_1, \beta_2, \phi_{11}, \dots, \phi_{s-1,1}, \phi_{12}, \dots, \phi_{s-1,2}$

4.1 Modelo 1

Como descrito anteriormente, para o modelo 1 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg2}, \quad (4.2)$$

e $\phi_{tg} = \phi$; $t = 1, \dots, s-1$ e $g = 1, \dots, G$. Assim,

$$p_{tg} = \frac{\exp(\beta_0 + \beta_1 x_{tg2})}{1 + \exp(\beta_0 + \beta_1 x_{tg2})}.$$

Desta forma, a verossimilhança associada ao modelo 1 é dada por:

$$\begin{aligned} L &\propto \left[\prod_{t=1}^{s-1} \prod_{g=1}^G (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{z_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1 x_{t+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{m_{t+1,g}} \times \phi^{\sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t (z_{t+1,g} + m_{t+1,g})}, \end{aligned} \quad (4.3)$$

onde Δ_t representa o número de anos entre os tempos t e $t+1$, e

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \left(\frac{p_{k+1,g}}{1 - p_{k+1,g}} \right) \prod_{i=t}^k \phi^{\Delta_i} \times (1 - p_{i+1,g}) \right\}, \quad (4.4)$$

onde $\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg2}$.

A distribuição condicional completa *a posteriori* de β é dada por

$$\pi(\beta \mid \phi, D) \propto L \times \pi(\beta).$$

Logo, a respectiva log-condicional completa *a posteriori* de β é dada por:

$$\begin{aligned} l = \log \pi(\beta \mid \phi, D) &\propto \sum_{g=1}^G \sum_{t=1}^{s-1} c_{tg} \log \chi_{tg} + \sum_{g=1}^G \sum_{t=1}^{s-1} m_{t+1,g} (\beta_0 + \beta_1 x_{t+1,g2}) \\ &- \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \log(1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})) \\ &- \frac{1}{2} (\beta_0^2 + \beta_1^2). \end{aligned} \quad (4.5)$$

Como descrito na seção 3.1, a distribuição condicional completa *a posteriori* de ϕ é dada por

$$\begin{aligned} \pi(\phi \mid \beta, D) &\propto \left\{ \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \phi^{\sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g})} \right\} \times \phi^{\alpha-1} (1 - \phi)^{\gamma-1} \\ &= \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \phi^{\sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) + (\alpha-1)} \times (1 - \phi)^{\gamma-1}. \end{aligned} \quad (4.6)$$

Note que ϕ é função de $\chi_{tg} \forall t$ e $\forall g$.

4.2 Modelo 2

Para o modelo 2 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg2}, \quad (4.7)$$

e $\phi_{tg} = \phi_g$; $t = 1, \dots, s-1$; $g = 1, \dots, G$.

Desta forma, a verossimilhança associada ao modelo 2 é dada por:

$$\begin{aligned} L &\propto \left[\prod_{t=1}^{s-1} \prod_{g=1}^G (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{z_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1 x_{t+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{m_{t+1,g}} \times \prod_{g=1}^G \phi_g^{\sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g})}. \end{aligned} \quad (4.8)$$

Logo, a respectiva log-condicional completa *a posteriori* de β também é dada pela expressão (4.5). No entanto, χ_{tg} é expresso por

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \left(\frac{p_{k+1,g}}{1 - p_{k+1,g}} \right) \prod_{i=t}^k \phi_g^{\Delta_i} \times (1 - p_{i+1,g}) \right\}, \quad (4.9)$$

onde $\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg2}$.

A distribuição condicional completa *a posteriori* do vetor ϕ é dada a seguir

$$\pi(\phi_1, \dots, \phi_G \mid \beta, D) = \prod_{g=1}^G \pi(\phi_g \mid \beta, D),$$

sob a suposição de que os ϕ_1, \dots, ϕ_G são permutáveis. Desta forma, para cada ϕ_g , $g = 1, \dots, G$,

$$\begin{aligned} \pi(\phi_g \mid \beta, D) &\propto \left\{ \left[\prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \phi_t^{\sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g})} \right\} \times \phi_g^{\alpha-1} (1 - \phi_g)^{\gamma-1} \\ &= \left[\prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \phi_g^{\sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) + (\alpha-1)} \times (1 - \phi_g)^{\gamma-1}, \end{aligned} \quad (4.10)$$

para $g = 1, \dots, G$. Note que ϕ_g é função de $\chi_{tg} \forall t$.

4.3 Modelo 3

Para o modelo 3 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg2}, \quad (4.11)$$

e $\phi_{tg} = \phi_t$; $t = 1, \dots, s-1$ e $g = 1, \dots, G$. Portanto, a verossimilhança associada ao modelo 3 é dada por

$$\begin{aligned} L &\propto \left[\prod_{t=1}^{s-1} \prod_{g=1}^G (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{z_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1 x_{t+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{m_{t+1,g}} \times \prod_{t=1}^{s-1} \phi_t^{\sum_{g=1}^G (z_{t+1,g} + m_{t+1,g})}. \end{aligned} \quad (4.12)$$

Portanto, a respectiva log-condicional completa *a posteriori* de β é novamente dada pela expressão (4.5). No entanto, χ_{tg} é dado por

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \left(\frac{p_{k+1,g}}{1 - p_{k+1,g}} \right) \prod_{i=t}^k \phi_i \times (1 - p_{i+1,g}) \right\}, \quad (4.13)$$

onde $\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_2}$.

Note que

ϕ_1 é função de χ_{1g} , $g = 1, \dots, G$;

ϕ_2 é função de χ_{1g} e de χ_{2g} , $g = 1, \dots, G$;

\vdots

ϕ_t é função de $\chi_{1g}, \chi_{2g}, \dots, \chi_{tg}$, $g = 1, \dots, G$;

\vdots

ϕ_{s-1} é função de $\chi_{1g}, \chi_{2g}, \dots, \chi_{s-1,g}$, $g = 1, \dots, G$.

A distribuição condicional completa *a posteriori* do vetor ϕ é dada por

$$\pi(\phi_1, \dots, \phi_{s-1} \mid \beta, D) = \prod_{t=1}^{s-1} \pi(\phi_t \mid \beta, D),$$

sob a suposição de que os $\phi_1, \dots, \phi_{s-1}$ são permutáveis. Desta forma, para cada ϕ_t , $t = 1, \dots, s - 1$,

$$\begin{aligned} \pi(\phi_t \mid \beta, D) &\propto \left\{ \left[\prod_{g=1}^G (\chi_{tg})^{c_{tg}} \right] \phi_t^{\sum_{g=1}^G (z_{t+1,g} + m_{t+1,g})} \right\} \times \phi_t^{\alpha-1} (1 - \phi_t)^{\gamma-1} \\ &= \left[\prod_{g=1}^G \prod_{i=1}^t (\chi_{ig})^{c_{ig}} \right] \phi_t^{\sum_{g=1}^G (z_{t+1,g} + m_{t+1,g}) + (\alpha-1)} \times (1 - \phi_t)^{\gamma-1}. \end{aligned} \quad (4.14)$$

4.4 Modelo 4

Para o modelo 4 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_2}, \quad (4.15)$$

e $\phi_{tg}, t = 1, \dots, s-1; g = 1, \dots, G$. Logo, a verossimilhança associada ao modelo 4 é dada por

$$L \propto \left[\prod_{t=1}^{s-1} \prod_{g=1}^G (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{z_{t+1,g}} \\ \times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1 x_{t+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g2})} \right]^{m_{t+1,g}} \times \prod_{g=1}^G \prod_{t=1}^{s-1} \phi_{tg}^{(z_{t+1,g} + m_{t+1,g})}. \quad (4.16)$$

Portanto, a respectiva log-condicional completa *a posteriori* de β é dada, novamente, pela expressão (4.5). No entanto, χ_{tg} é dada por

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \left(\frac{p_{k+1,g}}{1 - p_{k+1,g}} \right) \prod_{i=t}^k \phi_{ig} \times (1 - p_{i+1,g}) \right\}, \quad (4.17)$$

onde $\logit(p_{tg}) = \beta_0 + \beta_1 x_{tg2}$.

Note que

ϕ_{1g} é função de χ_{1g} , $g = 1, \dots, G$;

ϕ_{2g} é função de χ_{1g} e de χ_{2g} , $g = 1, \dots, G$;

\vdots

ϕ_{tg} é função de $\chi_{1g}, \chi_{2g}, \dots, \chi_{tg}$, $g = 1, \dots, G$;

\vdots

$\phi_{s-1,g}$ é função de $\chi_{1g}, \chi_{2g}, \dots, \chi_{s-1,g}$, $g = 1, \dots, G$.

A distribuição condicional completa *a posteriori* do vetor ϕ é dada por

$$\pi(\phi_{11}, \dots, \phi_{1G}, \dots, \phi_{s-1,1}, \dots, \phi_{s-1,G} \mid \beta, D) = \prod_{g=1}^G \prod_{t=1}^{s-1} \pi(\phi_{tg} \mid \beta, D), \quad (4.18)$$

sob a suposição de que os ϕ_{tg} 's são permutáveis. Desta forma, para cada $\phi_{tg}, g = 1, \dots, G$ e $t = 1, \dots, s-1$,

$$\pi(\phi_{tg} \mid \beta, D) \propto \left[\prod_{i=1}^t (\chi_{ig})^{c_{ig}} \right] \phi_{tg}^{(z_{t+1,g} + m_{t+1,g})} \times \phi_{tg}^{\alpha-1} (1 - \phi_{tg})^{\gamma-1} \\ = \left[\prod_{i=1}^t (\chi_{ig})^{c_{ig}} \right] \phi_{tg}^{(z_{t+1,g} + m_{t+1,g} + (\alpha-1))} \times (1 - \phi_{tg})^{\gamma-1}. \quad (4.19)$$

4.5 Modelo 5

Considere que

$$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2},$$

e $\phi_{tg} = \phi$; $t = 1, \dots, s-1$ e $g = 1, \dots, G$, ou seja,

$$p_{tg} = \frac{\exp(\beta_{0g} + \beta_{1g}x_{tg_1})}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})}, \quad g = 1, \dots, G, \quad t = 1, \dots, s-1.$$

Desta forma, a verossimilhança associada ao modelo 5 é dada por:

$$\begin{aligned} L \propto & \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{z_{t+1,g}} \\ & \times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_{0g} + \beta_{1g}x_{tg_1})}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{m_{t+1,g}} \times \phi^{\sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g})}. \end{aligned} \quad (4.20)$$

Portanto, a respectiva log-condicional completa *a posteriori* de β é dada por

$$\begin{aligned} l = \log \pi(\beta | D) \propto & \sum_{g=1}^G \sum_{t=1}^{s-1} c_{tg} \log \chi_{tg} + \sum_{g=1}^G \sum_{t=1}^{s-1} m_{t+1,g} (\beta_{0g} + \beta_{1g}x_{t+1,g_2}) \\ & - \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \log(1 + \exp(\beta_{0g} + \beta_{1g}x_{t+1,g_2})) \\ & - \frac{1}{2}(\beta_{01}^2 + \beta_{11}^2) - \frac{1}{2}(\beta_{02}^2 + \beta_{12}^2), \end{aligned} \quad (4.21)$$

onde χ_{tg} é expresso de acordo com a fórmula (4.4), mas com $\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2}$. A distribuição condicional completa *a posteriori* de ϕ é dada pela fórmula (4.6) com a devida alteração de χ_{tg} .

4.6 Modelo 6

Considere que

$$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2},$$

e $\phi_{tg} = \phi_g$, $t = 1, \dots, s-1$, e $g = 1, \dots, G$.

Desta forma, a verossimilhança associada ao modelo 6 é dada por:

$$L \propto \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{z_{t+1,g}} \\ \times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_{0g} + \beta_{1g}x_{tg_1})}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{m_{t+1,g}} \times \prod_{g=1}^G \phi_g^{\sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g})}. \quad (4.22)$$

Portanto, a respectiva log-condicional completa *a posteriori* de β é dada pela expressão (4.21). No entanto, χ_{tg} é expresso de acordo com a fórmula (4.9), mas com $\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2}$. A distribuição condicional completa *a posteriori* de ϕ_g , $g = 1, \dots, G$ é dada pela fórmula (4.10), com a devida mudança em χ_{tg} .

4.7 Modelo 7

Considere que

$$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2},$$

e $\phi_{tg} = \phi_t$, $t = 1, \dots, s-1$, $g = 1, \dots, G$.

Desta forma, a verossimilhança associada ao modelo 7 será dada por:

$$L \propto \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{z_{t+1,g}} \\ \times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_{0g} + \beta_{1g}x_{tg_1})}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{m_{t+1,g}} \times \prod_{t=1}^{s-1} \phi_t^{\sum_{g=1}^G (z_{t+1,g} + m_{t+1,g})}. \quad (4.23)$$

Logo, a respectiva log-condicional completa *a posteriori* de β também é dada pela expressão (4.21). No entanto, χ_{tg} é expresso de acordo com a fórmula (4.13), mas com $\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2}$. A distribuição condicional completa *a posteriori* de ϕ_t , $t = 1, \dots, s-1$, é dada pela expressão (4.14), com a devida alteração em χ_{tg} .

4.8 Modelo 8

Considere que

$$\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2},$$

e $\phi_{tg}, t = 1, \dots, s-1$ e $g = 1, \dots, G$.

Desta forma, a verossimilhança associada ao modelo 8 é dada por:

$$\begin{aligned} L \propto & \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{z_{t+1,g}} \\ & \times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_{0g} + \beta_{1g}x_{tg_1})}{1 + \exp(\beta_{0g} + \beta_{1g}x_{tg_1})} \right]^{m_{t+1,g}} \times \prod_{g=1}^G \prod_{t=1}^{s-1} \phi_{tg}^{(z_{t+1,g} + m_{t+1,g})}. \end{aligned} \quad (4.24)$$

Assim, a respectiva log-condicional completa *a posteriori* de β é novamente dada pela expressão (4.21). No entanto, χ_{tg} é expresso de acordo com a fórmula (4.17), mas com $\text{logito}(p_{tg}) = \beta_{0g} + \beta_{1g}x_{tg_2}$.

A distribuição condicional completa *a posteriori* de ϕ_{tg} , $t = 1, \dots, s-1$ e $g = 1, \dots, G$ é dada pela fórmula (4.19), com a devida mudança de χ_{tg} .

4.9 Modelo 9

Para o modelo 9 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1x_{t,g_1} + \beta_2x_{t,g_2}, \quad (4.25)$$

e $\phi_{tg} = \phi$, $t = 1, \dots, s-1$ e $g = 1, \dots, G$, ou seja,

$$p_{tg} = \frac{\exp(\beta_0 + \beta_1x_{t,g_1} + \beta_2x_{t,g_2})}{1 + \exp(\beta_0 + \beta_1x_{t,g_1} + \beta_2x_{t,g_2})}, \quad t = 1, \dots, s-1 \text{ e } g = 1, \dots, G.$$

Desta forma, a verossimilhança associada ao modelo 9 é dada por

$$\begin{aligned} L \propto & \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1x_{t+1,g_1} + \beta_2x_{t+1,g_2})} \right]^{z_{t+1,g}} \\ & \times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1x_{t+1,g_1} + \beta_2x_{t+1,g_2})}{1 + \exp(\beta_0 + \beta_1x_{t+1,g_1} + \beta_2x_{t+1,g_2})} \right]^{m_{t+1,g}} \times \phi^{\sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g})}. \end{aligned}$$

Assim, a respectiva log-condicional completa *a posteriori* de β é dada por

$$\begin{aligned}
l = \log \pi(\beta \mid D) &\propto \sum_{g=1}^G \sum_{t=1}^{s-1} c_{tg} \log \chi_{tg} + \sum_{g=1}^G \sum_{t=1}^{s-1} m_{t+1,g} (\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2}) \\
&- \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \log(1 + \exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})) \\
&- \frac{1}{2} (\beta_0^2 + \beta_1^2 + \beta_2^2), \tag{4.26}
\end{aligned}$$

onde χ_{tg} é expresso de acordo com a fórmula (4.4), mas com $\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$. A distribuição condicional completa *a posteriori* de ϕ é dada pela expressão (4.6), com a devida alteração em χ_{tg} .

4.10 Modelo 10

Para o modelo 10 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2},$$

e $\phi_{tg} = \phi_g$, $t = 1, \dots, s-1$, $g = 1, \dots, G$.

Desta forma, a verossimilhança associada ao modelo 10 é dada por

$$\begin{aligned}
L &\propto \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})} \right]^{z_{t+1,g}} \\
&\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})} \right]^{m_{t+1,g}} \times \prod_{g=1}^G \phi_g^{\sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g})}.
\end{aligned}$$

Logo, a respectiva log-condicional completa *a posteriori* de β é dada pela expressão (4.26), onde χ_{tg} é expresso de acordo com a fórmula (4.9), mas com $\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$. A distribuição condicional completa *a posteriori* de ϕ é dada pela expressão (4.10), com a devida alteração em χ_{tg} .

4.11 Modelo 11

Para o modelo 11 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2},$$

e $\phi_{tg} = \phi_t, t = 1, \dots, s-1, g = 1, \dots, G$.

Desta forma, a verossimilhança associada ao modelo 11 é dada por

$$\begin{aligned} L &\propto \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})} \right]^{z_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})} \right]^{m_{t+1,g}} \times \prod_{t=1}^{s-1} \phi_t^{\sum_{g=1}^G (z_{t+1,g} + m_{t+1,g})}. \end{aligned}$$

Desta forma, a respectiva log-condicional completa *a posteriori* de β é expressa pela fórmula (4.26), onde χ_{tg} é dado pela expressão (4.13), mas com $\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$. A distribuição condicional completa *a posteriori* de ϕ é dada pela expressão (4.14), com a devida alteração em χ_{tg} .

4.12 Modelo 12

Para o modelo 12 temos

$$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2},$$

e $\phi_{tg}, t = 1, \dots, s-1, g = 1, \dots, G$.

Desta forma, a verossimilhança associada ao modelo 12 é dada por

$$\begin{aligned} L &\propto \left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{1}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})} \right]^{z_{t+1,g}} \\ &\times \prod_{g=1}^G \prod_{t=1}^{s-1} \left[\frac{\exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g_1} + \beta_2 x_{t+1,g_2})} \right]^{m_{t+1,g}} \times \prod_{g=1}^G \prod_{t=1}^{s-1} \phi_{tg}^{(z_{t+1,g} + m_{t+1,g})}. \end{aligned}$$

Assim, a respectiva log-condicional completa *a posteriori* de β é dada pela expressão (4.26), onde χ_{tg} é expresso de acordo com a fórmula (4.17), mas com

$\text{logito}(p_{tg}) = \beta_0 + \beta_1 x_{tg_1} + \beta_2 x_{tg_2}$. A distribuição condicional completa *a posteriori* de ϕ é dada pela expressão (4.19), com a devida alteração em χ_{tg} .

Recordando que, como descrito nas seções 3.3 e 3.3.1, na especificação da função geradora de candidatos de β faz-se necessário o cálculo da informação de Fisher observada, \hat{H} , relativa à *posteriori* de β , no próximo capítulo descreveremos, para cada um dos doze modelos, o cálculo de \hat{H} .

Capítulo 5

Obtenção das derivadas parciais da log- condicional completa *a posteriori* de β para a obtenção de \hat{H}

Neste capítulo descrevemos as derivadas parciais da log-condicional completa a posteriori de β para os 12 modelos descritos no Capítulo 4. Para descrevê-las faz-se necessário um estudo do termo χ_{tg} , em cada caso. Este estudo é feito a seguir.

5.1 Estudando χ_{tg}

5.1.1 Modelos 1, 5 e 9

Recordemos que, para a taxa de sobrevivência, ϕ , constante ao longo do tempo e dos grupos (modelos 1, 5, 9 - Capítulo 4), temos

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \left(\frac{p_{k+1,g}}{1 - p_{k+1,g}} \right) \prod_{i=t}^k \phi^{\Delta_i} \times (1 - p_{i+1,g}) \right\}. \quad (5.1)$$

Desta forma,

$$\chi_{tg}(\beta) = 1 - \sum_{k=t}^{s-1} \left\{ \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi^{\Delta_i} \frac{1}{1 + \exp(\beta' X_{i+1,g})} \right\}.$$

Seja

$$V_{kt} = \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi^{\Delta_i} \frac{1}{1 + \exp(\beta' X_{i+1,g})}, \quad (5.2)$$

e, portanto,

$$\log V_{kt} = \beta' X_{k+1,g} + \sum_{i=t}^k \Delta_i \log \phi - \sum_{i=t}^k \log[1 + \exp(\beta' X_{i+1,g})].$$

Então,

$$\chi_{tg}(\beta) = 1 - \sum_{k=t}^{s-1} V_{kt} = 1 - \sum_{k=t}^{s-1} \exp(\log V_{kt}). \quad (5.3)$$

5.1.2 Modelos 2, 6 e 10

Para o caso em que ϕ varia de grupo para grupo, mas permanece constante ao longo do tempo (modelos 2, 6, 10), temos

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi_g^{\Delta_i} \frac{1}{1 + \exp(\beta' X_{i+1,g})} \right\},$$

Assim, o termo V_{kt} é expresso por

$$V_{kt} = \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi_g^{\Delta_i} \frac{1}{1 + \exp(\beta' X_{i+1,g})}. \quad (5.4)$$

Então, χ_{tg} é expresso pela fórmula (5.3), com as devidas alterações no termo V_{kt} .

5.1.3 Modelos 3, 7 e 11

Nos modelos em que a taxa de sobrevivência, ϕ , varia a cada ocasião amostral, mas é a mesma para todos os grupos (modelos 3, 7, 11), temos

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi_i \frac{1}{1 + \exp(\beta' X_{i+1,g})} \right\}. \quad (5.5)$$

Desta forma,

$$V_{kt} = \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi_i \frac{1}{1 + \exp(\beta' X_{i+1,g})}. \quad (5.6)$$

Assim, χ_{tg} é expresso pela fórmula (5.3), com as devidas alterações no termo V_{kt} .

5.1.4 Modelos 4, 8 e 12

Para os modelos em que a taxa de sobrevivência, ϕ , varia com o tempo e o grupo (modelos 4, 8, 12), temos

$$\chi_{tg} = 1 - \sum_{k=t}^{s-1} \left\{ \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi_{ig} \frac{1}{1 + \exp(\beta' X_{i+1,g})} \right\}. \quad (5.7)$$

Desta forma,

$$V_{kt} = \exp(\beta' X_{k+1,g}) \prod_{i=t}^k \phi_{ig} \frac{1}{1 + \exp(\beta' X_{i+1,g})}. \quad (5.8)$$

Portanto, χ_{tg} é expresso pela fórmula (5.3) com as devidas mudanças em V_{kt} .

Nas próximas seções obteremos as derivadas parciais da log-condicional completa *a posteriori* do vetor β , necessárias ao cômputo da matriz de informação de Fisher estimada, \hat{H} , (vide expressão (3.6)), utilizada na função geradora de candidatos β^* de β (vide expressão (3.5)).

5.2 Obtenção das derivadas parciais da log- condicional completa *a posteriori* de β para a obtenção de \hat{H} para os modelos 1, 2, 3 e 4

As derivadas primeiras de l (vide expressão (4.5)) são dadas pelas expressões que se seguem:

$$\begin{aligned} \bullet \frac{\partial l}{\partial \beta_0} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial}{\partial \beta_0} c_{tg} \log \chi_{tg} + \sum_{g=1}^G \sum_{t=1}^{s-1} m_{t+1,g} \\ &\quad - \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta_0 + \beta_1 x_{t+1})}{1 + \exp(\beta_0 + \beta_1 x_{t+1})} - \beta_0. \end{aligned}$$

$$\begin{aligned} \bullet \frac{\partial l}{\partial \beta_1} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial}{\partial \beta_1} c_{tg} \log \chi_{tg} + \sum_{g=1}^G \sum_{t=1}^{s-1} m_{t+1,g} x_{t+1,g_2} \\ &\quad - \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g_2} \exp(\beta_0 + \beta_1 x_{t+1,g_2})}{1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})} - \beta_1. \end{aligned}$$

Assim, as derivadas parciais de segunda ordem de l são dadas por:

$$\begin{aligned} \bullet \frac{\partial^2 l}{\partial \beta_0^2} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0^2} c_{tg} \log \chi_{tg} \\ &\quad - \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{(1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2}))(\exp(\beta_0 + \beta_1 x_{t+1,g_2}))}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} \\ &\quad + \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{(\exp(\beta_0 + \beta_1 x_{t+1,g_2}))(\exp(\beta_0 + \beta_1 x_{t+1,g_2}))}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} \\ &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0^2} c_{tg} \log \chi_{tg} - \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta_0 + \beta_1 x_{t+1,g_2})}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} - 1. \end{aligned}$$

$$\begin{aligned}
\bullet \frac{\partial^2 l}{\partial \beta_1^2} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_1^2} c_{tg} \log \chi_{tg} \\
&- \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g_2}^2 \exp(\beta_0 + \beta_1 x_{t+1,g_2}) [(1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2}))]}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} \\
&+ \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g_2} (\exp(\beta_0 + \beta_1 x_{t+1,g_2})) x_{t+1,g_2} (\exp(\beta_0 + \beta_1 x_{t+1,g_2}))}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} - 1 \\
&= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_1^2} c_{tg} \log \chi_{tg} \\
&- \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g_2}^2 (\exp(\beta_0 + \beta_1 x_{t+1,g_2}))}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} - 1.
\end{aligned}$$

$$\begin{aligned}
\bullet \frac{\partial^2 l}{\partial \beta_0 \partial \beta_1} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0 \partial \beta_1} c_{tg} \log \chi_{tg} \\
&- \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g_2} \exp(\beta_0 + \beta_1 x_{t+1,g_2}) [1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} \\
&+ \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g_2} \exp(2(\beta_0 + \beta_1 x_{t+1,g_2}))}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2} \\
&= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0 \partial \beta_1} c_{tg} \log \chi_{tg} \\
&- \sum_{g=1}^G \sum_{t=1}^{s-1} x_{t+1,g_2} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta_0 + \beta_1 x_{t+1,g_2})}{[1 + \exp(\beta_0 + \beta_1 x_{t+1,g_2})]^2}.
\end{aligned}$$

A partir da fórmula (5.3) podemos expressar as derivadas primeiras de χ_{tg} , com respeito a β_0 , considerando V_{kt} como expresso em (5.2), (5.4), (5.6) e (5.8), que incorporam as modelagens para ϕ , da seguinte forma:

$$\frac{\partial}{\partial \beta_0} \chi_{tg} = - \sum_{k=t}^{s-1} \frac{\partial}{\partial \beta_0} \exp(\log V_{kt}) = - \sum_{k=t}^{s-1} \frac{\partial \log V_{kt}}{\partial \beta_0} \exp(\log V_{kt}). \quad (5.9)$$

Mas,

$$\frac{\partial \log V_{kt}}{\partial \beta_0} = 1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1,g_2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g_2})}. \quad (5.10)$$

Portanto,

$$\bullet \frac{\partial}{\partial \beta_0} \chi_{tg} = - \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1, g_2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})} \right] V_{kt}.$$

Procedendo de forma análoga, pode-se obter as derivadas primeiras de χ_{tg} com respeito a β_1 :

$$\bullet \frac{\partial}{\partial \beta_1} \chi_{tg} = - \sum_{k=t}^{s-1} \left[x_{k+1, g_2} - \sum_{i=t}^k \frac{x_{i+1, g_2} \exp(\beta_0 + \beta_1 x_{i+1, g_2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})} \right] V_{kt}.$$

Através das expressões (5.9) e (5.10), podemos obter as derivadas segundas de χ_{tg} com respeito a β .

$$\begin{aligned} \bullet \frac{\partial^2}{\partial \beta_0^2} \chi_{tg} &= - \sum_{k=t}^{s-1} \left[\frac{\partial^2 \log V_{kt}}{\partial \beta_0^2} \exp(\log V_{kt}) + \frac{\partial}{\partial \beta_0} \exp(\log V_{kt}) \frac{\partial \log V_{kt}}{\partial \beta_0} \right] \\ &= - \sum_{k=t}^{s-1} \left\{ \frac{\partial^2 \log V_{kt}}{\partial \beta_0^2} \exp(\log V_{kt}) + \left[\frac{\partial \log V_{kt}}{\partial \beta_0} \right]^2 \exp(\log V_{kt}) \right\}. \end{aligned}$$

Logo,

$$\begin{aligned} \bullet \frac{\partial^2}{\partial \beta_0^2} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1, g_2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})]^2} V_{kt} + \\ &\quad - \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1, g_2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})} \right]^2 V_{kt}. \end{aligned}$$

Analogamente,

$$\begin{aligned} \bullet \frac{\partial^2}{\partial \beta_1^2} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1, g_2})^2 \exp(\beta_0 + \beta_1 x_{i+1, g_2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})]^2} V_{kt} + \\ &\quad - \sum_{k=t}^{s-1} \left[x_{i+1, g_2} - \sum_{i=t}^k \frac{(x_{i+1, g_2}) \exp(\beta_0 + \beta_1 x_{i+1, g_2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})} \right]^2 V_{kt}. \end{aligned}$$

De (5.9) temos:

$$\begin{aligned}
\bullet \frac{\partial^2}{\partial \beta_0 \beta_1} \chi_{tg} &= - \sum_{k=t}^{s-1} \left\{ \frac{\partial}{\partial \beta_1} \left[\frac{\partial \log V_{kt}}{\partial \beta_0} \right] V_{kt} + \frac{\partial \log V_{kt}}{\partial \beta_1} \cdot \frac{\partial \log V_{kt}}{\partial \beta_0} V_{kt} \right\} \\
&= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1, g_2}) \exp(\beta_0 + \beta_1 x_{i+1, g_2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})]^2} V_{kt} \\
&\quad - \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1, g_2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})} \right] \\
&\quad \times \left[(x_{k+1, g_2}) - \sum_{i=t}^k \frac{(x_{i+1, g_2}) \exp(\beta_0 + \beta_1 x_{i+1, g_2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1, g_2})} \right] V_{kt}.
\end{aligned}$$

Assim tem-se:

$$\bullet \sum_{t=1}^{s-1} \frac{\partial}{\partial \beta_0} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \frac{1}{\chi_{tg}} \frac{\partial}{\partial \beta_0} \chi_{tg}.$$

$$\bullet \sum_{t=1}^{s-1} \frac{\partial}{\partial \beta_1} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \frac{1}{\chi_{tg}} \frac{\partial}{\partial \beta_1} \chi_{tg}.$$

$$\begin{aligned}
\bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0^2} c_{tg} \log \chi_{tg} &= \sum_{t=1}^{s-1} c_{tg} \left[- \frac{\frac{\partial}{\partial \beta_0} \chi_{tg}}{\chi_{tg}^2} \frac{\partial}{\partial \beta_0} \chi_{tg} + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_0^2} \chi_{tg} \right] \\
&= \sum_{t=1}^{s-1} c_{tg} \left[- \frac{1}{\chi_{tg}^2} \left(\frac{\partial}{\partial \beta_0} \chi_{tg} \right)^2 + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_0^2} \chi_{tg} \right].
\end{aligned}$$

$$\begin{aligned}
\bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_1^2} c_{tg} \log \chi_{tg} &= \sum_{t=1}^{s-1} c_{tg} \left[- \frac{\frac{\partial}{\partial \beta_1} \chi_{tg}}{\chi_{tg}^2} \frac{\partial}{\partial \beta_1} \chi_{tg} + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_1^2} \chi_{tg} \right] \\
&= \sum_{t=1}^{s-1} c_{tg} \left[- \frac{1}{\chi_{tg}^2} \left(\frac{\partial}{\partial \beta_1} \chi_{tg} \right)^2 + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_1^2} \chi_{tg} \right].
\end{aligned}$$

$$\bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0 \beta_1} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[- \frac{\frac{\partial}{\partial \beta_1} \chi_{tg}}{\chi_{tg}^2} \frac{\partial}{\partial \beta_0} \chi_{tg} + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_0 \beta_1} \chi_{tg} \right].$$

5.3 Obtenção das derivadas parciais da log- condicional completa *a posteriori* de β para obtenção de \hat{H} para os modelos 5, 6, 7, e 8.

Procedendo de forma análoga a realizada para os modelos 1, 2, 3 e 4, obteremos as derivadas necessárias à avaliação de \hat{H} para os modelos 5, 6, 7 e 8. As derivadas primeiras de χ_{tg} com respeito a β para os referidos modelos são dadas pelas expressões que se seguem. Novamente consideramos V_{kt} como expresso em (5.2), (5.4), (5.6) e (5.8).

$$\bullet \frac{\partial}{\partial \beta_{0g}} \chi_{tg} = - \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})} \right] V_{kt},$$

e

$$\bullet \frac{\partial}{\partial \beta_{1g}} \chi_{tg} = - \sum_{k=t}^{s-1} \left[x_{k+1,g_2} - \sum_{i=t}^k \frac{x_{i+1,g_2} \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})} \right] V_{kt}.$$

As derivadas segundas de χ_{tg} são dadas pelas expressões a seguir:

$$\bullet \frac{\partial^2}{\partial \beta_{0g}^2} \chi_{tg} = \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{\exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{[1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})]^2} V_{kt} +$$

$$- \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})} \right]^2 V_{kt},$$

$$\bullet \frac{\partial^2}{\partial \beta_{1g}^2} \chi_{tg} = \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1,g_2})^2 \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{[1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})]^2} V_{kt} +$$

$$- \sum_{k=t}^{s-1} \left[x_{i+1,g_2} - \sum_{i=t}^k \frac{(x_{i+1,g_2}) \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})} \right]^2 V_{kt},$$

e

$$\begin{aligned}
\bullet \frac{\partial^2}{\partial \beta_{0g} \beta_{1g}} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1,g_2}) \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{[1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})]^2} V_{k_t} \\
&- \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})} \right] \\
&\times \left[(x_{k+1,g_2}) - \sum_{i=t}^k \frac{(x_{i+1,g_2}) \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})}{1 + \exp(\beta_{0g} + \beta_{1g} x_{i+1,g_2})} \right] V_{k_t}.
\end{aligned}$$

Destas formas, as derivadas parciais de segunda ordem, com respeito a β , da log-verossimilhança são expressas por:

$$\begin{aligned}
\bullet \frac{\partial^2 l}{\partial \beta_{0g}^2} &= \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_{0g}^2} c_{tg} \log \chi_{tg} - \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta_{0g} + \beta_{1g} x_{t+1,g_2})}{[1 + \exp(\beta_{0g} + \beta_{1g} x_{t+1,g_2})]^2} - 1, \\
\bullet \frac{\partial^2 l}{\partial \beta_{1g}^2} &= \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_{1g}^2} c_{tg} \log \chi_{tg} \\
&- \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g_2}^2 (\exp(\beta_{0g} + \beta_{1g} x_{t+1,g_2}))}{[1 + \exp(\beta_{0g} + \beta_{1g} x_{t+1,g_2})]^2} - 1,
\end{aligned}$$

e

$$\begin{aligned}
\bullet \frac{\partial^2 l}{\partial \beta_{0g} \partial \beta_{1g}} &= \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_{0g} \partial \beta_{1g}} c_{tg} \log \chi_{tg} \\
&- \sum_{t=1}^{s-1} x_{t+1,g_2} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta_{0g} + \beta_{1g} x_{t+1,g_2})}{[1 + \exp(\beta_{0g} + \beta_{1g} x_{t+1,g_2})]^2}.
\end{aligned}$$

Onde,

$$\bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_{0g}^2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\frac{1}{\chi_{tg}^2} \left(\frac{\partial}{\partial \beta_{0g}} \chi_{tg} \right)^2 + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_{0g}^2} \chi_{tg} \right],$$

$$\bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_{1g}^2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\frac{1}{\chi_{tg}^2} \left(\frac{\partial}{\partial \beta_{1g}} \chi_{tg} \right)^2 + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_{1g}^2} \chi_{tg} \right],$$

e

$$\bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_{0g} \beta_{1g}} c_{t,2} \log \chi_{t,2} = \sum_{t=1}^{s-1} c_{t,2} \left[-\frac{\frac{\partial}{\partial \beta_{12}} \chi_{t,2}}{\chi_{t,2}^2} \frac{\partial}{\partial \beta_{02}} \chi_{t,2} + \frac{1}{\chi_{t,2}} \frac{\partial^2}{\partial \beta_{02} \beta_{12}} \chi_{t,2} \right].$$

5.4 Obtenção das derivadas parciais da log- condicional *a posteriori* de β para obtenção de \hat{H} para os modelos 9, 10, 11 e 12.

Realizando procedimentos análogos aos anteriores obtemos as derivadas para os modelos 9, 10, 11 e 12, com V_{k_t} expresso como em (5.2), (5.4), (5.6) e (5.8). As derivadas com relação χ_{tg} são expressas por

$$\begin{aligned}
 \bullet \frac{\partial}{\partial \beta_0} \chi_{tg} &= - \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] V_{k_t}. \\
 \bullet \frac{\partial}{\partial \beta_1} \chi_{tg} &= - \sum_{k=t}^{s-1} \left[x_{k+1,g1} - \sum_{i=t}^k \frac{x_{i+1} \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] V_{k_t}. \\
 \bullet \frac{\partial}{\partial \beta_2} \chi_{tg} &= - \sum_{k=t}^{s-1} \left[x_{k+1,g2} - \sum_{i=t}^k \frac{x_{i+1} \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] V_{k_t}. \\
 \bullet \frac{\partial^2}{\partial \beta_0^2} \chi_{tg} &= - \sum_{k=t}^{s-1} \left\{ \frac{\partial^2 \log V_{k_t}}{\partial \beta_0^2} \exp(\log V_{k_t}) + \left[\frac{\partial \log V_{k_t}}{\partial \beta_0} \right]^2 \exp(\log V_{k_t}) \right\}. \\
 \frac{\partial^2}{\partial \beta_0^2} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})]^2} V_{k_t} + \\
 &\quad - \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right]^2 V_{k_t}.
 \end{aligned}$$

Analogamente,

$$\begin{aligned}
 \bullet \frac{\partial^2}{\partial \beta_1^2} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1,g1})^2 \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})]^2} V_{k_t} + \\
 &\quad - \sum_{k=t}^{s-1} \left[x_{i+1,g1} - \sum_{i=t}^k \frac{(x_{i+1,g1}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right]^2 V_{k_t}.
 \end{aligned}$$

$$\begin{aligned} \bullet \frac{\partial^2}{\partial \beta_2^2} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1,g2})^2 \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})]^2} V_{kt} + \\ &- \sum_{k=t}^{s-1} \left[x_{i+1,g2} - \sum_{i=t}^k \frac{(x_{i+1,g2}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right]^2 V_{kt}. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial \beta_0 \beta_1} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1,g1}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})]^2} V_{kt} \\ &- \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] \\ &\times \left[(x_{k+1,g1}) - \sum_{i=t}^k \frac{(x_{i+1,g2}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] V_{kt}. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial \beta_0 \beta_2} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1,g2}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})]^2} V_{kt} \\ &- \sum_{k=t}^{s-1} \left[1 - \sum_{i=t}^k \frac{\exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] \\ &\times \left[(x_{k+1,g2}) - \sum_{i=t}^k \frac{(x_{i+1,g2}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] V_{kt}. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial \beta_1 \beta_2} \chi_{tg} &= \sum_{k=t}^{s-1} \sum_{i=t}^k \frac{(x_{i+1,g1})(x_{i+1,g2}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{[1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})]^2} V_{kt} \\ &- \sum_{k=t}^{s-1} \left[x_{i+1,g1} - \sum_{i=t}^k \frac{x_{i+1,g1} \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] \\ &\times \left[(x_{k+1,g2}) - \sum_{i=t}^k \frac{(x_{i+1,g2}) \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})}{1 + \exp(\beta_0 + \beta_1 x_{i+1,g1} + \beta_2 x_{i+1,g2})} \right] V_{kt}. \end{aligned}$$

Assim, as derivadas parciais de segunda ordem de l são dadas por:

$$\bullet \frac{\partial^2 l}{\partial \beta_0^2} = \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0^2} c_{tg} \log \chi_{tg} - \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta' X_{t+1,g})}{[1 + \exp(\beta' X_{t+1,g})]^2} - 1.$$

$$\begin{aligned} \bullet \frac{\partial^2 l}{\partial \beta_1^2} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_1^2} c_{tg} \log \chi_{tg} \\ &- \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g1}^2(\exp(\beta' X_{t+1,g}))}{[1 + \exp(\beta' X_{t+1,g})]^2} - 1. \end{aligned}$$

$$\begin{aligned} \bullet \frac{\partial^2 l}{\partial \beta_2^2} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_2^2} c_{tg} \log \chi_{tg} \\ &- \sum_{g=1}^G \sum_{t=1}^{s-1} (z_{t+1,g} + m_{t+1,g}) \frac{x_{t+1,g2}^2(\exp(\beta' X_{t+1,g}))}{[1 + \exp(\beta' X_{t+1,g})]^2} - 1. \end{aligned}$$

$$\begin{aligned} \bullet \frac{\partial^2 l}{\partial \beta_0 \partial \beta_1} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0 \partial \beta_1} c_{tg} \log \chi_{tg} \\ &- \sum_{g=1}^G \sum_{t=1}^{s-1} x_{t+1,g1} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta' X_{t+1,g})}{[1 + \exp(\beta' X_{t+1,g})]^2}. \end{aligned}$$

$$\begin{aligned} \bullet \frac{\partial^2 l}{\partial \beta_0 \partial \beta_2} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0 \partial \beta_2} c_{tg} \log \chi_{tg} \\ &- \sum_{g=1}^G \sum_{t=1}^{s-1} x_{t+1,g2} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta' X_{t+1,g})}{[1 + \exp(\beta' X_{t+1,g})]^2}. \end{aligned}$$

$$\begin{aligned} \bullet \frac{\partial^2 l}{\partial \beta_1 \partial \beta_2} &= \sum_{g=1}^G \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_1 \partial \beta_2} c_{tg} \log \chi_{tg} \\ &- \sum_{g=1}^G \sum_{t=1}^{s-1} x_{t+1,g1} x_{t+1,g2} (z_{t+1,g} + m_{t+1,g}) \frac{\exp(\beta' X_{t+1,g})}{[1 + \exp(\beta' X_{t+1,g})]^2}, \end{aligned}$$

onde,

$$\bullet \sum_{t=1}^{s-1} \frac{\partial}{\partial \beta_0} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \frac{1}{\chi_{tg}} \frac{\partial}{\partial \beta_0} \chi_{tg}.$$

$$\bullet \sum_{t=1}^{s-1} \frac{\partial}{\partial \beta_1} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \frac{1}{\chi_{tg}} \frac{\partial}{\partial \beta_1} \chi_{tg}.$$

$$\begin{aligned}
& \bullet \sum_{t=1}^{s-1} \frac{\partial}{\partial \beta_2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \frac{1}{\chi_{tg}} \frac{\partial}{\partial \beta_2} \chi_{tg}. \\
& \bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0^2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\frac{1}{\chi_{tg}^2} \left(\frac{\partial}{\partial \beta_0} \chi_{tg} \right)^2 + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_0^2} \chi_{tg} \right]. \\
& \bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_1^2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\frac{1}{\chi_{tg}^2} \left(\frac{\partial}{\partial \beta_1} \chi_{tg} \right)^2 + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_1^2} \chi_{tg} \right]. \\
& \bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_2^2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\frac{1}{\chi_{tg}^2} \left(\frac{\partial}{\partial \beta_2} \chi_{tg} \right)^2 + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_2^2} \chi_{tg} \right]. \\
& \bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0 \beta_1} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\frac{\frac{\partial}{\partial \beta_1} \chi_{tg}}{\chi_{tg}^2} \frac{\partial}{\partial \beta_0} \chi_{tg} + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_0 \beta_1} \chi_{tg} \right]. \\
& \bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_0 \beta_2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\left(\frac{\partial}{\partial \beta_2} \chi_{tg} \frac{\partial}{\partial \beta_0} \chi_{tg} \right) \frac{1}{\chi_{tg}^2} + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_0 \beta_2} \chi_{tg} \right]. \\
& \bullet \sum_{t=1}^{s-1} \frac{\partial^2}{\partial \beta_1 \beta_2} c_{tg} \log \chi_{tg} = \sum_{t=1}^{s-1} c_{tg} \left[-\left(\frac{\partial}{\partial \beta_2} \chi_{tg} \frac{\partial}{\partial \beta_1} \chi_{tg} \right) \frac{1}{\chi_{tg}^2} + \frac{1}{\chi_{tg}} \frac{\partial^2}{\partial \beta_1 \beta_2} \chi_{tg} \right].
\end{aligned}$$

Em suma, a partir destas derivadas, pode-se avaliar \hat{H}^{-1} , necessária na descrição da função geradora de candidatos q descrita pela densidade

$$N(\hat{\beta}, \hat{H}^{-1}),$$

a ser utilizada no algoritmo de Metropolis-Hastings. Desta forma, a probabilidade de aceitação é

$$\alpha(\beta, \beta^*) = \min\{1, B\},$$

onde

$$B = \frac{\pi(\beta | D)}{\pi(\beta^* | D)} \times \frac{g(\beta^*)}{g(\beta)},$$

sendo $g(\beta)$ uma *priori* normal multivariada com vetor de médias $\mathbf{0}$ e uma matriz de covariâncias $\sigma^2 I$, onde σ^2 é considerado fixo e grande, de modo a se obter uma *priori* vaga.

No próximo capítulo descreveremos as funções geradoras de candidatos para as taxas de sobrevivência, ϕ , em cada um dos modelos.

Capítulo 6

Funções geradoras de candidatos para ϕ

Neste capítulo descreveremos as funções geradoras de candidatos para ϕ em cada um dos modelos descritos no Capítulo 4. Para isso, agruparemos tais modelos em 4 casos: Caso 1 - modelos 1, 5 e 9; Caso 2 - modelos 2, 6 e 10; Caso 3 - modelos 3, 7 e 11, e Caso 4 - modelos 4, 8 e 12.

6.1 Caso 1 - Modelos 1, 5 e 9:

Considere, como em Poole (2002), a função geradora de candidatos de ϕ , dada por

$$q(\phi^*|\phi) = \begin{cases} \frac{1}{2k\phi^*(1-\phi^*)}; & -k < \log\left(\frac{\phi^*}{1-\phi^*}\right) - \log\left(\frac{\phi}{1-\phi}\right) < k. \\ 0 & \text{c.c.} \end{cases} \quad (6.1)$$

Desta forma, a probabilidade de aceitação de ϕ^* é $\alpha(\phi, \phi^*) = \min\{1, B_1\}$, onde

$$B_1 = \frac{\pi(\phi^* | \cdot)}{\pi(\phi | \cdot)} \times \frac{q(\phi | \phi^*)}{q(\phi^* | \phi)}, \quad (6.2)$$

ou seja, considerando a expressão (4.6) para a distribuição condicional completa *a posteriori* de ϕ ,

$$B_1 = \frac{\left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg}^*)^{c_{tg}} \right] \times \phi^* \{ \sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) + \alpha \} \times (1 - \phi^*)^\gamma}{\left[\prod_{g=1}^G \prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \times \phi \{ \sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) + \alpha \} \times (1 - \phi)^\gamma}. \quad (6.3)$$

Seja $C_1 = \log B_1$, de forma que o cálculo da probabilidade de aceitação é facilitado através da operação $\alpha(\phi, \phi^*) = \min\{1, \exp(C_1)\}$, onde

$$\begin{aligned} C_1 &= \sum_{g=1}^G \sum_{t=1}^{s-1} (\log \chi_{tg}^* - \log \chi_{tg}) \\ &+ \left\{ \left[\sum_{g=1}^G \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) \right] + \alpha \right\} [\log \phi^* - \log \phi] \\ &+ \gamma [\log(1 - \phi^*) - \log(1 - \phi)]. \end{aligned} \quad (6.4)$$

6.2 Caso 2 - Modelos 2, 6 e 10

Considerando a função geradora de candidatos de $\phi_g, g = 1, \dots, G$ dada por

$$q(\phi_g^* | \phi_g) = \begin{cases} \frac{1}{2k\phi_g^*(1-\phi_g^*)}; & -k < \log\left(\frac{\phi_g^*}{1-\phi_g^*}\right) - \log\left(\frac{\phi_g}{1-\phi_g}\right) < k \\ 0 & \text{c.c.} \end{cases} \quad (6.5)$$

a probabilidade de aceitação de ϕ_g^* é $\alpha(\phi_g, \phi_g^*) = \min\{1, B_{2g}\}$, onde

$$B_{2g} = \frac{\pi(\phi_g^* | \cdot)}{\pi(\phi_g | \cdot)} \times \frac{q(\phi_g | \phi_g^*)}{q(\phi_g^* | \phi_g)}, \quad (6.6)$$

ou seja, considerando a expressão (4.10) para a distribuição condicional completa *a posteriori* de ϕ_g ,

$$B_{2g} = \frac{\left[\prod_{t=1}^{s-1} (\chi_{tg}^*)^{c_{tg}} \right] \times \phi_g^* \{ \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) + \alpha \} \times (1 - \phi_g^*)^\gamma}{\left[\prod_{t=1}^{s-1} (\chi_{tg})^{c_{tg}} \right] \times \phi_g \{ \sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) + \alpha \} \times (1 - \phi_g)^\gamma} \quad (6.7)$$

Desta forma, $\alpha(\phi_g, \phi_g^*) = \min\{1, \exp(C_{2g})\}$, onde

$$\begin{aligned} C_{2g} &= \sum_{t=1}^{s-1} (\log \chi_{tg}^* - \log \chi_{tg}) \\ &+ \left\{ \left[\sum_{t=1}^{s-1} \Delta_t(z_{t+1,g} + m_{t+1,g}) \right] + \alpha \right\} [\log \phi_g^* - \log \phi_g] \\ &+ \gamma [\log(1 - \phi_g^*) - \log(1 - \phi_g)], \quad g = 1, \dots, G. \end{aligned} \quad (6.8)$$

6.3 Caso 3 - Modelos 3, 7 e 11:

Considerando a função geradora de candidatos de ϕ_t , $t = 1, \dots, s-1$, dada por

$$q(\phi_t^* | \phi_t) = \begin{cases} \frac{1}{2k\phi_t^*(1-\phi_t^*)}; & -k < \log\left(\frac{\phi_t^*}{1-\phi_t^*}\right) - \log\left(\frac{\phi_t}{1-\phi_t}\right) < k. \\ 0 & \text{c.c.} \end{cases} \quad (6.9)$$

Desta forma, a probabilidade de aceitação de ϕ_t^* é $\alpha(\phi_t, \phi_t^*) = \min\{1, B_{3t}\}$, onde

$$B_{3t} = \frac{\pi(\phi_t^* | \cdot)}{\pi(\phi_t | \cdot)} \times \frac{q(\phi_t | \phi_t^*)}{q(\phi_t^* | \phi_t)}, \quad (6.10)$$

ou seja, considerando a expressão (4.14) para a distribuição condicional completa *a posteriori* de ϕ_t ,

$$B_{3t} = \frac{\left[\prod_{g=1}^G \prod_{i=1}^t (\chi_{ig}^*)^{c_{ig}} \right] \times \phi_t^* \{ \sum_{g=1}^G (z_{t+1,g} + m_{t+1,g}) + \alpha \} \times (1 - \phi_t^*)^\gamma}{\left[\prod_{g=1}^G \prod_{i=1}^{s-1} (\chi_{ig})^{c_{ig}} \right] \times \phi_t \{ [\sum_{g=1}^G (z_{t+1,g} + m_{t+1,g})] + \alpha \} \times (1 - \phi_t)^\gamma} \quad (6.11)$$

Desta forma, $\alpha(\phi_t, \phi_t^*) = \min\{1, \exp(C_{3t})\}$, onde

$$\begin{aligned} C_{3t} &= \sum_{g=1}^G \sum_{i=1}^t (\log \chi_{ig}^* - \log \chi_{ig}) \\ &+ \left\{ \left[\sum_{g=1}^G (z_{t+1,g} + m_{t+1,g}) \right] + \alpha \right\} [\log \phi_t^* - \log \phi_t] \\ &+ \gamma [\log(1 - \phi_t^*) - \log(1 - \phi_t)]; t = 1, \dots, s-1. \end{aligned} \quad (6.12)$$

6.4 Caso 4 - Modelos 4, 8 e 12:

Considerando a função geradora de candidatos de ϕ_{tg} , $t = 1, \dots, s-1$ e $g = 1, \dots, G$, dada por

$$q(\phi_{tg}^* | \phi_{tg}) = \begin{cases} \frac{1}{2k\phi_{tg}^*(1-\phi_{tg}^*)}; & -k < \log\left(\frac{\phi_{tg}^*}{1-\phi_{tg}^*}\right) - \log\left(\frac{\phi_{tg}}{1-\phi_{tg}}\right) < k. \\ 0 & \text{c.c.} \end{cases} \quad (6.13)$$

Desta forma, a probabilidade de aceitação de ϕ_{tg}^* é $\alpha(\phi_{tg}, \phi_{tg}^*) = \min\{1, B_{4tg}\}$, onde

$$B_{4tg} = \frac{\pi(\phi_{tg}^* | \cdot)}{\pi(\phi_{tg} | \cdot)} \times \frac{q(\phi_{tg} | \phi_{tg}^*)}{q(\phi_{tg}^* | \phi_{tg})}, \quad (6.14)$$

ou seja, considerando a expressão (4.19) para a distribuição condicional completa *a posteriori* de ϕ_{tg} ,

$$B_{4tg} = \frac{[\prod_{i=1}^t (\chi_{ig}^*)^{c_{ig}}] \times \phi_{tg}^* \{(z_{t+1,g} + m_{t+1,g}) + \alpha\} \times (1 - \phi_{tg}^*)^\gamma}{[\prod_{i=1}^t (\chi_{ig})^{c_{ig}}] \times \phi_{tg} \{(z_{t+1,g} + m_{t+1,g}) + \alpha\} \times (1 - \phi_{tg})^\gamma} \quad (6.15)$$

Desta forma, $\alpha(\phi_{tg}, \phi_{tg}^*) = \min\{1, \exp(C_{4tg})\}$, onde

$$\begin{aligned} C_{4tg} &= \sum_{i=1}^t (\log \chi_{ig}^* - \log \chi_{ig}) \\ &+ (z_{t+1,g} + m_{t+1,g}) + \alpha [\log \phi_{tg}^* - \log \phi_{tg}] \\ &+ \gamma [\log(1 - \phi_{tg}^*) - \log(1 - \phi_{tg})]; t = 1, \dots, s - 1. \end{aligned} \quad (6.16)$$

No próximo capítulo apresentamos os valores estimados dos parâmetros para cada um dos 12 modelos descritos no Capítulo 4, bem como a escolha do modelo que melhor descreve os dados.

Capítulo 7

Aplicação a dados reais

7.1 Descrição dos dados

As baleias bowheads têm sido fotografadas desde a metade dos anos 70. A maioria das fotografias se destina a estudos que têm por objetivo estimar tamanho e peso das baleias. Nos anos de 1981 a 1986 e em 1998 as baleias foram fotografadas durante sua migração de verão, no Mar de Beaufort (Alaska). Uma grande amostra foi coletada nas primaveras entre anos 1985 e 1994, exceto em 1988 e 1993.

Tabela 7.1: Tabela de capturas e recapturas - Baleias bowhead

Ano	1981	1982	1983	1984	1985	1986	1987	1989	1990	1991	1992	1994	1998
1981	24	0	0	0	4	1	0	3	1	0	0	0	0
1982		67	0	3	6	4	0	5	0	1	1	1	0
1983			5	0	0	0	0	0	0	0	1	0	0
1984				65	9	4	0	1	0	2	2	0	1
1985					259	20	3	5	3	4	4	3	3
1986						163	0	7	3	0	3	0	1
1987							22	0	1	1	0	0	0
1989								88	1	0	2	0	0
1990									58	1	0	0	0
1991										68	2	0	0
1992											61	1	0
1994												16	1
1998													18

Na diagonal da Tabela (7.1) descreve-se o número total de capturas de cada

ano. Nas outras células tem-se o número de baleias recapturadas pela primeira vez a cada ano. Por exemplo, dos 24 animais capturados em 1981, quatro foram recapturados pela primeira vez em 1985, um em 1986, três em 1989 e um em 1990. Podemos perceber que não ocorreram recapturas nos anos de 1982 e 1983, o que faz com que nesses anos os dados contenham pouca informação a respeito das probabilidades de captura. Portanto, Zeh *et al.* (2002) optaram por transferir o número de capturas de 1982 para 1981 e de 1984 para 1983, razão pela qual utiliza-se o termo Δ_t em expressões como a dada pela fórmula (4.3). Segundo os autores, esse procedimento não torna os dados viciados, uma vez que filhotes não fazem parte do conjunto de dados, ou seja as baleias capturadas em 1982 já estavam vivas em 1981 e aquelas capturadas em 1984 já existiam em 1983. Além disso, devido a baixa taxa de acúmulo de marcas naturais, essas baleias certamente já estavam marcadas no ano anterior ao da sua captura.

Os métodos utilizados por Zeh *et al.* (2002), para escolha de um conjunto de fotografias para análises de captura-recaptura produziram um conjunto de dados que incluem 2079 fotografias de 795 baleias, das quais apenas 102 (12.8 %) foram vistas em mais de uma ocasião amostral. Destas 102 baleias, 85 foram vistas em 2 anos, 15 em 3 anos e 1 em 4 anos.

Os dados de esforço amostral por classe de maturidade (altamente marcadas e moderadamente marcadas) e a matriz das histórias de captura, utilizados nesta dissertação, foram fornecidos, via e-mail, por Judy Zeh. Não nos foi dada autorização para fornecer mais detalhes sobre os mesmos. Os códigos de cada um dos 12 modelos foram escritos na linguagem Fortran. Na seleção dos modelos utilizou-se a fórmula (C.32), do apêndice C, para a verossimilhança marginal de cada modelo. Os diagnósticos de convergência foram feitos com o auxílio do *software* CODA. Na próxima seção passaremos a relatar alguns aspectos mais relevantes destas análises.

7.2 Diagnósticos de Convergência

A partir dos modelos discutidos no Capítulo 4, foi possível estimar, utilizando os códigos escritos na linguagem Fortran, as estimativas Bayesianas dos parâmetros relativos a cada modelo. Em todos os casos os resultados foram obtidos considerando-se uma cadeia com 20.000 observações das quais as 4.000 primeiras foram descartadas uma vez que correspondem ao período de *Burn-in*. Efetuou-se uma amostragem sistemática com período 10, de modo a se obter uma amostra pseudo-aleatória. Finalmente, considerou-se na análise Bayesiana, uma amostra de tamanho 1.600.

O CODA (Convergence Diagnostics and Output Analysis) é um conjunto de funções do *software* S-plus que pode ser utilizado para monitorar a convergência de processos de MCMC. Este *software* pode ser obtido gratuitamente através do site <http://www.mrc-bsu.ac.uk>.

Utilizando o CODA podemos realizar análises gráficas, como a da trajetória da cadeia e da aproximação da densidade *a posteriori*. Como exemplo, apresentamos as figuras 7.1 e 7.2, que foram obtidas para as 20.000 iterações dos parâmetros β_{01} e β_{11} (grupo 1), referentes ao modelo 5.

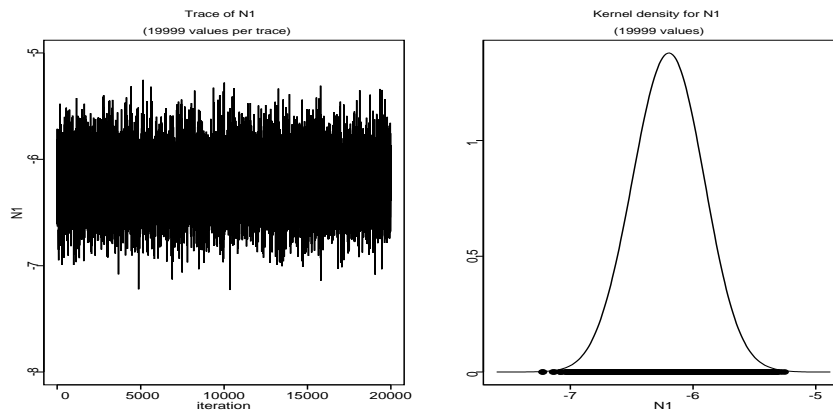


Figura 7.1: Trajetória da cadeia e densidade a posteriori - β_{01} .

Também podem ser obtidas estatísticas sumárias como média, desvio padrão e percentis, e diagnósticos de convergência. Nesta dissertação utilizamos um destes

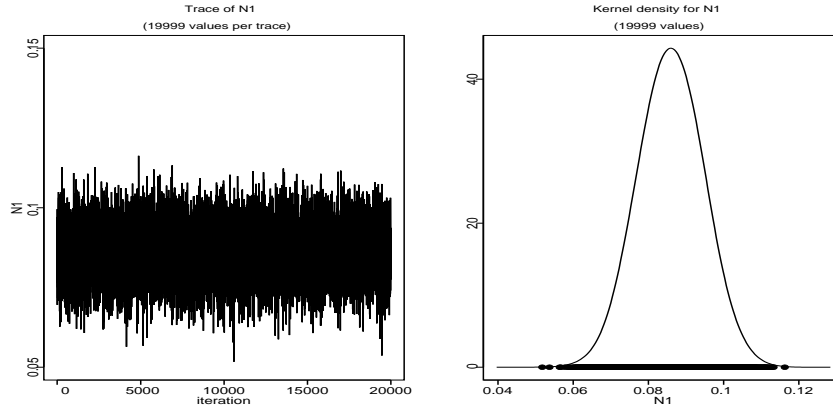


Figura 7.2: Trajetória da cadeia e densidade a posteriori - β_{11} .

diagnósticos, o critério de Gelman e Rubin (1992).

Gelman e Rubin propuseram um critério de convergência baseado em análise de variância. Neste critério utiliza-se 2 ou mais cadeias paralelas começando de diferentes pontos. Este método é baseado na comparação da variância *entre* as cadeias, E , e da variância *dentro* das cadeias, D .

Considere m cadeias paralelas, cada uma de tamanho n . Seja $\psi_{ij}, i = 1, \dots, n; j = 1, \dots, m$, a observação i da cadeia j . Desta forma, sejam as variâncias entre e dentro das cadeias dadas por, respectivamente,

$$E = \frac{n}{m-1} \sum_{j=1}^m (\bar{\psi}_{.j} - \bar{\psi})^2 \quad e \quad D = \frac{1}{n(m-1)} \sum_{j=1}^m \sum_{i=1}^n (\psi_{ij} - \bar{\psi}_{.j})^2,$$

onde $\bar{\psi}_{.j}$ é a média das observações da cadeia j , e $\bar{\psi}$ é a média dessas médias. A variância marginal *a posteriori* de ψ , $V(\psi|y)$, pode ser estimada através da média ponderada de D e E , isto é,

$$\hat{V}(\psi|y) = \frac{n-1}{n} D + \frac{1}{n} E. \quad (7.1)$$

Esta quantidade superestima $V[\psi|y]$ se a distribuição inicial for superdispersa, enquanto $\hat{V}(\psi|y)$ é não viesada sob estacionaridade da cadeia ($n \rightarrow \infty$).

Para n finito, a “variância dentro”, D , deve subestimar $V(\psi|y)$, uma vez que as sequências individuais (trajetórias) não foram percorridas tempo suficiente de

modo a permitir uma circulação adequada na gama de valores possíveis de ψ . Desta forma, D irá incorporar menor variabilidade do que $V(\psi|y)$. No limite, isto é, para $n \rightarrow \infty$, o valor esperado de D se aproxima de $V(\psi|y)$.

Monitora-se a convergência da cadeia em estudo através do chamado “fator de redução potencial de escala”, R , onde

$$\hat{R} = \sqrt{\frac{\hat{V}(\psi|y)}{D}}, \quad (7.2)$$

À medida que n cresce, \hat{R} declina para 1,0. Valores $\hat{R} \approx 1$ sugerem que o período de *burn-in* pode terminar e o processo iterativo continuar.

Com base no exposto, foi possível examinar a convergência das cadeias referentes aos 12 modelos. Para todos eles, os valores de \hat{R} ficaram muito próximos de 1, o que indica convergência. Para ilustrar as análises empreendidas, tomamos como exemplo o modelo 5 (vide seção 4.5). Na Tabela 7.2 descrevemos os percentis associados a cada parâmetro, o valor estimado de R e a taxa de aceitação, de cada parâmetro, nos esquemas de Metropolis-Hastings.

Tabela 7.2: Teste de Gelman e Rubin (CODA) - Modelo 5

Parâmetro	Percentis					\hat{R}	Taxa de aceitação
	2,5%	25%	50%	75%	97,5%		
β_{01}	-6,70	-6,38	-6,20	-6,03	-5,70	1	0,5126
β_{11}	0,0702	0,0805	0,0860	0,0913	0,1010	1	0,5126
β_{02}	-5,08	-4,90	-4,81	-4,71	-4,54	1	0,4958
β_{12}	0,0594	0,0654	0,0684	0,0714	0,0776	1	0,4958
ϕ	0,959	0,979	0,988	0,994	0,999	1,01	0,6568

Segundo Gelman *et al.* (2003), capítulo 11, a taxa de aceitação do algoritmo de Metropolis- Hastings, para cada parâmetro, deve estar por volta de 40%.

De acordo com os valores de \hat{R} e das taxas de aceitação, não há razão para duvidarmos da convergência de cada uma das cadeias simuladas para o modelo 5.

Como já mencionado, análises similares foram empreendidas para cada modelo e não observou-se anomalia no processo de convergência das cadeias.

7.3 Sensibilidade das inferências à escolha da *priori*

A fim de investigar uma possível sensibilidade das inferências à escolha da *priori* para ϕ , repetiu-se as simulações para cada um dos modelos utilizando-se uma *priori* $Beta(0, 5; 0, 5)$. Verificou-se que os resultados são ligeiramente afetados pela escolha da *priori*, como é possível observar comparando-se as tabelas 7.3 e 7.4, que apresentam estatísticas sumárias para os parâmetros do modelo 5. Os outros modelos apresentaram comportamento similar e, portanto, os resultados serão omitidos.

Tabela 7.3: Estimativas Modelos 5 - *Priori* $Beta(0, 5; 0, 5)$

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{01}	-6,233172	-6,22902700	0,24754410	(-6,727528;-5,755112)
β_{11}	0,086367	0,08631346	0,00771986	(0,071788;0,102111)
β_{02}	-4,842389	-4,84188700	0,13237720	(-5,108382;-4,578900)
β_{12}	0,068915	0,06887237	0,00457100	(0,059634;0,077907)
ϕ	0,999021	1,00000000	0,00510674	(0,987244;1,000000)

A Tabela 7.5 apresenta as probabilidades de captura obtidas com o modelo 5 utilizando-se *priori* $Beta(0, 5; 0, 5)$. Comparando os resultados da Tabela 7.5 com aqueles exibidos na Tabela 7.6, observa-se que as probabilidades de captura não são afetadas pela escolha da *priori*.

Na próxima seção apresentaremos os resultados obtidos para os 12 modelos.

Tabela 7.4: Estimativas Bayesianas dos parâmetros do Modelo 5 - Priori $Beta(1, 0; 1, 0)$

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{01}	-6,198187	-6,191242	0,255731	(-6,678162;-5,704580)
β_{11}	0,085770	0,085714	0,007865	(0,069395;0,100952)
β_{02}	-4,804657	-4,803448	0,132966	(-5,074073;-4,546703)
β_{12}	0,068389	0,068513	0,004491	(0,059092 0,077087)
ϕ	0,985296	0,987953	0,011191	(0,958538;0,999430)

Tabela 7.5: Probabilidades de captura Modelo 5 - *Priori Beta(0, 5; 0, 5)*

	Probabilidade de captura grupo 1	Probabilidade de captura grupo 2
p_2	0,010924	0,032447
p_3	0,102117	0,158121
p_4	0,042130	0,080888
p_5	0,004253	0,013505
p_6	0,009207	0,055001
p_7	0,006535	0,030352
p_8	0,014109	0,028388
p_9	0,007120	0,021697
p_{10}	0,003285	0,010285
p_{11}	0,002537	0,011787

Tabela 7.6: Probabilidades de captura para o modelo 5 - *Priori Beta(1, 0; 1, 0)*

	Probabilidade de captura grupo 1	Probabilidade de captura grupo 2
p_2	0,011176	0,002623
p_3	0,102753	0,159926
p_4	0,042677	0,082334
p_5	0,004380	0,013959
p_6	0,009431	0,056179
p_7	0,006710	0,031162
p_8	0,014408	0,029163
p_9	0,007307	0,022339
p_{10}	0,003390	0,010654
p_{11}	0,002623	0,012197

7.4 Estimativas dos parâmetros e seleção dos modelos

Os resultados das análise dos 12 modelos estão descritos nas Tabelas 7.7 a 7.18, que são apresentadas nas seções seguintes. Estas tabelas apresentam a média, a mediana, o desvio padrão e o intervalo de 95% de credibilidade para as estimativas dos parâmetros.

7.4.1 Modelo 1

Neste modelo foram estimados 3 parâmetros, a taxa de sobrevivência foi considerada constante para os dois grupos (indivíduos altamente marcados e moderadamente marcados) e para todas as ocasiões amostrais. As estatísticas sumárias relativas são apresentadas na Tabela 7.7. Os altos valores estimados de ϕ são razoáveis, uma vez que as baleias bowheads apresentam longevidade acima de 60 anos e são animais protegidos pelo governo americano, sendo muito pouco sujeitos à caça predatória.

Tabela 7.7: Estimativas Bayesianas dos parâmetros - Modelo 1

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,17046600	-5,16739400	0,11991350	(-5,42281700;-4,93598100)
β_1	0,07243979	0,07245659	0,00398639	(0,064461628;0,080312125)
ϕ	0,98421370	0,98675040	0,01133825	(0,95825710; 0,99925450)

7.4.2 Modelo 2

Neste modelo considera-se que as taxas de sobrevivência são variáveis por grupos. Nota-se que a taxa de sobrevivência é maior para indivíduos do grupo 2 (altamente marcados), conforme pode-se observar na Tabela 7.8. Os valores referentes ao vetor β , calculados para o modelo 2, foram similares àqueles apresentados para o modelo 1. Aparentemente, os indivíduos do grupo 2, isto é, os indivíduos altamente marcados, e, portanto, mais velhos, têm maiores taxas de sobrevivência do que os indivíduos moderadamente marcados.

Tabela 7.8: Estimativas Bayesianas dos parâmetros do Modelo 2

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,10093100	-5,10268200	0,12615770	(-5,34141500;-4,84794600)
β_1	0,07143174	0,07145116	0,00419152	(0,06315895;0,07965635)
Grupo 1	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
ϕ_1	0,87494760	0,87545650	0,03822634	(0,79881110;0,94815030)
Grupo 2	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
ϕ_2	0,99188040	0,99397340	0,00731062	(0,97285880;0,99976770)

Tabela 7.9: Estimativas Bayesianas dos parâmetros do Modelo 3

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,08470600	-5,08598700	0,12929780	(-5,34557400;-4,83683300)
β_1	0,07037174	0,07036075	0,00421894	(0,06216974;0,07916916)
ϕ_1	0,92329850	0,94085790	0,06689334	(0,74719240;0,99766250)
ϕ_2	0,93498370	0,94945590	0,05593689	(0,78686440;0,99787040)
ϕ_3	0,95458710	0,96517710	0,03978015	(0,84869300;0,99827680)
ϕ_4	0,93895940	0,95383350	0,05236441	(0,80191090;0,99793570)
ϕ_5	0,94180690	0,95568220	0,05005768	(0,81372650;0,99786870)
ϕ_6	0,90788410	0,92611580	0,07468619	(0,72769030;0,99690540)
ϕ_7	0,89565490	0,91691330	0,08509713	(0,68574780;0,99638920)
ϕ_8	0,91600090	0,93557730	0,07201350	(0,73779950;0,99759420)
ϕ_9	0,82823200	0,85340870	0,12887950	(0,53567340;0,99384680)
ϕ_{10}	0,81022120	0,84786200	0,14876790	(0,45508290;0,99318270)

7.4.3 Modelo 3

As estatísticas sumárias das estimativas dos parâmetros de interesse no modelo 3 são apresentadas na Tabela 7.9. Neste modelo considera-se que as taxas de sobrevivência mudam a cada ocasião amostral, porém são as mesmas para os dois grupos. A menor taxa de sobrevivência é observada na ocasião amostral 10. Os valores referentes ao parâmetro β foram similares aos anteriores. Aparentemente há diferenças entre as taxas de sobrevivência ao longo do tempo.

7.4.4 Modelo 4

As estatísticas sumárias referentes às estimativas do vetor β são similares às dos outros três modelos, como pode-se observar na Tabela 7.10. Entretanto, as taxas de sobrevivência são estimadas levando-se em conta diferenças entre as ocasiões amostrais e os grupos. A cada tempo, aparentemente, essas taxas são menores para

o grupo 1 (indivíduos mais jovens). Observa-se também que as menores taxas de sobrevivência estimadas ocorrem na 8ª e na 10ª ocasião amostral, para os grupos 1 e 2, respectivamente.

7.4.5 Modelo 5

Os resultados para este modelo são apresentados na Tabela 7.11. Observa-se que os vetores de coeficientes β são aparentemente distintos para cada grupo. Observa-se que β_0 obtido para o grupo 1 é inferior ao obtido para o grupo 2. O contrário ocorre com β_1 . Observa-se que os valores estimados do coeficientes de β relativos a este modelo são um pouco diferentes daqueles obtidos para os 4 primeiros modelos.

7.4.6 Modelo 6

A Tabela 7.12 apresenta os valores estimados dos parâmetros referentes a este modelo. As estimativas foram separadas por grupos. Com base neste modelo, indivíduos do grupo 1 aparentemente apresentam taxas de sobrevivência inferiores àquelas obtidas para o grupo 2. O valor de β_0 para o grupo 2 é aparentemente superior ao observado para o grupo 1. O contrário acontece com β_1 .

7.4.7 Modelo 7

Este modelo é similar aos 2 anteriores. Porém, as taxas de sobrevivência variam a cada ocasião amostral mas não variam por grupo. Os resultados estão sumarizados na Tabela 7.13. Os valores estimados de β são similares aos obtidos para os modelos 5 e 6. Como observado no modelo 3, a menor taxa de sobrevivência ocorre na 10ª ocasião amostral.

Tabela 7.10: Estimativas Bayesianas dos parâmetros do Modelo 4

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,003058	-5,002992	0,13041950	(-5,270505;-4,750452)
β_1	0,068974	0,068931	0,00422628	(0,060895;0,077467)
Grupo 1	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
ϕ_1	0,857843	0,883214	0,114445	(0,5704383;0,9960593)
ϕ_2	0,823230	0,843039	0,124955	(0,5336744;0,9935265)
ϕ_3	0,752273	0,762501	0,147685	(0,447929;0,984968)
ϕ_4	0,795671	0,823198	0,148770	(0,452200;0,992832)
ϕ_5	0,832040	0,862863	0,132758	(0,508260;0,995050)
ϕ_6	0,713303	0,736901	0,186693	(0,319821;0,983718)
ϕ_7	0,592590	0,597851	0,223635	(0,165965;0,973847)
ϕ_8	0,484928	0,460143	0,255283	(0,067438;0,961234)
ϕ_9	0,615664	0,634881	0,246993	(0,1312766;0,9824478)
ϕ_{10}	0,652598	0,695994	0,241576	(0,1311922 ; 0,9884815)
Grupo 2	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
ϕ_1	0,909255	0,930250	0,077173	(0,7112176 ; 0,9975390)
ϕ_2	0,942945	0,958222	0,052311	(0,8047747 ; 0,9980406)
ϕ_3	0,963109	0,973886	0,034981	(0,8728070 ; 0,9992582)
ϕ_4	0,950404	0,962732	0,0448721	(0,8343815 ; 0,9986646)
ϕ_5	0,950830	0,962786	0,043343	(0,8392145 ; 0,9982010)
ϕ_6	0,932906	0,949928	0,060353	(0,7757007 ; 0,9977344)
ϕ_7	0,931083	0,946918	0,059188	(0,7809858 ; 0,9980952)
ϕ_8	0,932586	0,950074	0,060669	(0,7738064 ; 0,9983968)
ϕ_9	0,848732	0,879047	0,121771	(0,5525324 ; 0,9957372)
ϕ_{10}	0,797620	0,831801	0,158013	(0,4384403 ; 0,9935435)

Tabela 7.11: Estimativas Bayesianas dos parâmetros do Modelo 5

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{01}	-6,198187	-6,191242	0,255731	(-6,678162;-5,704580)
β_{11}	0,085770	0,085714	0,007865	(0,069395;0,100952)
β_{02}	-4,804657	-4,803448	0,132966	(-5,074073;-4,546703)
β_{12}	0,068389	0,068513	0,004491	(0,059092 0,077087)
ϕ	0,985296	0,987953	0,011191	(0,958538;0,999430)

Tabela 7.12: Estimativas Bayesianas dos parâmetros do Modelo 6

Grupo 1	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{01}	-6,158736	-6,162416	0,274778	(-6.683573;-5.621523)
β_{11}	0,085297	0,085549	0,008328	(0,068603;0,101035)
ϕ_1	0,963515	0,97038080	0,028094	(0,895514;0,998530)
Grupo 2	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{02}	-4,804087	-4,804466	0,137446	(-5,063724;-4,530573)
β_{12}	0,068400	0,068454	0,004551	(0,059499;0,077510)
ϕ_2	0,983888	0,986370	0,028094	(0,9522576;0,9994939)

Tabela 7.13: Estimativas Bayesianas dos parâmetros do Modelo 7

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{01}	-6,106841	-6,102318	0,259610	(-6,616526 ; -5.587447)
β_{11}	0,083626	0,083575	0,008011	(0,068088 ; 0.100035)
β_{02}	-4,762350	-4,761167	0,144617	(-5.036450 ; -4.455168)
β_{12}	0,067741	0,067843	0,004719	(0,058035; 0,076543)
ϕ_1	0,924148	0,940411	0,065766	(0,751625 ; 0,997886)
ϕ_2	0,940341	0,954136	0,051249	(0,8117099 ; 0,998223)
ϕ_3	0,956812	0,968179	0,039093	(0,848439 ; 0,998632)
ϕ_4	0,936714	0,952256	0,054821	(0,799786 ; 0,998342)
ϕ_5	0,944939	0,957693	0,047488	(0,823834 ; 0,998416)
ϕ_6	0,908437	0,926195	0,074092	(0,721270 ; 0,997171)
ϕ_7	0,905473	0,925039	0,077452	(0,712944 ; 0,997299)
ϕ_8	0,921355	0,940834	0,067684	(0,749056 ; 0,997968)
ϕ_9	0,831405	0,859728	0,129428	(0,526985 ; 0,994922)
ϕ_{10}	0,815272	0,848314	0,146297	(0,436306 ; 0,993998)

7.4.8 Modelo 8

A Tabela 7.14 sumariza os resultados obtidos por meio do modelo 8, que estão separados por grupo. Os valores estimados de β se apresentam similarmente àqueles observados para os modelos 5, 6 e 7. As taxas de sobrevivência aparentemente são, novamente, inferiores para o grupo 1. Observa-se que, assim como no modelo 4, as menores taxas de sobrevivência estimadas são observadas na 8ª e na 10ª ocasião amostral, para os grupos 1 e 2, respectivamente.

7.4.9 Modelo 9

A Tabela 7.15 sumariza os resultados obtidos para este modelo, em que considera-se que as probabilidades de captura são expressas em função da ocasião amostral e do grupo, e as taxas de sobrevivência são constantes ao longo do tempo e grupos.

7.4.10 Modelo 10

Os resultados sumarizados na Tabela 7.16, referentes ao modelo 10, são similares aos obtidos para o modelo 9, exceto no que se refere à taxa de sobrevivência. Esta taxa é calculada levando-se em conta a presença de grupos. Conforme observado em modelos descritos anteriormente, a taxa de sobrevivência é aparentemente inferior para indivíduos do grupo 1.

7.4.11 Modelo 11

Os resultados obtidos para β , referentes ao modelo 11, são similares aos observados nos modelos 9 e 10, e estão sumarizados na Tabela 7.17. Com base nesta tabela pode-se perceber que a menor taxa de sobrevivência é aquela associada à 10ª ocasião amostral, como relatado para os modelos 3 e 7.

Tabela 7.14: Estimativas Bayesianas dos parâmetros do Modelo 8

Grupo 1	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{01}	-5,938024	-5,935936	0,285322	(-6,504622; -5,370239)
β_{11}	0,080830	0,080678	0,008542	(0,064562; 0,097625)
ϕ_1	0,879502	0,907643	0,103492	(0,622010 ; 0,996606)
ϕ_2	0,874543	0,900116	0,100201	(0,622200 ; 0,995844)
ϕ_3	0,855282	0,880128	0,114160	(0,579649 ; 0,995498)
ϕ_4	0,850371	0,877874	0,121682	(0,546887 ; 0,995217)
ϕ_5	0,863628	0,893382	0,115387	(0,572578 ; 0,996731)
ϕ_6	0,771437	0,804217	0,170461	(0,380419 ; 0,992597)
ϕ_7	0,676812	0,705694	0,215472	(0,222802 ; 0,987779)
ϕ_8	0,576501	0,584803	0,251593	(0,116955 ; 0,971852)
ϕ_9	0,631649	0,660403	0,246083	(0,113047 ; 0,980693)
ϕ_{10}	0,649785	0,687798	0,243222	(0,134033 ; 0,987939)
Grupo 2	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_{02}	-4,755953	-4,757072	0,144351	(-5,039323 ; -4,477774)
β_{12}	0,067691	0,067699	0,004810	(0,058338 ; 0,077162)
ϕ_1	0,891100	0,911587	0,088493	(0,675554 ; 0,997052)
ϕ_2	0,932259	0,947888	0,057259	(0,785785 ; 0,997693)
ϕ_3	0,956648	0,968487	0,038946	(0,855952 ; 0,998700)
ϕ_4	0,931325	0,948214	0,059030	(0,785976 ; 0,997704)
ϕ_5	0,936379	0,949492	0,053788	(0,798598 ; 0,997956)
ϕ_6	0,909508	0,929550	0,074810	(0,724510 ; 0,996825)
ϕ_7	0,912656	0,933407	0,076057	(0,711114 ; 0,997482)
ϕ_8	0,927291	0,943375	0,065734	(0,757765 ; 0,998005)
ϕ_9	0,833640	0,864226	0,130252	(0,525855 ; 0,994808)
ϕ_{10}	0,783206	0,814112	0,161486	(0,400558 ; 0,991757)

Tabela 7.15: Estimativas Bayesianas dos parâmetros do Modelo 9

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,851070	-5,853185	0,129228	(-6,094250 ; -5,601648)
β_1	0,072736	0,072713	0,004433	(0,063951; 0,081319)
β_2	0,913117	0,912903	0,070351	(0,7765803 ; 1,054496)
ϕ	0,986272	0,989057	0,011281	(0,958938 ; 0,999291)

Tabela 7.16: Estimativas Bayesianas dos parâmetros do Modelo 10

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,834293	-5,836321	0,124458	(-6,076948 ; -5,581363)
β_1	0,072412	0,072528	0,004399	(0,063484; 0,080914)
β_3	0,910190	0,909073	0,068105	(0,7715107 1,039896)
ϕ_1	0,949547	0,953557	0,028299	(0,8956260 ; 0,9986417)
ϕ_2	0,962707	0,963542	0,011294	(0,957726 ; 0,999472)

7.4.12 Modelo 12

Na Tabela 7.18 apresentamos as estatísticas sumárias referentes às estimativas dos parâmetros relativos ao modelo 12. As estimativas do vetor β são similares àquelas obtidas para os modelos 9, 10 e 11. Novamente, como descrito para os modelos 4 e 8, as menores taxas de sobrevivência são observadas nas ocasiões amostrais 8 e 10, para os grupos 1 e 2, respectivamente. As taxas de sobrevivência associadas ao grupo 1 são, aparentemente, ligeiramente menores do que as associadas ao grupo 2.

Tabela 7.17: Estimativas Bayesianas dos parâmetros do Modelo 11

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,812055	-5,819418	0,139283	(-6,075274; -5,530347)
β_1	0,071226	0,071426	0,004632	(0,061898 ; 0,080020)
β_3	0,978811	0,980267	0,067352	(0,848689; 1,112888)
ϕ_1	0,923052	0,940484	0,064417	(0,763554 ; 0,996952)
ϕ_2	0,941959	0,955989	0,051710	(0,797836 ; 0,997951)
ϕ_3	0,959222	0,970307	0,037081	(0,864743 ; 0,998905)
ϕ_4	0,933700	0,948236	0,056812	(0,786246 ; 0,998186)
ϕ_5	0,940368	0,954307	0,051556	(0,806924 ; 0,998604)
ϕ_6	0,907158	0,919975	0,072670	(0,723999 ; 0,996624)
ϕ_7	0,902974	0,921288	0,080007	(0,704254 ; 0,998055)
ϕ_8	0,920630	0,938389	0,069090	(0,735176 ; 0,997083)
ϕ_9	0,829046	0,856334	0,130277	(0,502070 ; 0,993308)
ϕ_{10}	0,802981	0,831988	0,147468	(0,446416 ; 0,993182)

Tabela 7.18: Estimativas Bayesianas dos parâmetros do Modelo 12

Parâmetro	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
β_0	-5,773796	-5,772348	0,133050	(-6,036100 ; -5,518706)
β_1	0,070765	0,070643	0,004500	(0,062011 ; 0,079384)
β_3	0,970382	0,971215	0,071804	(0,826608 ; 1,106594)
Grupo 1	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
ϕ_1	0,890429	0,916268	0,094432	(0,644663 ; 0,996745)
ϕ_2	0,896753	0,919620	0,088187	(0,657201 ; 0,996012)
ϕ_3	0,866753	0,892635	0,108881	(0,600612 ; 0,996557)
ϕ_4	0,857282	0,882601	0,116132	(0,568237 ; 0,995828)
ϕ_5	0,861056	0,889188	0,117174	(0,569637 ; 0,996287)
ϕ_6	0,768809	0,803340	0,170695	(0,373094 ; 0,991533)
ϕ_7	0,689196	0,712909	0,203524	(0,269035 ; 0,984531)
ϕ_8	0,561452	0,557847	0,256967	(0,107706 ; 0,979982)
ϕ_9	0,625239	0,649786	0,249605	(0,125115 ; 0,986818)
ϕ_{10}	0,644353	0,674399	0,238558	(0,153387 ; 0,982742)
Grupo 2	Média	Mediana	Desvio Padrão	Intervalo de credibilidade
ϕ_1	0,884864	0,903137	0,089390	(0,668495 ; 0,996486)
ϕ_2	0,926788	0,944622	0,063204	(0,774090 ; 0,997542)
ϕ_3	0,953136	0,963363	0,040952	(0,848896 ; 0,998334)
ϕ_4	0,928924	0,945333	0,060059	(0,778390 ; 0,997290)
ϕ_5	0,932039	0,947498	0,059782	(0,771032 ; 0,997420)
ϕ_6	0,915619	0,932339	0,070006	(0,738242 ; 0,997697)
ϕ_7	0,918155	0,937344	0,069499	(0,745036 ; 0,997001)
ϕ_8	0,929385	0,948437	0,064013	(0,761159 ; 0,998134)
ϕ_9	0,836734	0,865042	0,128331	(0,528436 ; 0,994533)
ϕ_{10}	0,783189	0,813751	0,162326	(0,405271 ; 0,993326)

Tabela 7.19: Estimador de média harmônica

Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
0,11154840	11,48341879	0,00009344	0,00037793	909,18559551	86,53675489
Modelo 7	Modelo 8	Modelo 9	Modelo 10	Modelo 11	Modelo 12
0,42680304	0,00023917	409,65528090	166,90604801	0,18647249	0,00816022

7.4.13 Seleção de modelos e conclusões

A fim de escolher o modelo que melhor descreve os dados, utilizou-se o fator de Bayes calculado a partir do *estimador de média harmônica* para a verossimilhança marginal de cada modelo, conforme descrito no Anexo C. A Tabela 7.19 apresenta os valores estimados das verossimilhanças marginais para cada um dos modelos.

Segundo os procedimentos descritos no anexo C, para a tomada de decisão a respeito do melhor modelo, basta obter a razão entre os *estimadores de média harmônica* das verossimilhanças marginais de cada modelo e compará-los. Procedendo desta forma obtém-se a Tabela 7.20. A partir dos valores apresentados na coluna referente ao modelo 5, (em negrito) desta tabela, verifica-se que, comparado com todos os outros modelos, o modelo 5 é o que fornece maior Fator de Bayes. Portanto, o modelo 5 é o que melhor descreve os dados.

A fim de comparar resultados Bayesianos com os frequentistas para a seleção de modelos utilizou-se o *software* MARK. O MARK é um programa que foi criado na Universidade do Colorado nos EEUU e é um programa útil na análise frequentista de dados de captura-recaptura, tanto de populações abertas quanto fechadas. A plataforma está disponível gratuitamente, no endereço eletrônico

<http://www.cnr.colostate.edu/~gwhite/mark.htm>.

Este *software* foi utilizado por Zeh *et al.* (2002) em suas análises frequentistas. Os resultados obtidos através do MARK estão na Tabela 7.21.

Comparando as estimativas para a verossimilhança verifica-se que, com exceção dos modelos 5 a 8, os valores são bem diferentes, o que indica que os modelos têm comportamentos diversificados no ajuste dos dados, sendo os modelos 5 e 6 os que maximizam a verossimilhança.

Os modelos 5 e 6 foram os que apresentaram menor AIC e, portanto, representam

Tabela 7.20: Fator de Bayes obtidos por meio dos estimadores de média harmônica

	Numerador			
Denominador	Modelo 1	Modelo 2	Modelo 3	Modelo 4
Modelo 1	1,00000000	102,94562053	0,00083768	0,00338803
Modelo 2	0,00971387	1,00000000	0,00000814	0,00003291
Modelo 3	1193,76935932	122893,32746633	1,00000000	4,04452739
Modelo 4	295,15670036	30385,08967254	0,24724768	1,00000000
Modelo 5	0,00012269	0,01263045	0,00000010	0,00000042
Modelo 6	0,00128903	0,13269990	0,00000108	0,00000437
Modelo 7	0,26135802	26,90566317	0,00021894	0,00088549
Modelo 8	466,39184808	48012,99821124	0,39068841	1,58014996
Modelo 9	0,00027230	0,02803191	0,00000023	0,00000092
Modelo 10	0,00066833	0,06880169	0,00000056	0,00000226
Modelo 11	0,59820297	61,58237619	0,00050110	0,00202673
Modelo 12	13,66977621	1407,24359488	0,01145094	0,04631362
Denominador	Modelo 5	Modelo 6	Modelo 7	Modelo 8
Modelo 1	8150,59321774131	775,77767505	3,82616923	0,00214412
Modelo 2	79,1737732569115	7,53580066	0,03716690	0,00002083
Modelo 3	9729928,44360636	926099,61811513	4567,56359021	2,55958453
Modelo 4	2405702,20012481	228975,97878002	1129,31948491	0,63285133
Modelo 5	1	0,09518052	0,00046943	0,00000026
Modelo 6	10,5063518581397	1,00000000	0,00493204	0,00000276
Modelo 7	2130,22287515676	202,75571425	1,00000000	0,00056038
Modelo 8	3801370,23376092	361816,38356380	1784,49413819	1,00000000
Modelo 9	2,21939185922355	0,21124286	0,00104186	0,00000058
Modelo 10	5,44728969604392	0,51847585	0,00255715	0,00000143
Modelo 11	4875,70908887689	464,07251106	2,28882581	0,00128262
Modelo 12	111416,785298473	10604,70720978	52,30287713	0,02930964
Denominador	Modelo 9	Modelo 10	Modelo 11	Modelo 12
Modelo 1	3672,44440583	1496,26578951	1,67167341	0,07315409
Modelo 2	35,67363417	14,53452592	0,01623841	0,00071061
Modelo 3	4384051,60547464	1786196,25291321	1995,59249050	87,32910771
Modelo 4	1083946,57307901	441632,87329329	493,40560650	21,59191897
Modelo 5	0,45057388	0,18357753	0,00020510	0,00000898
Modelo 6	4,73388772	1,92873015	0,00215484	0,00009430
Modelo 7	959,82278492	391,06105862	0,43690525	0,01911941
Modelo 8	1712798,13339985	697846,16678670	779,65484906	34,11846988
Modelo 9	1,00000000	0,40743048	0,00045519	0,00001992
Modelo 10	2,45440645	1,00000000	0,00111723	0,00004889
Modelo 11	2196,86715918	895,07064264	1,00000000	0,04376099
Modelo 12	50201,49318627	20453,61849938	22,85140136	1,00000000

Tabela 7.21: Modelos ajustados através do *software* Mark

	AIC	Delta AIC	AIC weight	Verossimilhaça	Par	Deviance
Modelo 1	1061,171	30,46	0	0	2	217,314
Modelo 2	1047,005	16,29	0,00008	0,0003	3	201,134
Modelo 3	1065,203	34,49	0	0	6	213,264
Modelo 4	1039,724	9,01	0,00304	0,011	7	185,754
Modelo 5	1030,714	0	0,27528	1	4	182,826
Modelo 6	1030,714	0	0,27528	1	4	182,826
Modelo 7	1032,167	1,45	0,13315	0,4837	6	180,229
Modelo 8	1032,085	1,37	0,13874	0,504	6	180,146
Modelo 9	1032,473	1,76	0,11423	0,415	3	186,603
Modelo 10	1034,304	3,59	0,04575	0,1662	4	186,415
Modelo 11	1040,61	9,9	0,00195	0,0071	8	184,604
Modelo 12	1036,9	6,19	0,01249	0,0454	8	180,894

melhor os dados. Entretanto, opta-se pelo modelo 5 por apresentar menor número de parâmetros, confirmando a conclusão obtida por meio do procedimento Bayesiano.

No caso Bayesiano, as probabilidades de captura e taxa de sobrevivência relativas ao modelo 5 selecionado estão descritas na Tabela 7.22. Podemos observar que, com execução da ocasião amostral 2, as probabilidades de captura são maiores para indivíduos do grupo 2. Isto acontece porque baleias mais marcadas são mais fáceis de serem capturadas.

O fato do modelo 5, que melhor descreve os dados observados das bowheads, prever alta taxa de sobrevivência, que é constante ao longo do tempo e igual para ambos os grupos, pode ser explicado pelo fato das bowheads serem animais de vida muito longa (por volta de 60 anos), terem baixa taxa de reprodução, de modo que a taxa de sobrevivência não se altera bruscamente entre ocasiões amostrais, além do fato destes animais não sofrerem pressões demográficas causadas pela caça

comercial. Caso as taxas de sobrevivência de cada grupo (moderadamente marcados e altamente marcados) sejam distintas, como sugeriram os valores estimados de ϕ para alguns modelos, esta diferença é pequena demais para justificar um modelo com um parâmetro para cada grupo.

Tabela 7.22: Probabilidades de captura e sobrevivência para o modelo 5 (Bayesiano)

	Probabilidade de captura grupo 1	Probabilidade de captura grupo 2
p_2	0,011176	0,002623
p_3	0,102753	0,159926
p_4	0,042677	0,082334
p_5	0,004380	0,013959
p_6	0,009431	0,056179
p_7	0,006710	0,031162
p_8	0,014408	0,029163
p_9	0,007307	0,022339
p_{10}	0,003390	0,010654
p_{11}	0,002623	0,012197
ϕ	0,985296	0,985296

Na análise frequentista do modelo 5 (vide Tabela 7.23) observa-se que as estimativas das probabilidades foram próximas das Bayesianas (vide tabela 7.22) e tanto mais próximas para o caso do grupo 2, cujos tamanhos amostrais são maiores, fornecendo melhores estimativas frequentistas.

No caso frequentista, a Tabela 7.23 apresenta um valor estimado de ϕ que iguala a unidade. Sabemos que a taxa de sobrevivência das bowheads é alta, mas certamente uma estimativa $\phi = 1,0$ é de pouco valor inferencial. O intervalo de 95% de confiança apresenta limite superior maior do que a unidade, incluindo valores que não pertencem ao espaço paramétrico. Desta forma, as estimativas Bayesianas, apesar da complexidade na obtenção, fornecem resultados mais confiáveis do que as frequentistas.

Tabela 7.23: Estimativas frequentistas para os parâmetros referentes ao modelo 5

Grupo	Parâmetro	Intervalos de confiança			
		Estimativa	Desvio Padrão	Limite inferior	Limite superior
1	p_2	0,009363	0,002415	0,005641	0,015502
1	p_3	0,101774	0,029929	0,056286	0,177122
1	p_4	0,036594	0,007644	0,024231	0,054909
1	p_5	0,002591	0,001111	0,001117	0,005996
1	p_6	0,020498	0,004143	0,013774	0,030404
1	p_7	0,008485	0,002280	0,005006	0,014348
1	p_8	0,007689	0,002154	0,004436	0,013295
1	p_9	0,005180	0,001710	0,002710	0,009880
1	p_{10}	0,001743	0,000852	0,000668	0,004537
1	p_{11}	0,002125	0,000974	0,000865	0,005213
2	p_2	0,031373	0,003333	0,025458	0,038606
2	p_3	0,154482	0,028047	0,107098	0,217719
2	p_4	0,082814	0,011044	0,063584	0,107193
2	p_5	0,015754	0,002426	0,011642	0,021286
2	p_6	0,027704	0,003090	0,022249	0,034448
2	p_7	0,021575	0,002747	0,016800	0,027669
2	p_8	0,037771	0,003860	0,030891	0,046110
2	p_9	0,022970	0,002821	0,018045	0,029199
2	p_{10}	0,013035	0,002247	0,009291	0,018259
2	p_{11}	0,010780	0,002070	0,007394	0,015692
	ϕ	1,000000	0,000011	0,999978	1,000022

Modelo 5 (Ajuste Bayesiano) - Densidade Estimada

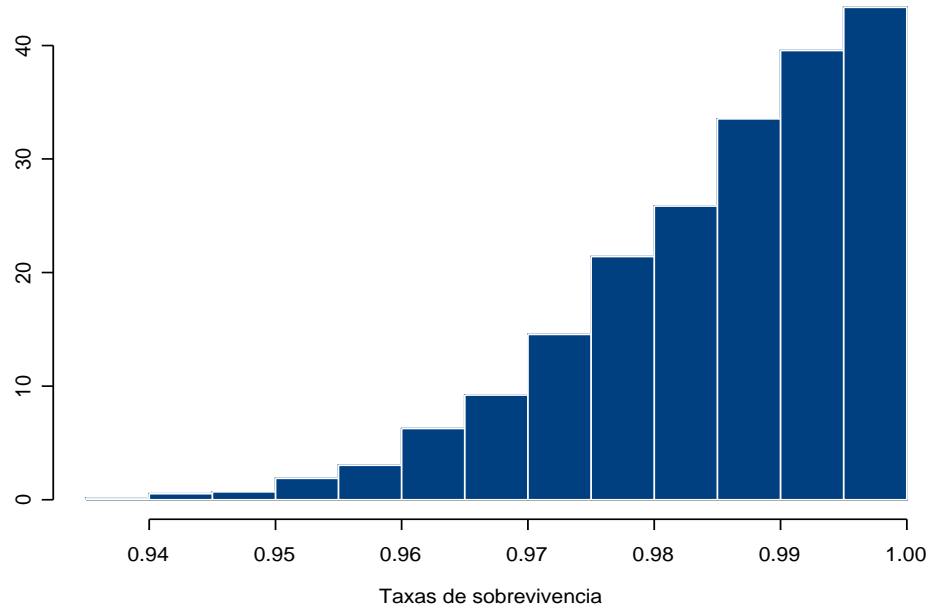


Figura 7.3: Densidade a posteriori da taxa de sobrevivência

A figura 7.3 descreve, para o modelo 5 (ajuste Bayesiano), a densidade estimada das taxas de sobrevivência ϕ , uma distribuição assimétrica e concentrada em valores muito altos.

Capítulo 8

Conclusões e considerações futuras

Nesta dissertação observamos que o uso da ligação logística para modelar probabilidades de captura, em modelos de captura-recaptura para populações abertas, apresenta grande flexibilidade na incorporação de efeitos devidos a covariáveis.

Para dados esparsos, como é o caso dos dados das baleias bowhead, a diminuição do número de parâmetros a estimar, proporcionado pelo uso de modelos logísticos, torna as estimativas dos parâmetros mais estáveis.

Neste trabalho um total de 12 modelos Bayesianos, considerando ligação logística para as probabilidades de captura, foram avaliados. Tais modelos incorporam a possibilidade de se levar em conta heterogeneidades nas probabilidades de captura devidas a presença de grupos distintos de indivíduos na população. No caso das baleias bowhead temos dois grupos de indivíduos, as baleias moderadamente marcadas e as altamente marcadas (em geral baleias mais velhas), sendo estas últimas mais propensas a serem recapturadas devido ao maior acúmulo de marcas naturais, facilitando a identificação dos indivíduos pelo biólogo que faz o processo de análise das fotos dos indivíduos. Neste estudo não considerou-se a possibilidade das probabilidades de captura serem constantes ao longo do tempo, uma vez que trabalhos anteriores, tais como os de Zeh *et al.* (2002), indicam que, para o caso das baleias bowhead, esta não seria uma possibilidade factível.

Verificamos que o uso de técnicas de simulação de Monte Carlo em Cadeias de Markov (MCMC) são apropriadas para a obtenção de sumários *a posteriori* de

interesse. Em contraste com as respectivas inferências obtidas com a análise frequentista dos modelos propostos, observamos que os modelos Bayesianos proporcionam valores estimados, em especial para taxas de sobrevivência, que são mais coerentes. Nos modelos frequentistas, ajustados com o auxílio do *software* MARK, em quase a totalidade dos valores estimados das taxas de sobrevivência, ϕ , obtivemos valores estimados $\hat{\phi} = 1, 0$, com intervalos de confiança apresentando limite superior maior que um. Em muitos casos, os limites inferior e superior dos intervalos de confiança foram iguais, igualando a unidade. O que se pode depreender de tais valores estimados para ϕ é que as taxas de sobrevivência das bowheads são, de fato, muito altas. Porém, definitivamente, não igualam a unidade. O fato dos intervalos de confiança frequentistas serem calculados com base em expressões assintóticas, envolvendo supostas grandes populações e amostras, faz com que tais estimativas sejam sujeitas a vários vícios para o caso de amostras tão diminutas como no caso das que dispomos para as baleias bowhead. A vantagem dos estimadores Bayesianos sobre os frequentistas é que eles não envolvem restrição alguma relativa ao caso de populações finitas e pequenas amostras. O procedimento amostral sugerido por Poole (2002) garante que as taxas de sobrevivência estimadas caiam dentro do espaço paramétrico de ϕ .

Neste trabalho encontramos que o modelo mais adequado para os dados das bowheads foi o modelo 5, que estabelece mesmas taxas de sobrevivência para ambos os grupos de indivíduos, sendo estas constantes ao longo do tempo, e probabilidades de captura modeladas segundo distintas ligações logísticas, envolvendo o esforço amostral para cada grupo de indivíduos. Estes resultados são consistentes com os encontrados por Zeh *et al.* (2002), que, sem considerar a modelagem Bayesiana logística para as probabilidades de captura, encontrou que o modelo $\{\phi, p_t\}$ (taxas de sobrevivência constantes e probabilidades de captura variando ao longo do tempo) foi o mais adequado para os dados das bowheads. Além disso, estes autores também estimaram as taxas de sobrevivência das bowheads como sendo muito altas, em torno de 0,987, enquanto a encontrada neste trabalho foi de 0,985.

Observou-se que as inferências sobre ϕ são ligeiramente afetadas pelo uso de distintas *prioris* vagas. Isto torna-se mais claro quando se comparam os valores médios

e desvios-padrão das estimativas. Para uma *priori* Beta(1,0;1,0), a média e o desvio-padrão dos valores estimados foram, respectivamente, 0,985296 e 0,011191, enquanto para a *priori* Beta(0,5;0,5), a média e o desvio-padrão dos valores estimados foram, respectivamente, 0,999021 e 0,005106. Observou-se também que as inferências sobre as probabilidades de captura, calculadas em função das transformações logísticas, não são afetadas pela escolha da *priori*. Este é um resultado importante, uma vez que há vários relatos na literatura apontando a sensibilidade das estimativas de probabilidades de captura à escolha da *priori* quando calculadas sem levar em conta a transformação logística (Zeh *et al.*, 2002, Brooks *et al.*, 2000).

Para trabalhos futuros temos a seguintes considerações:

(a) Os dados de esforço amostral, por categoria de indivíduo, de que dispomos no momento, ainda não são precisos. Novas análises precisarão ser empreendidas quando tivermos dados de melhor qualidade.

(b) A sensibilidade das inferências sobre ϕ à escolha da *priori* é uma questão que pode ser tratada utilizando-se *prioris* menos vagas, talvez com o uso de métodos Bayesianos empíricos. Outra possibilidade é também o ajuste das taxas de sobrevivência através de ligações logísticas.

(c) Pode-se tentar outras funções de ligação para as probabilidades de captura (e taxas de sobrevivência), tal como a ligação probito.

Referências Bibliográficas

- [1] Aitkin, M. (1991) Bayes Factor *J. R. Statist. Soc.*, **B**, 111-142
- [2] Akaike, H. (1973) Information theory and extension of the maximum likelihood principle. *Second International Symposium on Information Theory*. B. Petrov and F. Czakil (eds), 267-281.
- [3] Brooks, S. P.; Catchpole, E. A.; e Morgan, B. J. T. (2000). Bayesian animal survival estimation. *Statistical Science*, **15**, 357-376.
- [4] Coles, S. (1999). *Statistical Inference (Bayesian Inference)*.
- [5] da-Silva, C.Q.; Rodrigues, J.; Leite, J.G.; Milan, L.A. (2003). Bayesian estimation of the size of a closed population using photo-id data with part of the population uncatchable. *Communications In Statistics-Simulation And Computation*, **32**, n. 3.
- [6] Chapman, D.G. (1951). Some properties of the hypergeometric distribution with applications to zoological censuses. *Univ. Calif. Publ. Stat.* **1**, 131-160.
- [7] Chib S.; Greeberg, E. (1995). Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, **49**, n.4, 327-335.
- [8] Cormack, R. M. (1964) Estimates of survival from the sighting of marked animals. *Biometrika*, **51**, 429-438
- [9] Cormack, R. M. (1989) Loglinear models for capture-recapture. *Biometrics*, **45**, 395-413

- [10] Gamerman, D. (1996). Simulação Estocástica via Cadeias de Markov *Instituto de Matemática UFRJ*.
- [11] Geisser, S. e Eddy, W. F. (1979). A predictive approach to model selection. *Journal of the American Statistician Association*, **74**, 153-160.
- [12] Gelfand, A.E.; Dey, D. K.; Chang, H. (1992). Model determination using predictive distribution with implementation via sampling-based methods. *Bayesian Statistics*. Oxford University.
- [13] Gelfand, A.E., Sahu, S.K. and B.P. Carlin (1996): Efficient parametrization for generalized linear mixed models. In Bernardo, J.M., Berger, J.O., Dawid, A.P. and A.F.M. Smith (eds) *Bayesian Statistics*, **5**, 165-180, Clarendon Press, Oxford.
- [14] Gelman, A.; Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences *Statistical Science*, **7**, 457-511.
- [15] Gelman, A.; Carlin, J. B.; Stern, H. S.; Rubin, D. B. (2003) Bayesian Data Analysis. *Tests in Statistical Science Series*.
- [16] Gilks, W. R. e Roberts, G. O. (1996) Strategies for improving MCMC. *Markov Chain Monte Carlo Practice* (eds W. R. Gilks, S. Richardson e D. J. Spiegelhalter) 75-78.
- [17] Gray, R. (2002). Advanced Statistical Computing *Course Notes*.
- [18] Hosmer, D. W.; Lemeshow, S. (1989). Applied Logistic Regression *Wiley Series in Probability and Statistics*.
- [19] Jeffreys, H. (1935) Some tests of significance, treated by the theory of probability *Proc. Camb. Phil. Soc.*, **31**, 203-222.
- [20] Jeffreys, H. (1961) *Theory of Probability*, Terceira edição.
- [21] Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration stochastic model. *Biometrika*, **52**, 225-247.

- [22] Katz, R. W. (1981), On some criteria for estimating the order of Markov chain, *Technometrics*, **23**, 243-249.
- [23] Mardia, K.V.; Kent, J.T.; Bibby, J.M. (1979) *Multivariate Analysis* Academic Press.
- [24] Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A.; Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemistry and Physics* **21**, 1087-1091.
- [25] Müller, P. (1993). A generic approach to posterior integration and Gibbs sampling. *Journal of the American Statistician Association*
- [26] Paulino, C. D.; Turkman, M. A. A.; Murteira, B. (2003) *Estatística Bayesiana*, 351-359
- [27] Petersen, C. G. J. (1896). The yearly immigration of young plaice into Limfjord from the German sea, etc. *Report of the Danish Biological Station*, **6**, 1-48
- [28] Pollock, K. H.; Nichols, J. D.; Brownie, C.; Hines, J.E. (1990). Statistical Inference for Captur-Recapture Experiments *Wildlife Monographs* A Publication of The Wildlife Society .
- [29] Poole, D. (2002). Bayesian Estimation of Survival From Mark-Recapture Data. *American Statistica Association and the International Biometric Society - Journal of Agricultural, Biological, and Environment Statistics*, **7**, n. 2, 264-276.
- [30] Punt, A. E. e Butterworth, D. S. (1999) On assessment of the Bering-Chkchi-Beaufort Seas stock of bowhead whales (*Balaena mysticetus*) using a Bayesian approach. *Journal of Cetacean Research an Management*, **1**, 53-71.
- [31] Raftery, A. E. (1994) Bayesian Modelo Selection in Social Research. *Sociological Methodology*.
- [32] Rubin, D. B. (1984) Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Statist.*, **12**, 1151-1172.

- [33] Rugh, D. J.; Braham, H. W.; Miller, G. W. (1992) Methods for photographic identification of bowhead whales, *Balaena mysticetus*. *Canadian Journal of Zoology*, **70**, 617-624
- [34] Rugh, D. J.; Zeh, J. E.; Koski, W. R.; Baraff, L. S.; Miller, G. W.; Shelden, K. E. W. (1998). An improved system for scoring photo quality and whale identification in aerial photographs of bowhead whales. *Report of the international Whaling Commission*, **48**, 501-512.
- [35] Seber, G. A. F. (1965). A note on the multiple recapture census. *Biometrika*, **52**, 249-259.
- [36] Seber, G. A. F. (1970) Estimating time-specific survival and reporting rates for adult birds from band returns. *Biometrika*, **57**, 313-318
- [37] Seber, G. A. F. (1982). *The estimation of Animal Abundance and Related Parameters*. 2^a ed. New York: MacMillan.
- [38] Shibata, R. (1976). Selection of the order of an autorregressive model by Akaike's Information Criterion. *Biometrika*, **63**, 117-126.
- [39] Tierney, L.; Kadane, J. B. (1986) Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistician Association*, **81**, 82-86.
- [40] Whichee, B. J.; Zeh, J. E.; Rugh, D. J.; Koski, W. R.; Miller, G. W. (1996). Estimation of adult bowhead whale survival rates. Artigo SC/48/AS12, apresentado ao International Whaling Commission Scientific Committee, Maio 1996.
- [41] White, G. C.; Burnham, K. P.; (1999) Program MARK: Survival Estimation from populations of marked animals. *Bird Study*, **49** (supl.), 120-138.
- [42] Zeh, J.; Poole, D.; Miller, G.; Koski, W.; Baraff, L.; Rugh, D. (2002). Survival of Bowhead Whales, *Balaena mysticetus*, Estimated from 1981-1998 - Photoidentification Data. *Biometrics*, **58**, 832-840.

Apêndice A

Descrição do Algoritmo de Metropolis-Hastings

O algoritmo de Metropolis-Hastings nos dá um método geral para construir uma cadeia de Markov com distribuição estacionária dada por uma função de probabilidade arbitrária (ou uma densidade aproximada) $f(x)$ que, em geral, constitui nossa distribuição de interesse (distribuição alvo) cuja complexidade faz com que a geração de dados a partir desta distribuição não seja trivial. Por exemplo, as distribuições condicionais completas *a posteriori*, descritas no Capítulo 3,

$$\pi(\beta|\phi, D) \propto V_{\phi, \beta} \times L_{\beta} \times \pi(\beta)$$

e

$$\pi(\phi_{jg}|\phi_{1g}, \dots, \phi_{j-1g}, \dots, \phi_{s-1g}, \dots, \phi_{1G}, \dots, \phi_{tG}; \beta, D) \propto V_{\phi, \beta} \times L_{\phi} \times \pi(\phi),$$

nossas distribuições alvo, ou de equilíbrio, não têm forma conhecida. Portanto, na obtenção das distribuições empíricas associadas a elas, que constituem aproximações, faz-se necessário recorrer a métodos do tipo MCMC, como por exemplo, o algoritmo de Metropolis-Hastings. Tal algoritmo é baseado no método de amostragem por aceitação-rejeição que descreveremos a seguir.

A.1 Amostragem por Aceitação-Rejeição

De acordo com Chib e Greenberg (1995), o objetivo dos métodos de amostragem por aceitação-rejeição, AR, é gerar amostras a partir de uma *distribuição-alvo* absolutamente contínua $\pi(x) = f(x)/k$, onde $x \in \mathbf{R}^d$, $f(x)$ é uma densidade não normalizada, e k é a constante normalizadora (possivelmente desconhecida). Seja $h(x)$ uma densidade que pode ser simulada por algum método conhecido, e suponha que exista uma constante c , conhecida, tal que $f(x) \leq c.h(x)$ para todo x . Então, para gerar G_t a partir de $\pi(\cdot)$, considere os seguintes passos:

- (1) Gere um candidato z a partir de $h(\cdot)$ e um valor u a partir de uma distribuição $U(0, 1)$.
- (2) Se $f(z) \leq c.h(z)u$, aceite z e faça $G_t = z$. Caso contrário, rejeite z e volte para o passo (1).

Intuitivamente, o processo de geração/aceitação de candidatos para $\pi(x)$ a partir de $h(\cdot)$, pode ser ilustrado como a seguir.

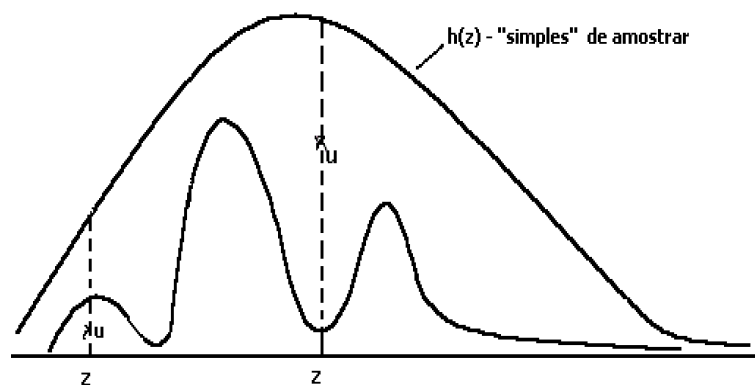


Figura A.1: Princípios do algoritmo de Metropolis-Hastings.

Segundo Chib e Greenberg (1995), o número esperado de iterações entre os passos (1) e (2), de modo a se obter uma aceitação, é dado por c^{-1} . O método de rejeição é otimizado ao fixarmos

$$c = \sup_x \frac{f(x)}{h(x)}.$$

No entanto, mesmo esta escolha de c pode resultar em um número muito grande de rejeições.

A.2 Princípios do Algoritmo M-H

A noção de uma distribuição geradora de candidatos também é utilizada no algoritmo de Metropolis-Hastings, M-H, no entanto, a grande diferença é que, no algoritmo de M-H, a função geradora de candidatos, denotada por $q(x, y)$, deve satisfazer a várias propriedades de cadeias de Markov, entre estas, a *reversibilidade*.

Para garantir a reversibilidade da cadeia de Markov a ser gerada, a *probabilidade de aceitação*, $\alpha(x, y)$, de um candidato y , partindo do “estado” x da cadeia, é tal que

$$\alpha(x, y) = \begin{cases} \min \left[\frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1 \right], & \text{se } \pi(y)q(y, x) > 0, \\ 1, & \text{caso contrário.} \end{cases}$$

O algoritmo de M-H é sumarizado através dos passos a seguir:

- (1) Para $j = 1, 2, \dots, N$.
- (2) Gere y a partir de $q(x^{(j)}, \cdot)$ e gere u de $U(0, 1)$.
- (3) Se $u \leq \alpha(x^{(j)}, y)$, aceite y e faça $x^{(j+1)} = y$. Caso contrário, rejeite y e faça $x^{(j+1)} = x^{(j)}$ e volte para (2).
- (4) A partir dos passos (1) a (3) obtenha a sequência de valores gerados $\{x^{(1)}, x^{(2)}, \dots, x^{(N)}\}$

Os valores gerados na sequência acima são considerados como provenientes da distribuição alvo, $\pi(x)$, somente após a cadeia ter passado um estágio transiente e o efeito dos pontos de partida (valores iniciais) terem se tornado tão pequenos que podem ser ignorados. Nessa situação, dizemos que a cadeia atingiu o *equilíbrio* ou que a cadeia *convergiu*.

Apêndice B

A escolha de $q(x, y)$

Quanto à questão da escolha de distribuições $q(x, y)$, apresentaremos as três formas mais comuns descritas na literatura.

B.1 Cadeia do tipo Passeio Aleatório

Uma família de funções geradoras de candidatos, que é descrita no trabalho de Metropolis *et al.* (1953), é dada por

$$q(x, y) = g(x - y),$$

onde $g(\cdot)$ é uma densidade multivariada (veja Müller, 1993). O candidato y é selecionado de acordo com o processo $y = x + z$, onde z é chamado de um *incremento aleatório*, que segue distribuição g . Como o valor-candidato y iguala ao valor atual adicionado de um *ruído*, o caso em questão é chamado de um *passeio aleatório*. Escolhas possíveis de $g(\cdot)$ incluem as distribuições normal multivariada e t-multivariada.

No caso da distribuição Normal-Multivariada, uma escolha pode ser $N(\mathbf{0}, \hat{H})$, onde $\hat{H} = \left\{ \frac{d^2}{dx dx'} \log[\pi(\hat{x})] \right\}^{-1}$, onde \hat{x} representa a moda de $\pi(x)$. A distribuição t-multivariada com matriz de dispersão \hat{H} também pode ser uma escolha apropriada.

Se g é simétrica, então

$$\alpha(x, y) = \min \left\{ \frac{\pi(y)g(x - y)}{\pi(x)g(y - x)}, 1 \right\} = \min \left\{ \frac{\pi(y)}{\pi(x)}, 1 \right\}.$$

B.2 Cadeia Independente

Seja g uma densidade definida no mesmo espaço de probabilidade de π e faça

$$q(x, y) = g(y).$$

Portanto, os valores candidatos são gerados independentemente do valor atual da cadeia. Novamente, escolhas eficientes de g envolvem as distribuições Normal e t-multivariada com vetor de médias \hat{x} e matriz de dispersão \hat{H} . Nesse caso,

$$\alpha(x, y) = \min \left\{ \frac{\pi(y)g(x)}{\pi(x)g(y)}, 1 \right\}.$$

B.3 Cadeia do tipo Amostragem por Rejeição

Uma terceira opção de descrição de g , que é eficiente, quando disponível, consiste em explorar a forma que se conhece de $\pi(\cdot)$ de modo a se especificar uma função geradora de candidatos (veja Chib e Greenberg, 1995). Por exemplo, se $\pi(t)$ pode ser escrita como

$$\pi(t) \propto \psi(t)h(t),$$

onde $h(t)$ é uma densidade que pode ser amostrada e $\psi(t)$ é uniformemente limitada, então faça

$$q(x, y) = h(x),$$

obtendo uma cadeia independente para gerar candidatos y . Neste caso, a probabilidade de aceitação requer somente o cálculo da função ψ , e é dada por:

$$\alpha(x, y) = \min \left\{ \frac{\psi(y)}{\psi(x)}, 1 \right\}.$$

Devido a importância do tema nesta dissertação, no próximo capítulo faremos uma breve revisão de literatura sobre comparação de modelos sob o paradigma Bayesiano.

Apêndice C

A perspectiva Bayesiana e o fator de Bayes na seleção de modelos

C.1 O Fator de Bayes

Um modelo é definido pela especificação de uma distribuição conjunta de variáveis aleatórias $Y = \{Y_1, Y_2, \dots, Y_n\}$, que são observáveis, e por quantidades não observáveis $\theta = \{\theta_1, \theta_2, \dots, \theta_k\}$, que representa o vetor de parâmetros associado ao modelo. Na perspectiva Bayesiana, enquanto as inferências são advindas da distribuição *a posteriori* $P(\theta | Y)$, a seleção de modelos não o é. A distribuição *a posteriori* $P(\theta | Y)$ não permite que se compare modelos, uma vez que a modelos distintos correspondem a distintos conjuntos de parâmetros. Para contornar tal problema na comparação de modelos, faz-se uso da densidade marginal $P(Y)$ que é obtida através da marginalização com respeito a θ da densidade conjunta $P(Y; \theta)$, isto é,

$$P(Y) = \int P(Y | \theta)P(\theta)d\theta. \quad (\text{C.1})$$

Se Y_{obs} denota o vetor de dados observados e $P(Y | M_i)$ denota a densidade marginal sob o modelo M_i , $i = 1, \dots, k$, a quantidade

$$B_{ji} = \frac{P(Y_{obs} | M_j)}{P(Y_{obs} | M_i)} \quad (\text{C.2})$$

Tabela C.1: Regras de interpretação de Jeffreys para o fator de Bayes

Intervalo de variação de B_{ji}	INTERPRETAÇÃO
$1 \leq B_{ji} \leq 3$	Fraca evidência em favor de M_j
$3 \leq B_{ji} \leq 10$	Moderada evidência em favor de M_j
$10 \leq B_{ji} \leq 100$	Forte evidência em favor de M_j
$B_{ji} \geq 100$	Decisiva evidência em favor de M_j

denota o *Fator de Bayes* calculado em função dos modelos i e j , que proporciona o peso relativo da evidência do modelo M_j comparada ao modelo M_i .

Com base nas expressões (C.1) e (C.2),

$$B_{ji} = \frac{\int P(Y_{obs} | M_j; \theta_j)P(\theta_j | M_j)d\theta_j}{\int P(Y_{obs} | M_i; \theta_i)P(\theta_i | M_i)d\theta_i}, \quad (C.3)$$

onde $P(Y_{obs} | M_j; \theta_j)$ representa a verossimilhança sob o modelo M_j , enquanto $P(\theta_j|M_j)$ representa a distribuição *a priori* dos parâmetros do modelo M_j .

Uma vez que a probabilidade *a posteriori* de que o modelo M_j seja uma boa descrição para os dados é dada por

$$P(M_j | Y_{obs}) = \frac{P(Y_{obs} | M_j)P(M_j)}{\sum_{l=1}^k P(Y_{obs} | M_l)P(M_l)}, \quad (C.4)$$

o *Fator de Bayes* B_{ji} pode ser representado por:

$$B_{ji} = \frac{P(Y_{obs} | M_j)}{P(Y_{obs} | M_i)} = \frac{P(M_j | Y_{obs})}{P(M_i | Y_{obs})} \times \frac{P(M_i)}{P(M_j)} \quad (C.5)$$

ou seja, o *Fator de Bayes* B_{ji} é dado pelo produto entre a razão entre as *posteriores* e a razão entre as *prioris* relativas aos modelos M_i e M_j .

C.1.1 Interpretação

O uso do Fator de Bayes para comparação de modelos foi proposto por Jeffreys(1935), e em Jeffreys (1961) ele propôs regras para interpretação de B_{ji} que estão sumarizadas na Tabela C.1.

C.1.2 Avaliação do *Fator de Bayes* e a aproximação BIC

A avaliação do *Fator de Bayes* envolve o cálculo da verossimilhança integrada dada pela expressão (C.1), que pode ser uma integral multidimensional e intratável.

Algumas aproximações analíticas e outras numéricas foram propostas, e nesta seção discutiremos a aproximação BIC, que tanto é simples quanto acurada. Nesta aproximação seguiremos os passos descritos por Raftery (1994).

Seja $Y (Y_1, Y_2, \dots, Y_n)$ um vetor aleatório, e considere novamente a expressão (C.1), isto é,

$$P(Y) = \int P(Y | \theta)P(\theta)d\theta.$$

A aproximação de $P(Y)$ consiste em se aproximar o integrando da expressão acima para uma modelo qualquer.

Seja

$$g(\theta) = \log\{P(Y | \theta)P(\theta)\}. \quad (\text{C.6})$$

A derivação consiste em se considerar a expansão em série de Taylor de $g(\theta)$ em torno de $\tilde{\theta}$, o valor de θ que maximiza $g(\theta)$, ou seja, a moda *a posteriori*. A expansão é dada por:

$$g(\theta) = g(\tilde{\theta}) + (\theta - \tilde{\theta})^t g'(\tilde{\theta}) + \frac{1}{2}(\theta - \tilde{\theta})^t g''(\tilde{\theta})(\theta - \tilde{\theta}) + o(\|\theta - \tilde{\theta}\|^2), \quad (\text{C.7})$$

onde

$$g'(\theta) = \left(\frac{\partial g(\theta)}{\partial \theta_1}, \dots, \frac{\partial g(\theta)}{\partial \theta_k} \right)^t \quad (\text{C.8})$$

é o vetor das primeiras derivadas parciais de $g(\theta)$ com respeito a θ e $g''(\theta)$ é a matriz Hessiana, isto é, das segundas derivadas parciais de $g(\theta)$ cujo elemento (i, j) é $\frac{\partial^2 g(\theta)}{\partial \theta_i \partial \theta_j}$.

Note que $g'(\tilde{\theta}) = 0$ já que $\tilde{\theta}$ é o valor que maximiza $g(\theta)$. Então,

$$g(\theta) \approx g(\tilde{\theta}) + \frac{1}{2}(\theta - \tilde{\theta})^t g''(\tilde{\theta})(\theta - \tilde{\theta}). \quad (\text{C.9})$$

Desta forma, para θ próximo a $\tilde{\theta}$, e com base em (C.9),

$$\begin{aligned} P(Y) &= \int \exp[g(\theta)]d\theta \\ &\approx \exp[g(\tilde{\theta})] \int \exp \left[\frac{1}{2}(\theta - \tilde{\theta})^t g''(\tilde{\theta})(\theta - \tilde{\theta}) \right] d\theta. \end{aligned} \quad (\text{C.10})$$

Como se pode perceber, o integrando na equação (C.10) é proporcional a uma densidade normal multivariada, ou seja

$$P(Y) \approx \exp[g(\tilde{\theta})](2\pi)^{k/2}|A|^{-\frac{1}{2}}, \quad (\text{C.11})$$

onde k é o número de parâmetros no modelo e $A = -g''(\tilde{\theta})$. A utilização da equação (C.11) é denominada *Método de integração de Laplace*.

Segundo Tierney e Kadane (1986), o erro na equação (C.11) é da ordem de $O(n^{-1})$. Desta forma, com base nas expressões (C.6) e (C.11),

$$\log P(Y) = \log P(Y | \tilde{\theta}) + \log P(\tilde{\theta}) + \frac{k}{2} \log(2\pi) - \frac{1}{2} \log |A| + O(n^{-1}), \quad (\text{C.12})$$

onde $O(n^{-1})$ representa qualquer quantidade tal que $nO(n^{-1})$ converge para uma constante quando n cresce.

Para grandes amostras, $\tilde{\theta} \approx \hat{\theta}$ onde $\hat{\theta}$ representa o estimador de máxima verossimilhança de θ , e $A \approx ni(\hat{\theta})$, onde i é a matriz de informação de Fisher esperada para uma única observação. Desta forma, $i(\hat{\theta})$ é uma matriz ($k \times k$) cujo elemento (i, j) é dado por

$$-E \left[\frac{\partial^2 \log P(Y_1 | \theta)}{\partial \theta_i \partial \theta_j} \right]_{\theta=\hat{\theta}}. \quad (\text{C.13})$$

Portanto, $|A| \approx n^k |i(\hat{\theta})|$. Estas aproximações introduzem um erro da ordem de $O(n^{-1/2})$ na equação (C.12), de modo que

$$\log P(Y) = \log P(Y | \hat{\theta}) + \log P(\hat{\theta}) + \frac{k}{2} \log(\pi) - \frac{k}{2} \log n - \frac{1}{2} \log |i(\hat{\theta})| + O(n^{-1/2}). \quad (\text{C.14})$$

O termo $\log P(Y | \hat{\theta})$ é da ordem de $O(n)$, enquanto o termo $\frac{k}{2} \log n$ é da ordem de $O(\log n)$. Os demais termos à direita na equação (C.14) são da ordem de $O(1)$ ou menos, de modo que podem ser excluídos, resultando em

$$\log P(Y) = \log p(Y | \hat{\theta}) - \frac{k}{2} \log n + O(1). \quad (\text{C.15})$$

A equação (C.15) implica que o logaritmo da verossimilhança integrada, $\log P(Y)$, é igual a log-verossimilhança avaliada em $\theta = \hat{\theta}$, $\log P(Y | \hat{\theta})$, menos um fator de correção devido ao tamanho da amostra.

O termo de erro $O(1)$ em (C.15) sugere uma aproximação ainda grosseira. No entanto, a aproximação de $\log P(Y)$ pode ser melhorada ao observarmos que

$$\log P(\hat{\theta}) = -\frac{k}{2} \log(\pi) + \frac{1}{2} \log |i(\hat{\theta})|. \quad (\text{C.16})$$

Substituindo-se (C.16) em (C.14) temos que

$$\log P(Y) = \log P(Y | \hat{\theta}) - \frac{k}{2} \log n + O(n^{-1/2}), \quad (\text{C.17})$$

proporcionando uma aproximação muito melhor para tamanhos n de amostras moderados.

Utilizando a expressão (C.17), o Fator de Bayes, expresso pela equação (C.2), isto é,

$$B_{ji} = \frac{P(Y_{obs} | M_j)}{P(D | M_j)}, \quad (\text{C.18})$$

é escrito, de forma conveniente, na escala de duas vezes o logaritmo de B_{ji} , isto é,

$$2 \ln B_{ji} = 2 \left(\log P(Y | \hat{\theta}_j, M_j) - \log P(Y | \hat{\theta}_i, M_i) \right) - (k_j - k_i) \log n + O(1). \quad (\text{C.19})$$

Denota-se por BIC (Bayesian Information Criterion) a quantidade

$$2 \log P(Y | \hat{\theta}, M) - k \log n, \quad (\text{C.20})$$

relativa a um modelo M qualquer.

Desta forma, na comparação entre os modelos M_j e M_i tem-se que a diferença entre os BIC 's, relativos a cada modelo, iguala a expressão (C.19), isto é,

$$\begin{aligned} \Delta BIC_{ji} &= 2 \log B_{ji} \\ &= 2 \left(\log P(Y | \hat{\theta}_j, M_j) - \log P(Y | \hat{\theta}_i, M_i) \right) - (k_j - k_i) \log n. \end{aligned} \quad (\text{C.21})$$

C.1.3 O critério AIC

O critério AIC (*Akaike Information Criterion*) para a seleção de modelos foi desenvolvido por Akaike (1973). Segundo Akaike, ao se comparar uma classe

de modelos para certo conjunto de dados devemos escolher aquele que minimiza a expressão

$$AIC = -2(\log p(D | \theta_i, M_i)) + 2d, \quad (C.22)$$

onde d é o número de parâmetros do modelo.

Embora este processo seja simples, Shibata(1976) e Katz (1981) mostraram que o AIC tende a superestimar o número de parâmetros, mesmo assintoticamente. Além disso, segundo Paulino *et al.* (2003), ΔAIC só aproxima $2 \ln B_{ji}$ se a informação contida na distribuição *a priori* aumentar na mesma proporção que a informação contida na verossimilhança. Como se sabe, esta situação não é realística do ponto de vista da metodologia Bayesiana e, portanto, não é razoável sua utilização neste contexto.

C.1.4 Alguns problemas com o Fator de Bayes na seleção de modelos

Segundo Gelfand (1996), a formulação do Fator de Bayes possui alguns problemas,

- (1) **Especificação da Priori** - Em muitos modelos, pelo menos parte da especificação *a priori* é vaga, de modo que $P(\theta)$ é imprópria. No entanto, mesmo quando a *posteriori*, resultante de uma *priori* imprópria, é própria, $P(Y)$ (vide equação (C.1)) é imprópria. Deste modo, como avaliar se $P(Y_{obs})$ é grande? Como calibrar o Fator de Bayes? A resposta é que isto não é possível, de modo que não é possível interpretar tais quantidades.
- (2) **Cálculo do Fator de Bayes** - Mesmo no caso de $P(Y)$ própria, esta trata-se, geralmente, de uma densidade multidimensional que não pode ser avaliada explicitamente, devendo ser calculada como uma marginalização com respeito a θ (vide equação (C.1)). A integração k -dimensional não pode ser feita de forma trivial, de modo que o uso de métodos MCMC faz-se necessário em seu cálculo. No entanto, as integrações aproximadas de Monte Carlo de (C.1)

tendem, em geral, a serem instáveis. Dessa forma, como $P(Y_{obs})$ não pode ser calculado de modo acurado, o mesmo ocorre com o Fator de Bayes, que é descrito pela razão de tais integrais.

Deste modo, apesar de amplamente difundido, o uso do Fator de Bayes é frequentemente inapropriado em situações reais.

Discutiremos a seguir procedimentos alternativos na seleção de modelos.

C.2 Distribuições Preditivas

A densidade $P(Y)$ pode ser denotada por *densidade preditiva a priori*. Isto é, com base na expressão (C.1), $P(Y)$ é representada pelo valor esperado de $P(Y | \theta)$ com respeito a $P(\theta)$. Analisaremos outras duas densidades preditivas, a *densidade preditiva via validação cruzada* e a *densidade preditiva a posteriori*.

C.2.1 Densidade Preditiva via validação cruzada e Pseudo Fator de Bayes

Suponha que y é um conjunto de observações $\{y_r, r = 1, 2, \dots, n\}$. Então as densidades preditivas via validação cruzada são descritas pelo conjunto

$$\{P(y_r | Y_{(r)}); r = 1, \dots, n\}$$

onde $Y_{(r)}$ denota todos os elementos de Y exceto y_r .

As densidades preditivas via validação cruzada são usualmete calculadas pela equação

$$P(y_r | Y_{(r)}) = \int P(y_r | \theta, Y_{(r)})P(\theta | Y_{(r)})d\theta. \quad (\text{C.23})$$

No caso das observações serem condicionalmente independentes dado θ ,

$$P(y_r | \theta, Y_{(r)}) = P(y_r | \theta).$$

Usualmente, $P(y_r | Y_{(r)})$ é uma densidade própria, mesmo quando $P(Y)$ é imprópria, de modo que a interpretabilidade dos resultados na comparação dos

modelos é frequentemente possível. A partir da expressão (C.23) pode-se observar que $P(y_r | Y_{(r)})$ é própria quando $P(y_r | \theta, Y_{(r)})$ é derivada de uma densidade $P(Y | \theta)$ própria (o que usualmente é o caso) e $P(\theta | Y_{(r)})$ é própria (o que ocorre quando $P(\theta | Y)$ é própria).

Observa-se que quando $P(Y)$ é própria, o conjunto $\{P(y_r | Y_{(r)})\}$ é equivalente a $P(Y)$ no sentido de que $\{P(y_r | Y_{(r)})\}$ contém a mesma informação com respeito a performance do modelo que $P(Y)$.

O produto das densidades preditivas via validação cruzada,

$$\prod_{r=1}^n P(y_r | Y_{(r),obs})$$

foi proposto como uma aproximação ou representação de $P(Y)$. Desta forma, define-se a primeira versão do *pseudo Fator de Bayes* através da expressão

$$PB_{ji}^{(1)} = \prod_{r=1}^n \frac{P(y_{r,obs} | Y_{(r),obs}, M_j)}{P(y_{r,obs} | Y_{(r),obs}, M_i)}. \quad (C.24)$$

Se este fator apresentar valor maior do que 1, prefere-se M_j a M_i . No caso de haver vários modelos competidores pode-se calcular as densidades preditivas para cada modelo e escolher o modelo que apresentar maior valor. Para mais detalhes vide Geisser e Eddy (1979), e Gelfand *et al.* (1992).

C.2.2 Aspectos computacionais do cálculo da densidade preditiva via validação cruzada

Para a densidade preditiva via validação cruzada podemos escrever

$$\begin{aligned} P(y_r | Y_{(r)}) &= \frac{P(Y)}{P(Y_{(r)})} = \frac{P(Y)}{\int P(Y_{(r)}; \theta) d\theta} \\ &= \frac{1}{\int \frac{1}{P(Y)} P(Y_{(r)}; \theta) d\theta} = \frac{1}{\int \frac{P(Y_{(r)}; \theta)}{P(Y, \theta)} P(\theta | Y) d\theta} \\ &= \frac{1}{\int \frac{1}{P(y_r | Y_{(r)}, \theta)} P(\theta | Y) d\theta} = \frac{1}{E_{\theta|Y} \left[\frac{1}{P(y_r | Y_{(r)}, \theta)} \right]} \end{aligned} \quad (C.25)$$

Através do uso de métodos de integração de Monte Carlo para expressão (C.25), temos que

$$\hat{P}(y_r | Y_{(r),obs}) = \frac{1}{\frac{1}{m} \sum_{j=1}^m \frac{1}{P(y_r | Y_{(r),obs}, \theta_j^*)}}, \quad (\text{C.26})$$

que é a média harmônica de $\{P(y_r | Y_{(r),obs}, \theta_j^*); j = 1, \dots, m\}$. No caso de $\{y_r\}$ condicionalmente independente dado θ ,

$$\hat{P}(y_r | Y_{(r),obs}) = \frac{1}{\frac{1}{m} \sum_{j=1}^m \frac{1}{P(y_r | \theta_j^*)}}, \quad (\text{C.27})$$

onde $\{\theta_j^*; j = 1, \dots, m\}$ é uma amostra obtida a partir de $P(\theta | Y_{obs})$.

Desta forma os elementos descritos por (C.26) ou (C.27) são substituídos na fórmula (C.24), na estimação do pseudo fator de Basyes.

C.2.3 Densidade preditiva a posteriori

A densidade preditiva *a posteriori*, $P(Y | Y_{obs})$, é a densidade preditiva de um novo conjunto independente de observações sob o modelo em estudo, dado o conjunto atual de observações (Rubin,1984; Aitkin,1991).

Similarmente a expressão (C.1), calculamos

$$P(Y | Y_{obs}) = \int P(Y | \theta)P(\theta | Y_{obs})d\theta = E_{\theta|Y_{obs}}[P(Y | \theta)], \quad (\text{C.28})$$

uma vez que $P(Y | \theta) = P(Y | \theta, Y_{obs})$.

Neste caso, $P(Y | Y_{obs})$ pode ser estimada, via integração de Monte Carlo, por

$$\hat{P}(Y | Y_{obs}) = \frac{1}{m} \sum_{j=1}^m P(Y | \theta_j^*). \quad (\text{C.29})$$

Desta forma define-se outra versão do pseudo fator de Bayes através de

$$PB_{ji}^{(2)} = \frac{P(Y | Y_{obs}, M_j)}{P(Y | Y_{obs}, M_i)}. \quad (\text{C.30})$$

C.3 Estimador de Média harmônica de $P(Y)$

Finalmente, considere $h(\theta)$ uma densidade própria, tal que

$$\begin{aligned} P^{-1}(Y) &= \int P^{-1}(Y)h(\theta)d\theta = \int \frac{h(\theta)}{P(Y|\theta)P(\theta)}P(\theta|Y)d\theta \\ &= \int \frac{P(\theta|Y)}{P(Y|\theta)}\frac{h(\theta)}{P(\theta)}d\theta. \end{aligned}$$

Para $h(\theta) = P(\theta)$ própria,

$$P^{-1}(Y) = \int P^{-1}(Y|\theta)P(\theta|Y)d\theta. \quad (\text{C.31})$$

Deste modo, $P(Y)$ pode ser estimado, via integração de Monte Carlo, por

$$\hat{P}(Y) = \frac{1}{\frac{1}{m} \sum_{j=1}^m \frac{1}{P(Y|\theta_j)}}, \quad (\text{C.32})$$

onde os θ_j 's são gerados a partir da distribuição *a posteriori* $P(\theta|Y)$.

A expressão (C.32) é denominada *estimador de média harmônica* de $P(Y)$. É fato conhecido que amostras ocasionais de θ podem ser altamente inconsistentes com respeito a $P(Y|\theta)$. Além disso, a soma de parcelas do tipo de $(L(Y|\theta_j))^{-1}$ pode dominar o estimador, tornando-o muito instável.

Diante das opções para a comparação de modelos descritas neste capítulo, e considerando que a verossimilhança de Cormack não é escrita como o produto de verossimilhanças individuais, inviabilizando o uso das expressões (C.21), (C.24) e (C.30), embora cientes das deficiências da fórmula (C.32), nesta dissertação faremos uso de tal expressão na comparação de modelos.