

Universidade Estadual de Campinas
FACULDADE DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO

Inferência e Indicadores de Viabilidade para Modelos Estocásticos de Crescimento de Populações

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas, como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, por

Selene Maria Coelho Loibel

Mestre em Matemática Computacional - ICMC/USP

em 18 de junho de 2004, sob a orientação do

Prof. Dr. João Bosco Ribeiro do Val

perante a banca examinadora constituída por

João Bosco Ribeiro do Val
Luciano Martins Verdade
José Galvão Leite
Akebo Yamakami
Wagner Caradori do Amaral

FEEC/UNICAMP
ESALQ/USP
DEs/UFSCAR
FEEC/UNICAMP
FEEC/UNICAMP

Agradecimentos

Ao professor João Bosco Ribeiro do Val, pela orientação para a elaboração deste trabalho e pela paciência e amizade.

Ao meu marido Marinho Gomes de Andrade Filho, pela co-orientação, motivação e apoio durante todo o período em que me dediquei a este doutorado.

Ao professor Marcelo Dutra Fragoso, pelo apoio no período de visita ao Laboratório Nacional de Computação Científica - LNCC no qual iniciei o levantamento bibliográfico para este trabalho.

Ao Dr. Alfredo Ribeiro Freitas, pesquisador da Embrapa Pecuária Sudeste por sua colaboração no estudo do crescimento do rebanho bovino brasileiro.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPQ, pela bolsa concedida durante os quatro anos de doutorado.

Aos meus pais Izette Alves Coelho Loibel e Gilberto Francisco Loibel e ao meu filho Diogo Loibel Sandonato, por todo incentivo para prosseguir estudando e aprimorando meu conhecimento.

A todos os colegas, professores e funcionários da Universidade Estadual de Campinas.

Resumo

Este trabalho estuda a aplicação do modelo de crescimento de Richards a populações de animais e ao cálculo de certos indicadores da viabilidade das populações. A probabilidade da população em estudo atingir um determinado número de indivíduos e o tempo esperado para que ela atinja este valor são os dois indicadores estudados. A aplicação da teoria assintótica e a abordagem clássica do problema de inferência dos parâmetros do modelo de Richards apresentam algumas dificuldades devido ao mal condicionamento da matriz de informação de Fisher associada ao modelo, o que resulta na identificabilidade pobre do parâmetro de forma. Nesta tese mostramos que o modelo de Richards pode ser visto como um modelo logístico quando aplica-se aos dados a transformação de potência de Box e Cox. Com esse recurso podemos reduzir o problema original à inferência de um modelo linear, para em seguida, aplicar o método de máxima verossimilhança *profile*, tornando assim viável a inferência dos parâmetros. A análise bayesiana completa o painel relativo à inferência, motivada pela situação de pequenas amostras de dados, situação encontrada freqüentemente no estudo de populações autóctones. Para populações sujeitas a manejo, introduz-se a estimação do modelo com abate proporcional ao tamanho da população, na forma de um processo estocástico ou como uma fração determinística. As técnicas desenvolvidas são testadas em séries sintéticas e faz-se uma aplicação à modelagem e a análise dos indicadores de viabilidade da série do rebanho bovino e histórico de abates, obtendo-se projeções do tamanho do rebanho brasileiro nos próximos anos.

A obtenção de estimativas dos parâmetros do modelo com boa precisão foi o principal objetivo da tese, uma vez que, o cálculo dos indicadores de viabilidade da população depende diretamente dessas estimativas.

Abstract

This work focuses on the application of the Richards growth model on animal populations and apply it to the calculus of certain population viability indicators. The probability of the population to reach a certain level and the expected time to reach it are the two indicators treated here. First we apply the classical theory to the parameter inference problem, but some difficulties arise due to the application of the asymptotic theory to a generally small data sample. Moreover, direct application of maximization methods yields poor identifiability of the shape parameter. In this thesis, we show that the Richards model can be interpreted as a logistic model, by applying the Box and Cox power transformation to the original data. This understanding allows us to reduce the original problem to a linear inference, and apply the profile likelihood technique. The Bayesian analysis completes the approach to the inference problem, motivated in part, by the small size of data related to native animal populations. For populations subject to management, this work introduces the model estimation with a harvesting that is proportional to the population size, as a deterministic fraction or as a stochastic process. The technique is tested on generated series, and on an application involving the Gross Brazilian cattle herd. This study provides projections to the future population size.

The main objective of this thesis was to obtain estimates of the Richards model parameters with good precision, which allows reliable calculus of population viability indicators.

Índice

Capítulo 1: Apresentação do problema de análise e controle do crescimento populacional.....	01
1.1 O problema de crescimento populacional.....	01
1.1.1 Caracterização de uma população.....	03
1.1.2 Aspéctos econômicos do problema de crescimento populacional..	04
1.2 Os modelos de crescimento.....	05
1.2.1 Modelos com taxas de crescimento dependentes da densidade.....	08
1.3. Estrutura da tese.....	13
Capítulo 2: Modelos estocásticos de crescimento populacional.....	15
2.1 Modelos estocásticos de crescimento para uma espécie.....	15
2.2 Modelo de Richards.....	16
2.3 Modelos logístico e de Gompertz.....	17
2.4 Comparação entre os modelos de Richards, logístico e de Gompertz.....	19
2.5 Modelos de crescimento com intervenção.....	21
2.6 Cálculo de indicadores de viabilidade de uma população.....	23
2.6.1 Cálculo de indicadores da viabilidade de uma população com intervenção estocástica.....	28
Capítulo 3: Inferência clássica com bootstrap para o modelo de Richards.....	31
3.1 Transformação de Box e Cox e o modelo de Richards.....	31
3.2 Função de verossimilhança <i>profile</i>	33
3.3 Cálculo de intervalos de confiança para os parâmetros ρ , K e τ	35
3.4 Técnicas <i>bootstrap</i>	36
3.4.1 <i>Bootstrap</i> paramétrico.....	37
3.4.2 Intervalos de confiança <i>bootstrap</i>	39
3.5 Escolha de modelos: Teste da razão de verossimilhança <i>profile</i>	42
3.6 Aplicação da inferência clássica.....	44
Capítulo 4: Abordagem bayesiana para o modelo de Richards.....	51
4.1 Discussão sobre o uso de métodos clássicos e bayesianos.....	51
4.2 Inferência bayesiana para os modelos logístico e de Gompertz.....	54
4.3 Inferência bayesiana para o modelo de Richards.....	56
4.3.1 Algoritmo amostrador de Gibbs em conjunto com o algoritmo Metropolis- Hastings.....	59

4.4 Seleção de modelos bayesiana: Critério das densidades preditivas ordenadas...	60
4.5 Aplicação da inferência bayesiana.....	61
Capítulo 5: Inferência para o modelo de crescimento populacional com intervenção e o cálculo de indicadores da viabilidade da população com modelo de Richards...	65
5.1 O modelo de Richards com intervenção estocástica.....	65
5.1.1 Inferência clássica para o modelo de Richards com intervenção estocástica.....	66
5.1.2 Inferência bayesiana para o modelo de Richards com intervenção estocástica.....	69
5.1.3 Cálculo de indicadores da viabilidade com o modelo de Richards com intervenção estocástica.....	71
5.2 O modelo de Richards com intervenção determinística.....	71
5.2.1 Inferência clássica para o modelo de Richards com intervenção determinística.....	72
5.2.2 Inferência bayesiana para o modelo de Richards com intervenção determinística.....	73
5.2.3 Cálculo de indicadores de viabilidade com o modelo de Richards intervenção determinística.....	73
Capítulo 6: Aplicação utilizando dados reais: Avaliação do crescimento do rebanho bovino brasileiro.....	77
6.1 Resultados da inferência clássica.....	79
6.2 Resultados da inferência bayesiana.....	88
Capítulo 7: Conclusões e propostas futuras.....	95
7.1 Conclusões.....	95
7.2 Propostas futuras.....	99
7.2.1 O problema de controle ótimo de uma população Formulação por programação dinâmica estocástica.....	99
7.2.2 Modelos heterocedásticos para crescimento populacional.....	101
7.2.3 Abordagem bayesiana com densidades a priori informativas.....	103
Referências Bibliográficas.....	105
Apêndice.....	111

Tabela de Símbolos

N_t	Processo estocástico que representa o tamanho da população no tempo t
W_t	Processo de Wiener
ϵ_t	Ruído branco do modelo estocástico de crescimento
h_t	Processo estocástico que representa a intervenção na população
ω_t	Ruído branco do modelo estocástico de intervenção
ρ, K, q, τ	Parâmetros do modelo de crescimento de Richards
ρ	Taxa intrínseca de crescimento da população
K	Capacidade de suporte da população
q	Parâmetro de forma do modelo de crescimento de Richards
τ	Inverso da variância do ruído ϵ_t
α_0, α_1	Parâmetros da reparametrização do modelo de crescimento de Richards
β_0, β_1	Parâmetros do modelo de Richards - bayesiano
ϕ_0, ϕ_1, η	Parâmetros do modelo de intervenção
η	Inverso da variância do ruído ω_t
\mathcal{S}	Espaço de estados da cadeia de Markov associada ao processo N_t
N_S	Número de intervalos do espaço de estados \mathcal{S}
I_i	Intervalo do espaço de estados \mathcal{S}
P_{ij}	Probabilidade de transição entre os estados contidos em I_i e I_j
\mathbf{P}	Matriz de probabilidades de transição de estados
T_{ij}	Tempo de primeira passagem de I_i para I_j
μ_{ij}	Valor esperado de T_{ij}
T_∞	Tempo máximo do simulação

Capítulo 1

Apresentação do problema de análise e controle do crescimento populacional

Nesta tese tratamos de modelos estocásticos de crescimento de populações animais. Este estudo consiste na escolha de um modelo adequado, que se ajuste aos conjuntos de observações do número de indivíduos de uma população animal ao longo do tempo, e na inferência dos parâmetros dos modelos. A modelagem é uma ferramenta importante para se compreender como as incertezas ambientais afetam o crescimento de uma população, vide por exemplo, Boyce, 1992 ; Ludwig, 1996 ; Hertzler et al.,1997 e Holmes, 2001. Tem-se como objetivo calcular indicadores da viabilidade da população na forma da probabilidade da população atingir um determinado nível e do tempo esperado para a população atingir esse nível. A obtenção de estimativas dos parâmetros dos modelos com boa precisão foi o principal objetivo da tese, uma vez que, o cálculo dos indicadores de viabilidade da população depende diretamente dessas estimativas.

A motivação para o estudo desses modelos é a preocupação, nos dias atuais, com o problema de extinção ou aumento excessivo de populações selvagens que tiveram seu habitat modificado e também o problema do ponto de vista econômico de exploração de recursos renováveis como pesca, caça, criação para abate etc.

Neste capítulo apresentamos o problema de crescimento populacional, os impactos ecológicos que afetam esse crescimento e os modelos matemáticos do crescimento de uma população. No final do capítulo, apresentamos a estrutura da tese.

1.1 O problema de crescimento populacional

Uma população é definida como um grupo de organismos, geralmente da mesma espécie, que ocupa uma área definida durante um período de tempo específico. A população, mais do que os indivíduos, é o foco primário do manejo da vida selvagem.

Enquanto os biólogos monitoram as condições do habitat das populações, eles também determinam cuidadosamente o tamanho, a vitalidade e a diversidade das populações para medir a "saúde" do ecossistema.

Uma das questões básicas da ecologia se refere ao tamanho das populações de algumas espécies que pode variar muito em curto espaço de tempo. Os pesquisadores dessa área se esforçam para melhorar a confiabilidade das respostas à questão; infelizmente, respostas claras são raramente disponíveis. Porém, elas tornam-se mais plausíveis quando são baseadas em conhecimento quantitativo das populações selvagens.

Considerando-se uma espécie como o resultado da história evolutiva de seu patrimônio genético em relação a seu ecossistema, sua conservação só faz sentido se inserida no contexto do meio ambiente em que habita, assegurando-se acima de tudo a funcionalidade desse ecossistema. Esta preocupação tem norteado a filosofia conservacionista mais recente. Nela, considera-se como principal objetivo de trabalho os processos ecológicos que determinem a distribuição e abundância dos seres vivos e, a partir daí, a funcionalidade de seus ecossistemas (apud Verdade, 2004).

Um ponto importante neste estudo é descobrir por que as populações animais flutuam, algumas vezes extensamente, crescendo até se tornarem pragas ou diminuindo até chegar ao limite da extinção - um problema pertinente para os pesquisadores que trabalham com a vida selvagem.

Existem muitos processos físicos, químicos e biológicos responsáveis pela dinâmica de uma população animal. Para que seja possível modelar matematicamente tal dinâmica, é preciso entender alguns conceitos ecológicos importantes tais como reprodução, predação, competição, mutualismo etc (Petrere, 1992).

Além da preocupação com a conservação das espécies selvagens que não têm interesse econômico, os pesquisadores também trabalham com espécies cuja exploração econômica deve ser controlada para manter os níveis de produção sem provocar a extinção.

Estudos detalhados sobre as populações, no entanto, requerem a coleta e análise de dados sobre fatores óbvios como disponibilidade e qualidade de alimento, populações de predadores, conhecimento sobre quais espécies competem entre si, ocorrência de doenças, influências climáticas e o impacto causado por ações humanas como caça, pesca ou atividades que destroem os habitats como desmatamento e mineração. Em poucos estudos de campo encontramos esses dados.

1.1.1 Caracterização de uma população

As populações têm características que os indivíduos não têm, que podem ser expressas matematicamente. Por exemplo, uma população pode ser caracterizada pela densidade demográfica, que é o número de indivíduos por unidade de área. Para entender a dinâmica de crescimento de uma população, estudamos suas taxas de natalidade e mortalidade, definidas como o número de nascimentos e de mortes por mil, cem ou por um indivíduo por ano.

Uma população tem estrutura de idade, que é a distribuição do número de indivíduos com várias idades. Naturalmente, a proporção de indivíduos na idade de procriação em uma população afeta a taxa de nascimento e influencia fortemente o crescimento. Da mesma forma, o número de indivíduos a partir de certa idade afeta a taxa de mortalidade. Além dessas taxas, a razão entre os sexos também influencia o potencial de reprodução.

As populações também têm fecundidade e fertilidade específicas da espécie. A primeira se refere, por exemplo, ao número de "ovos" produzidos por fêmea ou à quantidade de esperma produzida por macho. Como a quantidade de esperma raramente influencia a taxa de nascimento, a fecundidade geralmente está relacionada mais com a quantidade de ovos produzidos. A fertilidade é a porcentagem de ovos férteis, e a produção é o número efetivo de descendentes produzidos em determinado período de tempo. Alguns autores medem a produção somente como o número de novos indivíduos que atingem a idade de procriação, pois os animais que morrem entre o nascimento e a maturidade sexual não são contados. Mudanças em parâmetros como a razão dos sexos e a distribuição de idade influenciam muito a produção, isto é, o número de descendentes pode variar muito dependendo da proporção de fêmeas e de animais em idade de procriação na população em um dado momento (Bolen, Robinson, 1999).

A obtenção de informação precisa sobre o tamanho de uma população selvagem é uma das tarefas mais difíceis para os pesquisadores. Os dados são apresentados em forma de estimativas e, muitas vezes, apresentam censura*. Algumas espécies, particularmente aquelas que vivem em áreas abertas, são contadas a partir de fotos aéreas, outras exigem um trabalho de amostragem no campo.

censura*: observações incompletas obtidas em períodos irregulares de tempo

Uma contagem completa é raramente possível em áreas onde a vegetação ou a topografia escondem os animais ou onde a população é muito grande. Em tais casos, uma estimativa pode ser feita com base em uma amostragem por *quadrats*, que consiste na contagem e/ou pesagem dos organismos encontrados em quadrados de tamanho e quantidade adequados para fornecerem uma estimativa da densidade na área amostrada. Esses *quadrats* podem ser ninhos ou outro tipo de abrigo em determinado espaço. Com o conhecimento do número médio de *quadrats* pode-se então inferir o tamanho da população.

Uma maneira muito comum para obter uma amostra de animais é o método da razão captura e re-captura. Uma primeira fase deste método consiste na captura e marcação de animais. A segunda fase é soltá-los e mais tarde coletar outra amostra. O tamanho da população é calculado usando-se a proporção de animais marcados que estão na segunda amostra.

Quando existe migração em uma população, a obtenção de estimativas do tamanho da população fica muito mais complexa. Nesses casos, é preciso contar com a colaboração de pesquisadores que se encontram nas diversas regiões para onde os animais podem migrar. Os pesquisadores que trabalham com essas espécies devem estar sempre em contato para a troca de dados.

1.1.2 Aspectos econômicos do problema de população e os impactos ecológicos

Em diversos países, já existem programas de controle de pesca e caça de algumas espécies com o objetivo de conservar suas populações e manter uma produção sustentável, mas no Brasil ainda estamos longe de atingir um controle efetivo da exploração dos recursos naturais. Devido à extensão territorial de nosso país e aos níveis de pobreza das famílias que vivem em algumas áreas, é praticamente impossível implantar programas de controle que garantam seu sustento por muito tempo.

A legislação brasileira que normaliza o uso da fauna silvestre pode ser considerada extremamente conservadora se comparada a países como os Estados Unidos e a Venezuela. No Brasil, a proibição da caça impede o uso legal de espécies econômicas, tornando assim ilícita a renda gerada e necessário o investimento em fiscalização em lugar de monitoramento. Dessa forma, a parcela da população local que não possa ou não deseje abrir mão do uso desses recursos, é levada à ilegalidade. Por outro lado, a parcela da população local que pode optar por formas legais de geração de

renda, normalmente o faz através de alterações mais profundas no ambiente como, por exemplo, substituindo a floresta por pasto ou agricultura. É importante notar que existe uma diferença socio-econômica e cultural entre a caça praticada por esporte e aquela praticada para subsistência, como é o caso dos caboclos da Amazônia ou os sertanejos nordestinos. O uso sustentável não leva a população à extinção, os programas extensivos de uso sustentável, ao contrário, geram renda permitindo a conservação do ambiente onde as espécies vivem. Através da utilização econômica de uma espécie é possível proporcionar a conservação do ecossistema, ou de sua funcionalidade ecológica (apud Verdade, 2004).

A fauna brasileira é riquíssima e é por seu valor econômico que muitos dos nossos animais são abatidos, sem respeito a seus períodos de reprodução, expondo-se muitas espécies ao risco de extinção. Entre os mais caçados, estão as emas e araras para retirada das penas, as tartarugas para utilização da carne e do casco, os jacarés e as lontras para o uso da pele na confecção de casacos, sapatos e bolsas. Para se ter uma idéia da quantidade de animais abatidos, 2.500 pessoas foram detidas no pantanal do Mato Grosso, somente em 1991, por comercializarem peles do jacaré-do-Pantanal, uma espécie ameaçada de extinção (Costa, 2001).

Uma vez extinta uma espécie animal ou vegetal, não existe mais possibilidade de recuperá-la. Não conhecendo sequer o número de espécies animais e vegetais existentes no planeta, assim como a sua utilidade, o homem não deveria correr o risco de perdê-las definitivamente.

1.2 Os Modelos de Crescimento

A escolha de um modelo matemático adequado para o estudo da dinâmica de uma população animal é uma das maiores dificuldades dos pesquisadores dessa área. A questão é se devemos usar um modelo muito realista mas que limita a obtenção de resultados que podem ser relevantes ou usar modelos mais simples mas que têm maior facilidade analítica.

Os modelos de crescimento podem ter vários níveis de complexidade e, dependendo do objetivo da pesquisa, o modelo considerado pode ser intratável do ponto de vista da inferência. Alguns tipos de modelos que podem ser usados são os modelos

determinísticos, os modelos estocásticos, os modelos para uma só espécie, os modelos para duas ou mais espécies (envolvendo competição, mutualismo e predação). Os modelos podem ainda se apresentar com estados discretos ou estados contínuos, modelos com estrutura etária, modelos que contêm a razão dos sexos, modelos com distribuição espacial etc.

Se desejamos analisar a viabilidade da população através do cálculo da probabilidade de extinção ou do tempo esperado até extinção, é recomendável optar por modelos simples e que levem em conta os parâmetros mais importantes, que são a capacidade de suporte (nível no qual a população se estabiliza na ausência de fatores ambientais aleatórios) e a taxa intrínseca de crescimento da população. Um modelo estocástico para uma população pode aproximar adequadamente as propriedades estatísticas de flutuações no tamanho da população.

No caso em que o objetivo é o cálculo de políticas de controle da população, também é preciso ponderar a complexidade do modelo com as dificuldades que este pode apresentar para esses cálculos.

Assim como existem dificuldades na obtenção de estimativas confiáveis para o tamanho da população, também encontramos dificuldades para obter dados de competição, mutualismo, predação, comportamento territorial, migração etc. Nesta tese estudamos os modelos estocásticos discretos para apenas uma espécie sem estrutura etária, razão dos sexos e distribuição espacial.

As populações apresentam padrões característicos de aumento denominados formas de crescimento populacional. Para fins de comparação, podemos designar dois padrões básicos, segundo as formas das curvas de crescimento: a curva em forma de J, na qual a densidade aumenta rapidamente, de forma exponencial, parando abruptamente quando a resistência ambiental ou outro limite se torna efetivo, e a curva em forma de S ou sigmoideal, na qual a população aumenta lentamente no início, depois mais rapidamente e logo a taxa de aumento vai diminuindo aos poucos até o equilíbrio ser alcançado e mantido. Essas formas contrastantes de crescimento podem ser combinadas ou modificadas, ou ambas as coisas, de várias maneiras, dependendo das peculiaridades dos diversos organismos e dos ambientes (Odum, 1988).

O primeiro modelo matemático, cuja forma de crescimento é em forma de J, foi proposto por Malthus em 1798. Este é o modelo mais simples para o crescimento populacional, pois considera uma taxa de crescimento constante (não dependente do tamanho da população): $r = b - m$, onde b é a taxa de natalidade e m é a taxa de

mortalidade específicas da população estudada. Denotando-se por N_t o número de indivíduos na população no tempo t , a equação proposta por Malthus para o crescimento populacional é dada por:

$$\frac{d\ln(N_t)}{dt} = r \quad (1.1)$$

Se $r > 0$, a população cresce exponencialmente para o infinito e, se $r < 0$, tem-se a extinção da população.

Sem escassez de alimento e espaço, sem doenças, parasitas ou predadores, a taxa de nascimento é máxima, limitada somente pela fisiologia da reprodução da espécie, e a taxa de mortalidade é mínima, com mortes apenas ocorrendo em indivíduos a partir de certa idade. Se as mortes não compensam os nascimentos, o resultado é de crescimento infinito.

Sob condições ideais, as populações em um certo período podem apresentar taxas de crescimento exponencial. Tais condições podem ocorrer quando uma pequena população é introduzida em um novo ambiente favorável a ela. Entre os casos revistos por Dasmann, somente a população de cervo-de-cauda-branca introduzida na reserva George em Michigan nos EUA realmente apresentou um verdadeiro crescimento exponencial (apud Bolen, Robinson, 1999).

Essas condições ideais têm sido criadas em laboratório para culturas de leveduras e populações de ratos, que são colocados em locais suficientemente grandes para crescer e receber alimentos. Esse modelo com taxa exponencial é adequado para organismos como bactérias, leveduras, alguns insetos e possivelmente alguns pequenos mamíferos para os quais a reprodução e o crescimento podem ser considerados contínuos.

Qualquer população pode apresentar uma taxa de crescimento irrestrita se for reduzida a um nível muito baixo, mas obviamente nenhuma população pode crescer exponencialmente por muito tempo. O suprimento de alimento não pode ser suficiente para a demanda de uma população que cresce sempre. O espaço e as condições de abrigo disponíveis também são limitados, e predadores podem acabar com um grande número de indivíduos além do aumento na ocorrência de doenças. A taxa de natalidade pode cair e/ou a taxa de mortalidade crescer e, eventualmente, a população deve parar de crescer.

1.2.1 Modelos com Taxas de Crescimento Dependentes da Densidade

A discussão anterior leva-nos à conclusão que quanto maior a população, maior o efeito de desaceleração do crescimento. Tal efeito tem sido definido matematicamente e aplicado às equações de crescimento na forma de um parâmetro que representa o número máximo de indivíduos que o ambiente pode sustentar e é chamado de capacidade de suporte. Quando o tamanho da população se aproxima deste valor, então a sua taxa de crescimento diminui. Devemos notar que esse parâmetro pode mudar de tempos em tempos quando a oferta de alimentos, as condições de abrigo, a água disponível e outros fatores ambientais variam com as estações e em anos sucessivos. Fatores tais como comportamento territorial e as respostas à "lotação" podem interagir com esses fatores externos, tal que o crescimento da população pode lentamente diminuir.

Um fator que causa alta mortalidade ou redução na taxa de natalidade quando a população se torna mais densa é chamado de fator dependente da densidade, isto é, se a probabilidade de um indivíduo nascer ou sobreviver é menor quando o número de animais na população se torna maior, um fator dependente da densidade está agindo para restringir o crescimento da população. Tais fatores incluem suprimento de alimentos, predação, doenças e comportamento territorial. Existem poucos fatores independentes da densidade e estes são principalmente relacionados com o clima (frio, chuva, enchentes etc). É razoável supor, portanto, que a taxa de crescimento depende da densidade da população.

A. Uma técnica matemática muito útil para a análise de taxas de extinção é baseada na aproximação por difusão, que é precisa se a mudança no tamanho da população por período é pequena. Então o tamanho da população pode ser aproximado por um processo estocástico contínuo. Essa aproximação nos permite calcular o tempo esperado até a extinção e o risco de extinção, mas existem dúvidas sobre a sua acuidade em populações com tamanho pequeno e taxa de crescimento alta. Além disso, a solução, usando-se a aproximação por difusão, envolve o cálculo de equações diferenciais parciais, o que pode se tornar intratável se assumirmos que os parâmetros demográficos, como a taxa de crescimento e a capacidade de suporte, também são variáveis aleatórias.

A maioria dos animais estudados procria sazonalmente, produzindo uma quantidade anual de descendentes, ou seja ela cresce aos saltos. Somente se pudéssemos acompanhar a população durante 50 ou 100 anos, é que esses saltos seriam menos visíveis. O manejo de populações requer períodos de tempo muito menores, portanto os modelos de crescimento discretos são mais adequados que os contínuos.

Os modelos discretos com taxa de crescimento dependente da densidade da população podem ser obtidos assumindo a aproximação discreta do modelo de Malthus, substituindo a taxa constante r , por uma taxa que depende da densidade, denotada por $r(N_t)$. Consideramos a equação diferencial

$$\frac{d\ln(N_t)}{dt} = r(N_t) \quad (1.2)$$

Uma aproximação discreta para (1.2) pode ser encontrada escrevendo:

$$\int_{t_0}^{t_0+\Delta t} d\ln(N_t) = \int_{t_0}^{t_0+\Delta t} r(N_t)dt$$

$$\ln\left(\frac{N_{t_0+\Delta t}}{N_{t_0}}\right) \approx r(N_{t_0})\Delta t$$

Fazendo $\Delta t = 1$, a equação diferença análoga a (1.2) é dada por:

$$\ln\left(\frac{N_{t+1}}{N_t}\right) = r(N_t)$$

Portanto, a forma geral dos modelos discretos, com taxa de crescimento dependente da densidade é:

$$N_{t+1} = N_t \exp\{r(N_t)\} \quad c.i. \quad N_0 > 0 \quad (1.3)$$

Entre os modelos com taxa de crescimento dependente da densidade, o modelo mais usado para estudar a dinâmica de populações é o modelo logístico que foi desenvolvido pelo matemático belga Pierre Verhulst em 1838, foi utilizado extensamente por Lotka e "redescoberto" por Pearl e Reed em 1930 (apud Odum, 1988). No modelo logístico, a taxa de crescimento é dada por:

$$r(N_t) = \rho \left[1 - \frac{N_t}{K}\right] \quad (1.4)$$

Portanto o modelo logístico é escrito na forma (1.3) como:

$$N_{t+1} = N_t \exp\left\{\rho \left[1 - \frac{N_t}{K}\right]\right\} \quad c.i. \quad N_0 > 0 \quad (1.5)$$

Um outro modelo também usado é aquele baseado na lei de mortalidade de Gompertz. O matemático holandês, Benjamim Gompertz, mostrou em 1825 que a taxa

de mortalidade cresce em progressão geométrica. Se tivermos um gráfico das taxas de morte em escala logarítmica, obtemos uma reta, que é conhecida como função de Gompertz. Este modelo tem sido usado para crescimento individual de organismos (Kroll, Tornero, 1994) e também é muito usado para crescimento de tumores, nesse caso considera-se o crescimento do número de células cancerosas (Cahales et al, 2001). A aproximação discreta para o modelo de Gompertz é escrita na forma (1.3) como:

$$N_{t+1} = N_t \exp \left\{ -\rho \ln \left(\frac{N_t}{K} \right) \right\} \quad (1.6)$$

Nesses modelos, o parâmetro ρ representa a taxa intrínseca de crescimento da população e não depende da densidade, ou seja, ela é simplesmente a diferença entre as taxas de natalidade e mortalidade da população em questão: $\rho = d - m$.

O parâmetro K é o nível de saturação ou capacidade de suporte da população, que é o tamanho limite para a população, a partir do qual a taxa de crescimento torna-se decrescente. Esse nível é imposto basicamente pelo meio, por escassez de alimentos e espaço.

Os modelos logístico e de Gompertz apresentados em (1.5) e (1.6), respectivamente, são determinísticos e, nesse caso, sua forma de crescimento é sigmoideal. Na Figura 1.1a) apresentamos um gráfico simulado de uma população com modelo logístico determinístico, no qual destacamos a capacidade de suporte, que é o nível em que a população se estabilizaria na ausência de fatores ambientais aleatórios. Supondo agora que uma população seja explorada para fins econômicos, existe uma relação entre o estoque de recursos (ou tamanho da população) e a taxa sustentável de exploração desses recursos.

Populações muito pequenas só podem fornecer uma produção pequena e populações muito grandes também são improdutivas, devido a interações competitivas que reduzem o desempenho individual.

No gerenciamento de produção máxima de uma população, é desejável manter a população aproximadamente na metade da capacidade de suporte (K). Este é o ponto no qual a taxa de crescimento é máxima, isto é, em qualquer outro nível o número produzido seria menor. Este nível é chamado de produção máxima sustentável e está em destaque na Figura 1.1a). Maximizando a taxa de crescimento logística com relação a N_t , obtemos a produção máxima sustentável, como pode ser visto na Figura 1b).

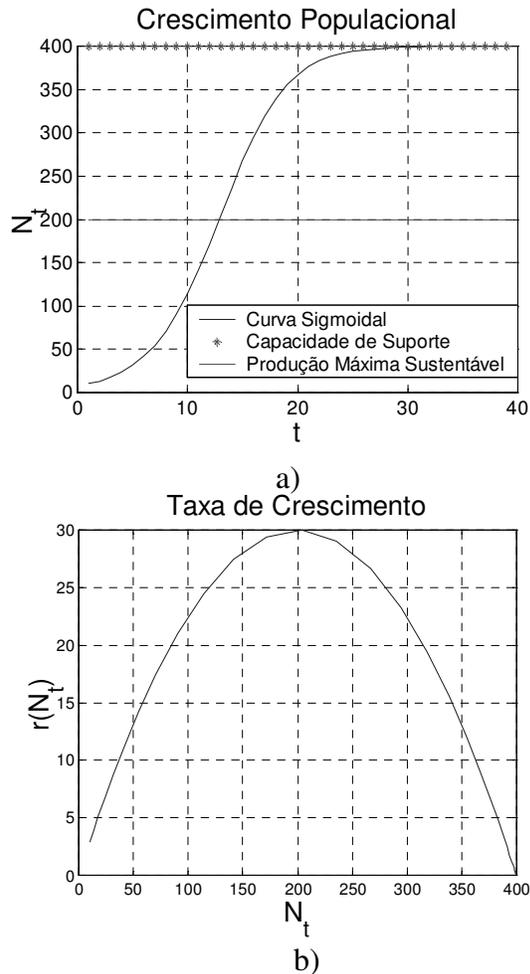


Figura 1.1: a) Crescimento logístico (curva sigmoidal) e b) Curva da taxa de crescimento

A primeira porção da curva sigmoidal dos modelos de crescimento pode servir como referência teórica com a qual os pesquisadores podem comparar a taxa de crescimento de populações observadas. Se o crescimento da população está próximo do ponto de inflexão da curva da Figura 1.1a), existe muito pouca coisa que pode ser feita para aumentar a taxa de crescimento desta população.

O modelo de Malthus (1.1) e os modelos logístico (1.5) e de Gompertz (1.6) podem ser melhor adequados para situações reais, já que, originalmente, não consideram a aleatoriedade dos processos naturais. Por exemplo, se uma enchente fatal ou um incêndio ocorre durante o período considerado, esta irregularidade na população não será explicada por esses modelos. A variabilidade devida a fatores ambientais (abióticos) deve ser introduzida no modelo por um processo estocástico e esse procedimento está apresentado, em detalhes, no capítulo 2.

O objetivo principal desta tese, além da escolha de um modelo adequado ao problema de crescimento de uma população, é propor técnicas de inferência que possibilitem boas estimativas dos parâmetros dos modelos e, conseqüentemente permitam cálculos precisos de indicadores da viabilidade de uma população animal.

A contribuição deste trabalho está no uso de um modelo, que é uma generalização dos modelos logístico e de Gompertz, denominado modelo de Richards. Consideramos o modelo de Richards estocástico e a tempo discreto e calculamos dois indicadores de viabilidade de uma população (probabilidade da população atingir uma meta em tempo determinado e o tempo esperado para a população atingir uma meta). O cálculo desses indicadores é feito com uma discretização do espaço de estados do processo de crescimento e considerando o modelamento por cadeia de Markov.

Do ponto de vista da estimação dos parâmetros do modelo de Richards consideramos a inferência clássica usando a semelhança entre a transformação de potência de Box e Cox para modelos lineares (Box e Cox, 1964) e a expressão da taxa de crescimento de Richards. Essa semelhança sugere o cálculo de estimadores de máxima verossimilhança *profile* (Barndorff-Nielsen e Cox, 1994; Murphy e Van der Vaart, 2000). Mostramos que o uso de técnicas clássicas associadas à técnica de *bootstrap* (Wehrens et al, 2000 ; Efron e Tibshirane, 1993) leva à estimativas pontuais e por intervalo razoáveis, mas estas são imprecisas quando a quantidade de observações do processo N_t é pequena. Este problema foi resolvido considerando a abordagem bayesiana para inferência e escolha de modelos de crescimento (Gelman et al, 1995 ; Box e Tiao, 1973). Os resultados bayesianos se mostraram mais robustos que os resultados clássicos com relação ao tamanho da amostra de N_t .

1.3 Estrutura da Tese

Nesta tese apresentamos, no capítulo 2, o modelo de crescimento estocástico de Richards e os seus casos particulares, além do modelo de crescimento com intervenção e os métodos para o cálculo de indicadores de viabilidade de uma população. No capítulo 3, destacamos a contribuição da tese na solução do problema de inferência clássica ao considerar a transformação de Box e Cox nos dados de população, contornando as dificuldades de cálculo de intervalos de confiança com as técnicas de *bootstrap*. Para escolha de modelos propomos o teste de razão de máxima verossimilhança *profile*, com a técnica bootstrap para a obtenção do valor crítico do teste. A dificuldade no uso da

teoria assintótica devida a pequenos tamanhos de amostra e dependência nas observações, foi contornada com a inferência bayesiana apresentada no capítulo 4. Neste capítulo também apresentamos o critério das preditivas ordenadas para seleção de modelos. No final dos capítulos 3 e 4, apresentamos a aplicação das técnicas propostas para dois conjuntos de dados gerados. No capítulo 5 introduzimos a inferência para o modelo de Richards com intervenção, calculamos dois indicadores da viabilidade de uma população usando o modelo de Richards. No capítulo 6, apresentamos uma aplicação contendo um conjunto de dados reais sobre o rebanho bovino brasileiro e fazemos avaliações e projeções sobre o tamanho deste rebanho para anos futuros. No capítulo 7, apresentamos as conclusões da tese e propostas para pesquisas futuras.

Capítulo 2

Modelos estocásticos de crescimento populacional

O crescimento de uma população inevitavelmente apresenta flutuações estocásticas provocadas por inúmeras causas imprevisíveis sendo a sobrevivência de espécies animais inerentemente incerta. A variabilidade devida a fatores ambientais (abióticos) pode ser introduzida no modelo por um processo estocástico. Além disso, é razoável supor que o crescimento depende do tamanho da população, ou seja, a taxa de crescimento não é simplesmente a diferença entre as taxas de natalidade e mortalidade, mas uma função do tamanho da população naquele período de tempo (Dennis et al, 1991).

Neste capítulo apresentamos alguns dos modelos estocásticos discretos para o crescimento de uma única espécie, sem considerar variáveis como estrutura etária, razão dos sexos e distribuição espacial.

2.1 Modelos estocásticos de crescimento para uma espécie

Para considerar os efeitos ambientais estocásticos, a forma geral dos modelos discretos com taxa de crescimento dependente da densidade (1.3) deve ser modificada adicionando à $r(N_t)$ um processo estocástico com distribuição conhecida.

As flutuações externas podem ser modeladas por um ruído colorido (processo correlacionado), o qual pode ser aproximado por um ruído branco (processo não correlacionado, com média zero) que denotamos por $\sigma\epsilon_t$, onde $\sigma > 0$ é o desvio padrão do processo e ϵ_t é um ruído branco padrão com variância unitária. Formalmente, ϵ_t é a derivada, no sentido de funções generalizadas do processo padrão de Wiener. A aproximação é feita por conveniência matemática e leva a uma equação diferencial estocástica de Stratonovich como uma aproximação razoável para o modelo de crescimento (Arnold, 1974). Assim, a forma geral dos modelos discretos com taxa de crescimento que depende do tamanho da população e que considera o efeito aleatório das condições ambientais pode ser escrita na seguinte forma:

$$N_{t+1} = N_t \exp\{r(N_t) + \epsilon_t\} \quad (2.1)$$

Assumindo-se que ϵ_t é um ruído branco com $E(\epsilon_t) = 0$, $E(\epsilon_t^2) = \tau^{-1}$ ($\tau = 1/\sigma^2 > 0$) e $E(\epsilon_t \epsilon_{t+k}) = 0$, $k = 1, 2, 3, \dots$

A equação (2.1) é a forma geral dos modelos de crescimento populacional estocásticos para apenas uma população, ou seja este tipo de modelo não considera predação, mutualismo e competição. No entanto, alguns indicadores da viabilidade da população são calculados considerando esse modelo mais simples e esses resultados são importantes para os ambientalistas e na exploração econômica dos animais.

2.2 Modelo de Richards

O modelo de Richards inicialmente era usado para estudar o crescimento em organismos individuais (por exemplo, peso \times idade) e só posteriormente começou a ser usado para crescimento populacional. Esse modelo é uma generalização de outros modelos como o logístico, o Gompertz, o Brody e o Bertalanffy. É interessante notar que todos esses modelos podem ser usados tanto para o crescimento dos indivíduos como para o crescimento das populações, mas não se sabe se existe uma relação de causa e efeito entre um caso e o outro. Atualmente, pode-se dizer apenas que existem algumas semelhanças nos padrões. Populações e comunidades não são "superorganismos", mas podem possuir propriedades análogas. Os modelos de Brody e de Bertalanffy são mais usados para o crescimento de organismos individuais e, portanto, não foram tratados nesta tese (Fitzhugh, 1974).

O modelo de Richards é um modelo do tipo (2.1) com uma taxa de crescimento $r(N_t)$ proposta por Richards em 1959. Podemos escrevê-lo na forma discreta como a seguir. Considerando o processo aleatório definido em (2.1) para representar a variabilidade devida a fatores ambientais, denotado por ϵ_t ,

$$N_{t+1} = N_t \exp \left\{ \frac{\rho}{q} \left[1 - \left(\frac{N_t}{K} \right)^q \right] + \epsilon_t \right\} \quad (2.2)$$

Observando a equação em (2.2) vemos que a taxa de crescimento de Richards é dada por

$$r(N_t) = \frac{\rho}{q} \left[1 - \left(\frac{N_t}{K} \right)^q \right] \quad (2.3)$$

Na equação (2.3), ρ/q representa a taxa intrínseca de crescimento e o parâmetro K é a capacidade de suporte da população.

O parâmetro q também representa um fator intrínseco da população, denominado parâmetro de forma. Quando o modelo é usado para o peso dos animais, este fator está relacionado com o metabolismo dos animais, o que facilita a interpretação. No caso de crescimento populacional, esse parâmetro descreve a forma como a taxa $r(N_t)$ diminui quando N_t aumenta, ou seja, se $q > 1$, temos uma queda suave; se $q < 1$, temos uma queda mais rápida; e, se $q = 1$, a queda é linear.

A taxa de Richards pode ser considerada geral porque, variando-se o valor do parâmetro q , temos uma infinidade de modelos. Para simplificar, podemos escrever a taxa de Richards na forma:

$$r(N_t) = \alpha_0 + \alpha_1 N_t^q \quad (2.4)$$

com $\alpha_0 = \rho/q$ e $\alpha_1 = -\rho/qK^q$. Assim, temos com este modelo a capacidade de suporte dada por $K = (-\alpha_0/\alpha_1)^{1/q}$ e a taxa intrínseca de crescimento dada por α_0 .

Na maioria dos trabalhos encontrados na literatura, os modelos logístico e Gompertz são os mais utilizados devido às dificuldades de se fazer inferência dos parâmetros do modelo não linear de Richards (Ludwig, 1996, 1999). A principal dificuldade está na identificabilidade do parâmetro de forma q , ou seja, na presença de variabilidade grande nos dados é difícil calcular uma boa estimativa deste parâmetro.

2.3 Modelos logístico e de Gompertz

A equação logística é baseada puramente na operação de fatores dependentes da densidade. Embora tenha sido demonstrado que o crescimento de uma grande variedade de populações segue um padrão sigmoidal, não decorre necessariamente daí que tais populações aumentam segundo a equação logística. Muitas equações matemáticas produzem uma curva sigmoidal. A equação logística representa um tipo de forma sigmoidal mínima de crescimento, uma vez que os efeitos limitantes de espaço e recursos começam no início do crescimento. Na maioria dos casos, esperaríamos, de início, um crescimento menos limitado, seguido por um retardamento à medida que aumentasse a densidade.

O modelo logístico equivale ao modelo de Richards com $q = 1$. Da expressão (2.3), temos a taxa de crescimento logística dada por:

$$r(N_t) = \rho \left[1 - \frac{N_t}{K} \right] \quad (2.5)$$

Para facilitar a inferência dos parâmetros, podemos escrever essa taxa na forma linear:

$$r(N_t) = \alpha_0 + \alpha_1 N_t \quad \text{com } \alpha_0 = \rho \text{ e } \alpha_1 = -\rho/K \Rightarrow K = -\alpha_0/\alpha_1 \quad (2.6)$$

Enquanto que a taxa logística é proporcional a N_t , a taxa de Gompertz é proporcional a $\ln(N_t)$. Esse modelo é menos usado que o logístico, mas tem sido muito usado para o crescimento de diversos tipos de tumores. O modelo de Gompertz pode ser obtido quando consideramos o limite da taxa de crescimento de Richards (2.3) quando $q \rightarrow 0$.

$$\lim_{q \rightarrow 0} r(N_t) = \lim_{q \rightarrow 0} \frac{\rho}{q} \left[1 - \left(\frac{N_t}{K} \right)^q \right] = \lim_{q \rightarrow 0} \frac{\rho}{q} \left[1 - \exp \left\{ q \ln \left(\frac{N_t}{K} \right) \right\} \right]$$

Aplicando a regra de L'Hôpital, temos a taxa Gompertz dada por:

$$\lim_{q \rightarrow 0} r(N_t) = \lim_{q \rightarrow 0} \left[-\rho \ln \left(\frac{N_t}{K} \right) \exp \left\{ q \ln \left(\frac{N_t}{K} \right) \right\} \right] = -\rho \ln \left(\frac{N_t}{K} \right)$$

Assim como a taxa logística, a taxa Gompertz pode ser escrita na forma linear:

$$r(N_t) = \alpha_0 + \alpha_1 \ln(N_t) \quad (2.7)$$

$$\text{com } \alpha_0 = \rho \ln(K) \text{ e } \alpha_1 = -\rho \Rightarrow K = \exp \{ -\alpha_0/\alpha_1 \}$$

2.4 Comparação entre os modelos de Richards, logístico e de Gompertz

Uma das maiores dificuldades nos estudos ecológicos é a seleção de um modelo adequado. Os pesquisadores da área frequentemente questionam: Por que o número atual de algumas espécies é tão pequeno se era muito maior há pouco tempo atrás? Sabemos que os dados reais podem apresentar uma variabilidade grande devida a fatores ambientais imprevisíveis como catástrofes ecológicas, por exemplo. Essa variabilidade

pode tornar a seleção do modelo mais complicada, já que acarreta a falta de identificabilidade de um ou mais parâmetros.

Ajustando o modelo de Richards, com variabilidade grande nos dados, podemos não estimar o parâmetro q de forma única, e isso significa que o modelo adequado para aqueles dados poderia ser o logístico, por exemplo.

Para exemplificar como a variabilidade pode dificultar a seleção do modelo adequado, simulamos o crescimento com os três modelos citados. A simulação foi feita com os seguintes valores para os parâmetros: $K = 500$, $\rho = 0,3$ e $N_0 = 50$. Para simular o modelo de Richards, usamos três valores do parâmetro q ($q = 0,5$, $q = 1,5$ e $q = 2$). A comparação dos modelos estocásticos com variabilidades diferentes foi feita com $\tau = 1/\sigma^2 = 100$ e $\tau = 10$.

Se consideramos os modelos determinísticos, vemos na Figura 2.1 que o crescimento é bem diferente para cada modelo e, portanto, não há problema na identificação.

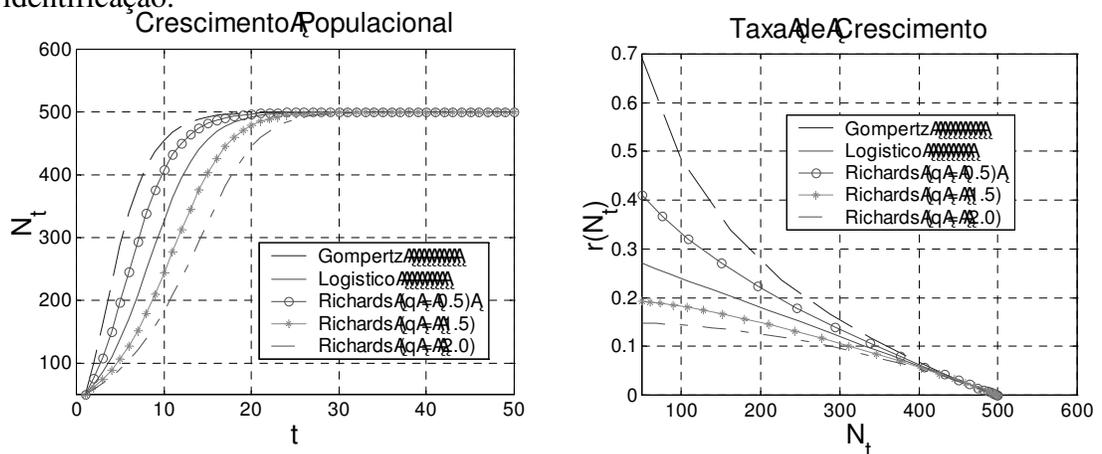


Figura 2.1 - Simulação com modelos determinísticos ($K = 500$, $\rho = 0,3$ e $N_0 = 50$)

Se assumimos modelos estocásticos, mas com variância do resíduo relativamente pequena ($\tau = 1/\sigma^2 = 100$), também vemos na Figura 2.2 que não há muito problema de identificabilidade.

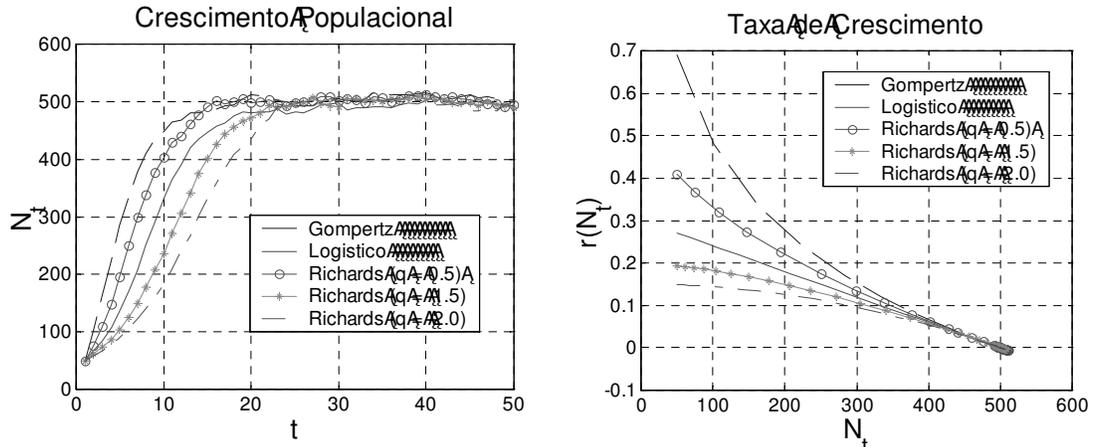


Figura 2.2 - Simulação com modelos estocásticos ($K = 500, \rho = 0,3 ; N_0 = 50$ e $\tau = 100$)

Na Figura 2.3, a simulação foi feita com uma variância maior ($\tau = 1/\sigma^2 = 10$) e, então, vemos o quanto os modelos se confundem. Neste caso, os critérios conhecidos para seleção de modelos podem falhar.

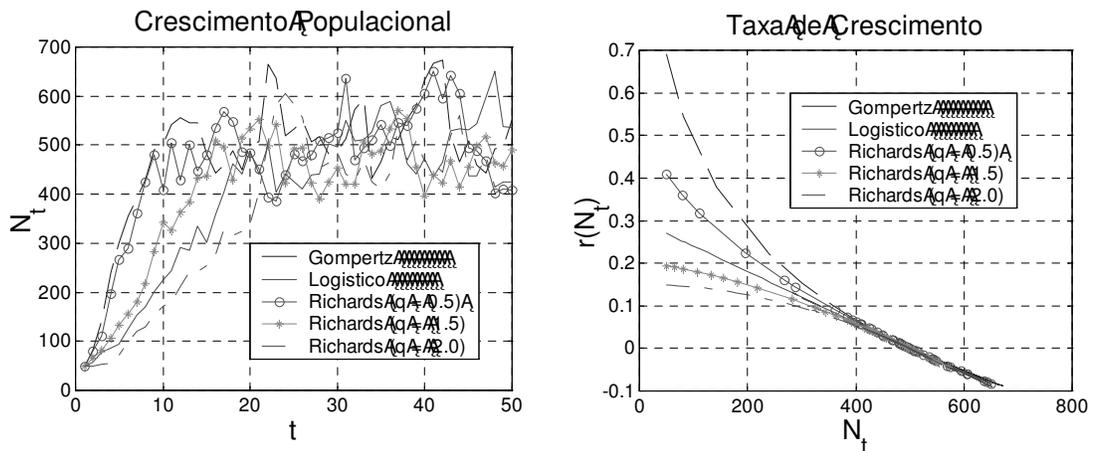


Figura 2.3 - Simulação com modelos estocásticos ($K = 500, \rho = 0,3 ; N_0 = 50$ e $\tau = 10$)

Do ponto de vista da estatística clássica, a escolha do modelo pode ser feita com o teste da razão de máxima verossimilhança *profile* e, do ponto de vista bayesiano, podemos usar o critério das preditivas ordenadas (Kass e Raftery, 1995). Os métodos para a seleção de modelos estão nos capítulos 3 e 4, respectivamente.

2.5 Modelos de crescimento com intervenção

O modelamento com intervenção torna-se muito importante quando queremos identificar os parâmetros do crescimento de populações que sofrem interferências humanas com objetivos econômicos ou ecológicos. Com essa abordagem, é possível fazer estimativas do risco de extinção e também do tempo até a extinção de populações devido a tais interferências. A intervenção entra no modelo como uma retirada de uma parte dos animais - retirada que pode ser pesca, caça ou abate no caso de criação de animais em cativeiro. A intervenção $\{h_t(N_t), t \geq 0\}$ pode ser modelada como uma função determinística ou estocástica do tamanho da população.

Consideramos o problema de crescimento determinístico com intervenção, definido como.

$$N_{t+1} = N_t \exp\{r(N_t)\} - h_t(N_t)N_t \quad (2.8)$$

Note que o termo estocástico, que envolve ϵ_t , não está presente. A intervenção, denotada por $h_t(N_t)$ pode ser escrita como uma função do tamanho da população. Essa hipótese é natural uma vez que a taxa de abate (ou retirada) em t deve ser influenciada pelo tamanho da população nesse instante, quando a decisão é tomada.

Consideramos que a taxa deve estar entre 0 e 1 e a escolha $h_t(N_t) = 1$ pode significar uma decisão drástica, se a população não se recupera rapidamente. A adoção dessa política não leva à extinção no passo seguinte uma vez que ao mesmo tempo que $h_t(N_t)$ é aplicado, no intervalo $[t, t + 1]$ a população se reproduz. No entanto é fácil mostrar que adotar $h_t(N_t) = 1$ pode levar a população, a longo prazo, a um nível muito baixo dependendo da sua taxa de reprodução $r(N_t)$.

A Assumindo que $\exp\{r(N_t)\} < 2$, ou seja o tamanho da população não duplica entre t e $t + 1$ *, podemos mostrar que o abate de 100% leva à extinção da população.

A Se $h_t(N_t) = 1$, denotamos $\delta_t = \exp\{r(N_t)\} - 1$, como $\exp\{r(N_t)\} < 2$, então $0 \leq \delta_t < 1$ e podemos escrever (2.11) como:

$$N_{t+1} = \delta_t N_t$$

Se $\delta_{\max} = \sup_t(\delta_t) < 1$, temos que:

$$N_{t+1} \leq \delta_{\max} N_t$$

*Caso de populações de mamíferos grandes como elefantes, rinocerontes, baleias, bois e etc.

Denotando o ponto de partida da população por N_0 , no intervalo de tempo $[0, T]$, a população é:

$$N_T \leq \delta_{\max}^T N_0$$

Assim, fazendo $T \rightarrow \infty$, temos que $N_T \rightarrow 0$.

O importante na exploração de recursos naturais (desenvolvimento sustentável) é obter o equilíbrio entre esses dois fatores de forma a permitir a intervenção (geralmente com apelo sócio-econômico) e ao mesmo tempo a preservação da população.

Ao considerar a intervenção estocástica podemos propor um modelo para $h_t(N_t)$, dado por:

$$h_t(N_t) = \phi_0 + \phi_1 N_t + \omega_t \quad (2.9)$$

onde ω_t é um ruído branco com $E(\omega_t) = 0$, $E(\omega_t^2) = \eta^{-1}$ e $E(\omega_t \omega_{t+k}) = 0$, $k = 1, 2, 3, \dots$. Nesse caso a identificação de modelos com intervenção só pode ser feita considerando a densidade conjunta $P(N_t, h_t(N_t))$. Este problema está apresentado em detalhes no capítulo 5.

2.6 Cálculo de indicadores da viabilidade de uma população

Para o cálculo de indicadores da viabilidade introduzimos uma partição no espaço de estados e modelamos o processo como uma cadeia de Markov a tempo discreto. Essa partição representa a discretização do espaço, assim ao considerar $N_t \in I_i = [a_i, b_i]$ estamos assumindo que $N_t = N^i = \frac{a_i + b_i}{2}$. Essa aproximação torna-se mais próxima do processo verdadeiro N_t a medida que diminuimos a amplitude dos intervalos $I_i, i = 1, \dots, N_S$, ou seja, fazendo $N_S \rightarrow \infty$. O espaço de estados \mathcal{S} pode ser decomposto na forma:

$$\mathcal{S} = \bigcup_{i=1}^{N_S} I_i \quad \text{com} \quad I_i \cap I_j = \emptyset, i \neq j$$

A probabilidade da população atingir um determinado estado I_j no instante $t + 1$, dado o histórico da população $\{N_u, u \leq t\}$ é escrita como:

$$P_{ij} = P(N_{t+1} \in I_j | N_u \in I_i, 0 \leq u \leq t, 0 \leq i, j \leq N_S, i \neq j)$$

(2.10)

Os modelos que consideramos estabelecem uma equação recursiva entre N_{t+1} e N_t dada, de forma geral, por:

$$N_{t+1} = g(N_t, \epsilon_t)$$

onde o processo $\{\epsilon_t, t \geq 0\}$ é independente e identicamente distribuído, tal que $E(\epsilon_t) = 0$ e

$$E(\epsilon_t \epsilon_{t+k}) = \begin{cases} \tau^{-1}, & k = 0 \\ 0, & k = 1, 2, 3, \dots \end{cases}$$

Nessas condições, $\{N_t, t \geq 0\}$ define um processo de Markov que relaciona a transição entre intervalos do espaço de estados, de forma que:

$$P_{ij} = P(N_{t+1} \in I_j | N_t \in I_i, 0 \leq i, j \leq N_S, i \neq j) \quad (2.11)$$

Seja \mathbf{P} uma matriz quadrada com as entradas P_{ij} , definida para todo $\{(i, j), 0 \leq i, j \leq N_S\}$. Então \mathbf{P} é chamada matriz de Markov sobre \mathbf{S} se satisfaz as seguintes propriedades:

$$P_{ij} \geq 0 \quad \text{e} \quad \sum_{j=1}^{N_S} P_{ij} = 1, \quad \forall 0 \leq i, j \leq N_S$$

Estamos interessados em calcular a probabilidade da cadeia sair de um estado I_i no instante t e atingir um determinado estado I_j no instante $t + s$, $s \geq 0$. Portanto, temos que calcular a matriz de probabilidade de transição em s passos denotada por \mathbf{P}^s .

Para qualquer $s \in [0, T]$, a probabilidade da cadeia sair do estado I_i para o estado I_j em s passos é $P\{N_s \in I_j | N_0 \in I_i\} = P_{ij}^s$, para todo $I_i, I_j \subset \mathbf{S}$ e $s > 0$.

Para calcular essa probabilidade podemos usar a equação de Chapman-Kolmogorov (Çinlar, 1975).

Para $0 < t < s$, temos:

$$P\{N_s \in I_j | N_0 \in I_i\} = \sum_k P\{N_s \in I_j, N_t \in I_k | N_0 \in I_i\} =$$

$$\sum_k P\{N_s \in I_j | N_t \in I_k\} P\{N_t \in I_k | N_0 \in I_i\} =$$

$$\sum_k P_{kj}^{s-t} P_{ik}^t = P_{ij}^s$$

Portanto a matriz de probabilidade de transição em s passos pode ser calculada com:

$$\mathbf{P}^s = \mathbf{P}^{s-t} \mathbf{P}^t$$

Em muitos problemas podemos estar interessados em calcular a probabilidade da população atingir qualquer estado acima de I_j . Assim denotando $G = \bigcup_{k=j}^{N_S} I_k$, temos:

$$P(N_s \in G | N_0 \in I_i) = \sum_{l \geq j} P(N_s \in I_l | N_0 \in I_i), \quad 1 \leq i < j \quad (2.12)$$

Ou seja, a probabilidade de atingir G é a soma das colunas l tal que $l \geq j$, na linha i da matriz \mathbf{P}^s .

Em alguns casos podemos estar interessados em calcular não somente a probabilidade da população atingir um estado I_j , em um tempo $s > 0$, mas queremos também calcular qual é o tempo esperado até que a população atinja esse estado. Dado que $N_0 \in I_i$, queremos calcular o valor esperado $E(T_{I_j})$ tal que N_s faz a primeira passagem por I_j . Existem inúmeras situações na prática onde o cálculo do tempo de primeira passagem é importante. Por exemplo, quando uma população encontra-se em estado de quase extinção e queremos calcular qual é o tempo esperado para que essa população saia da faixa de perigo. No caso de controle de populações, que crescem de forma explosiva, queremos calcular qual o tempo esperado até que a população atinja um nível que justifique uma intervenção de controle, por exemplo, um abate controlado da população.

Definição 1: Tempo de primeira passagem: Sejam I_i e I_j dois estados de uma cadeia de Markov $\{N_t, t \in \mathbb{Z}_+\}$ com espaço de estados discretos \mathcal{S} . Se para um dado tempo t , $N_t \in I_i$, o tempo de primeira passagem do estado I_i para o estado I_j é definido como :

$$T_{ij} = \inf\{s : s \in \mathbb{Z}_+, N_{t+s} \in I_j\} \quad (2.13)$$

Definição 2: Para qualquer $s \in \mathbb{Z}_+$ define-se a função de probabilidade de T_{ij} por:

$$g_{ij}(s) = P(T_{ij} = s) \quad I_i, I_j \in \mathcal{S}, s \in \mathbb{Z}_+ \quad (2.14)$$

Por convenção adotamos $g_{ij}(0) = 0$.

Definição 3: Valor esperado do tempo de primeira passagem:

$$\mu_{ij} = E[T_{ij}] \quad (2.15)$$

Proposição 1: Para todo par de estados I_i e $I_j \subset \mathcal{S}$, temos que:

$$g_{ij}(1) = P_{ij} \quad (2.16)$$

$$g_{ij}(t) = \sum_{k \neq j} P_{ik} g_{kj}(t-1) \quad (2.17)$$

onde P_{ij} é a probabilidade de transição a um passo de I_i para I_j .

Prova: Usando a Definição 2, para $t = 1$, temos:

$$P(T_{ij} = 1) = P(N_1 \in I_j | N_0 \in I_i)$$

Portanto

$$g_{ij}(1) = P_{ij}$$

A Sendo $P(T_{ij} = t) = P(N_t \in I_j | N_0 \in I_i)$ e considerando todos os possíveis estados I_k com $k \neq j$, temos que :

$$P(N_t \in I_j | N_0 \in I_i) = \sum_{k \neq j} P(N_t \in I_j | N_1 \in I_k) P(N_1 \in I_k | N_0 \in I_i)$$

A Então
$$g_{ij}(t) = \sum_{k \neq j} P_{ik} g_{kj}(t-1).$$

Proposição 2: $\mathcal{S} = \bigcup_{k=1}^{N_S} I_k$ com $I_i \cap I_j = \emptyset, \forall i \neq j$. Sejam $N_0 \in I_i$ e $\mu = (\mu_{ij}, 0 \leq i, j \leq N_S \text{ e } i \neq j)$, os tempos esperados até atingir o estado I_j , partindo de qualquer estado I_i . Podemos calcular μ resolvendo o sistema de equações dado por:

$$(I - Q)\mu = \mathbf{1} \quad (2.18)$$

onde I é a matriz identidade de dimensão $(N_S - 1)$, $\mathbf{1} = (1, \dots, 1)'$ é um vetor de dimensão $(N_S - 1) \times 1$ e Q é a matriz, também de dimensão $(N_S - 1)$, construída retirando-se a linha e a coluna j da matriz P .

Prova: Considerar que

$$E[T_{ij}|N_1 \in I_k] = \begin{cases} 1 & \text{se } k = j \\ 1 + \mu_{kj} & \text{se } k \neq j \end{cases} \quad (2.19)$$

Podemos escrever:

$$\mu_{ij} = E[T_{ij}] = E_{N_1}\{E[T_{ij}|N_1 \in I_k]\} = \sum_k E[T_{ij}|N_1 \in I_k]P_{ik} \quad (2.20)$$

Substituindo (2.19) em (2.20), temos:

$$\begin{aligned} \mu_{ij} &= 1.P_{ij} + \sum_{k \neq j} (1 + \mu_{kj})P_{ik} = P_{ij} + \sum_{k \neq j} P_{ik} + \sum_{k \neq j} \mu_{kj}P_{ik} \\ \mu_{ij} &= E[T_{ij}] = 1 + \sum_{k \neq j} \mu_{kj}P_{ik} \end{aligned} \quad (2.21)$$

Fixando j e escrevendo (2.21) para $\mu = (\mu_{ij}, 0 \leq i, j \leq N_S \text{ e } i \neq j)$, temos o sistema de equações dado por $(I - Q)\mu = \mathbf{1}$.

A solução do sistema de equações (2.18) é única se \mathbf{P} é uma matriz estocástica irredutível que representa a matriz de transição de uma cadeia de Markov com estados ergódicos (Heyman e Sobel, 1982).

Exemplo

Para calcular o tempo esperado até a extinção, considere o crescimento de uma população animal ameaçada de extinção e supor que o espaço de estados é o conjunto de todos os tamanhos de população dividido em 4 intervalos, denotado por $\mathcal{S} = \{I_0, I_1, I_2, I_3\}$.

Estamos interessados em calcular o tempo esperado até atingir o estado de extinção, que está contido no intervalo I_0 da cadeia, partindo de qualquer estado contido nos outros 3 intervalos. Denotar por μ_{i0} o tempo esperado para atingir o estado de extinção (contido em I_0) partindo-se de um estado contido no intervalo I_i , com $i = 1, 2, 3$.

Dada a matriz de probabilidade de transição, temos que:

$$\mu_{i0} = 1 + \sum_{k \neq 0} \mu_{ko}P_{ik}$$

Portanto, os valores de μ_{10}, μ_{20} e μ_{30} podem ser calculados resolvendo-se o sistema de equações dado por:

$$\begin{pmatrix} 1 - P_{11} & -P_{12} & -P_{13} \\ -P_{21} & 1 - P_{22} & -P_{23} \\ -P_{31} & -P_{32} & 1 - P_{33} \end{pmatrix} \begin{pmatrix} \mu_{10} \\ \mu_{20} \\ \mu_{30} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Supondo que não haja transições entre os estados intermediários, isto é $P_{ij} = 0$, para $i, j = 1, 2, 3$ e $i \neq j$, o tempo esperado para atingir o estado I_0 , saindo do estado I_i , é o inverso da probabilidade P_{i0} .

$$\mu_{i0} = \frac{1}{1 - P_{ii}} = \frac{1}{P_{i0}}$$

Na presença de intervenção, a matriz de probabilidade de transição depende de h_t , assim esse método só pode ser aplicado com h_t fixo, ou determinístico. Quando adotamos h_t estocástico, o cálculo dos indicadores de viabilidade só pode ser feito por simulação de Monte Carlo.

2.6.1 Cálculo de indicadores da viabilidade de uma população com intervenção estocástica

Na presença de uma intervenção estocástica podemos calcular a probabilidade da população atingir um determinado conjunto nível definido por $G = \{x \in \mathbb{R}_+ : x \leq g_0\}$, $0 < g_0 < \infty$ ou $G = \{x \in \mathbb{R}_+ : x \geq g_1\}$, $0 < g_1 < \infty$. Estes intervalos podem estar associados a um nível muito baixo, indicando um estado de quase extinção, ou uma meta superior, indicando explosão demográfica. Além do cálculo da probabilidade de extinção, podemos estar interessados em calcular qual é o tempo que a população levará para atingir esse nível.

Considerando o modelo (2.8) com intervenção estocástica dada por (2.9) e lembrando que ϵ_t e ω_t são processos independentes com ϵ_t *i.i.d.* $N(0, \tau_0^{-1})$ e ω_t *i.i.d.* $N(0, \eta^{-1})$, podemos simular o comportamento da população por um longo período de tempo, gerando possíveis trajetórias para os processos $\{N_t^{(j)}, h_t^{(j)}(N_t^{(j)}), j = 1, \dots, J\}$ para valores de tempo a partir de $t = T_0$, para um certo instante inicial T_0 de interesse, no qual o tamanho da população é conhecido. Estimamos desta forma, o risco ou

probabilidade da população atingir uma determinada meta contando, entre todas as J trajetórias geradas, em quais trajetórias a meta foi atingida.

Denotamos o tempo decorrido para uma trajetória do processo atingir a meta G por T_G , com $T_G = \inf\{k > T_0 : N_k \in G | N_{T_0}\}$, e o número associado a essas ocorrências por $\#\{T_G < \infty\}$. Assim, temos que a estimativa da probabilidade da população atingir a meta é dada por:

$$\hat{P}(N_k \in G | N_{T_0}) = \frac{\#\{T_G < \infty | N_{T_0}\}}{J} \quad (2.22)$$

O tempo para a população atingir a meta é um tempo de parada T_G e seu valor esperado pode ser estimado usando-se as trajetórias geradas. Basta calcular a média dos processos $T_G^{(j)}$, $j = 1, 2, \dots, J$, assim temos:

$$\bar{T}_G = \frac{1}{J} \sum_{j=1}^J T_G^{(j)} \quad (2.23)$$

Devemos adotar um número de trajetórias J muito grande para que a lei dos grandes números possa assegurar a convergência dos estimadores (2.22) e (2.23). Outra dificuldade encontrada na aplicação desta técnica é que, dependendo da política de intervenção adotada $\{h_t(N_t), t \geq T_0\}$, a meta pode se tornar inatingível, ou somente atingível com uma probabilidade muito baixa. Assim, é necessário estabelecer um limite para o tempo máximo de simulação, que denotamos por T_∞ , isto é $\{T_k < \infty\} \equiv \{T_k < T_\infty\}$ e se na simulação do processo N_t temos que $N_t < G$ para $t \leq T_\infty$ então consideramos que a meta não foi atingida nesta realização. Nesse caso, $T_G^{(j)} = T_\infty$.

A comparação da modelagem por cadeia de Markov com a simulação de Monte Carlo é apresentada no Capítulo 6, no estudo de caso.

Capítulo 3

Inferência clássica com *bootstrap* para o modelo de Richards

O uso da teoria assintótica e a abordagem clássica do problema de inferência dos parâmetros do modelo de Richards apresentam algumas dificuldades devido ao mal condicionamento da matriz de informação de Fisher associada ao modelo. A identificabilidade pobre do parâmetro de forma dificulta o cálculo de uma estimativa única (Omlin, Reichert, 1998). Nesta tese, mostramos que o modelo de Richards pode ser interpretado como um modelo logístico quando aplicamos aos dados a transformação de potência de Box e Cox (Box e Cox, 1964). Com essa interpretação podemos usar o método de máxima verossimilhança *profile* (Barndorff-Nielsen e Cox, 1994; Murphy e Van Der Vaart, 2000) tornando viável a inferência dos parâmetros. As estimativas por intervalo calculadas com base na teoria assintótica podem ser imprecisas e, portanto, propomos os métodos *bootstrap* para o cálculo de estimativas por intervalo.

3.1 Transformação de Box e Cox e o modelo de Richards

O cálculo de estimadores clássicos para os parâmetros do modelo de Richards com os métodos dos mínimos quadrados ou de máxima verossimilhança apresenta diversas dificuldades. Um problema encontrado é que o uso de métodos numéricos como os métodos de Newton e quase-Newton para maximizar a função de verossimilhança não é possível, pois a matriz hessiana para esse modelo pode ser singular (Bickel, 1977).

Para contornar essa dificuldade, interpretamos o modelo de Richards como um modelo logístico quando aplicamos aos dados a transformação de Box e Cox. Neste caso, propomos o cálculo dos estimadores de máxima verossimilhança *profile* para os parâmetros do modelo de Richards, onde o parâmetro de forma q coincide com o parâmetro da transformação de potência de Box e Cox. Para admitir essa coincidência, consideramos a transformação de potência de Box e Cox dada através da seguinte relação:

$$N_t^{(q)} = \begin{cases} N_t^q, & q > 0 \\ \ln(N_t), & q = 0 \end{cases} \quad (3.1)$$

Consideramos a equação do modelo de Richards original, com algumas mudanças na notação para diferenciar do modelo logístico.

Podemos reescrever o modelo de Richards com as reparametrizações $K_L = K_R^q$ e $z_t = q\epsilon_t$. Então, o modelo para os dados transformados tem a seguinte forma:

$$N_{t+1}^q = N_t^q \exp\left\{\rho \left[1 - \frac{N_t^q}{K_L}\right]\right\} \exp\{z_t\}, \quad q > 0 \quad (3.2)$$

Elevando ambos os lados da equação (3.2) à potência $1/q$, chegamos ao modelo de Richards apresentado no capítulo 2 em (2.5):

$$N_{t+1} = N_t \exp\left\{\frac{\rho}{q} \left[1 - \left(\frac{N_t}{K_R}\right)^q\right]\right\} \exp\{\epsilon_t\}, \quad q > 0 \quad (3.3)$$

considerando em (3.3) $K_R = K_L^{1/q}$ e o ruído $\epsilon_t = z_t/q$.

Para fazer inferência dos parâmetros do modelo (3.2), é conveniente adotar a notação $N_t^{(q)} = N_t^q$, $q > 0$, e escrever o modelo logístico para $N_t^{(q)}$, dado por:

$$N_{t+1}^{(q)} = N_t^{(q)} \exp\left\{\rho \left[1 - \frac{N_t^{(q)}}{K_L}\right]\right\} \exp\{z_t\} \quad (3.4)$$

Fazendo $Y_t^{(q)} = \ln(N_{t+1}^{(q)}/N_t^{(q)})$, temos a equação (3.4) escrita na forma de um modelo linear:

$$Y_t^{(q)} = \alpha_0 + \alpha_1 N_t^{(q)} + z_t \quad (3.5)$$

com $\alpha_0 = \rho$ e $\alpha_1 = -\frac{\rho}{K_L}$.

O ruído $z_t = q\epsilon_t$ é normalmente distribuído com $E(z_t) = 0$ e $E(z_t z_{t+k}) = 0$ para $k \neq 0$. Sabemos que $var(z_t) = q^2 var(\epsilon_t)$; então, denotando $var(\epsilon_t) = \tau^{-1}$, temos que $var(z_t) = \tau_0^{-1} = q^2 \tau^{-1}$.

O uso da transformação de Box e Cox, introduzida aqui, permite-nos tratar estatisticamente o modelo de Richards, que era evitado por não se ter forma apropriada na literatura que levasse a (3.5). Normalmente restringem-se os estudos aos modelos logístico e de Gompertz.

3.2 Função de Verossimilhança *Profile* (FVP)

Para estimar os parâmetros do modelo de Richards com a transformação de Box e Cox utilizando a FVP, consideramos a equação (3.5) do modelo logístico:

$$Y_t^{(q)} = \alpha_0 + \alpha_1 N_t^{(q)} + z_t$$

Com o modelo escrito nessa forma, não temos dificuldades para estimar α_0 e α_1 para valores fixos de q e, assim, podemos encontrar uma estimativa \hat{q} para q que maximiza a FVP.

Como $z_t \sim N(0, \tau_0^{-1})$, a função densidade condicional de $Y_t^{(q)}$ dado $N_t^{(q)}$ é Normal com média $\alpha_0 + \alpha_1 N_t^{(q)}$ e variância $\text{var}(z_t) = \tau_0^{-1}$, de forma que:

$$P(Y_t^{(q)} | N_t^{(q)}) = \left(\frac{\tau_0}{2\pi} \right)^{1/2} \exp \left\{ - \frac{\tau_0}{2} (Y_t^{(q)} - \alpha_0 - \alpha_1 N_t^{(q)})^2 \right\} \quad (3.6)$$

Como estamos interessados em calcular os valores que maximizam a função de verossimilhança dos dados originais, determinamos esta função considerando o jacobiano da transformação. Temos que se $N_t^{(q)} = N_t^q$, então:

$$P(N_{t+1} | N_t) = P(N_{t+1}^{(q)} | N_t^{(q)}) \left| \frac{dN_{t+1}^{(q)}}{dN_{t+1}^q} \right| = P(N_{t+1}^{(q)} | N_t^{(q)}) q N_{t+1}^{q-1} \quad (3.7)$$

Mas, de acordo com (3.5), temos:

$$P(N_{t+1}^{(q)} | N_t^{(q)}) = P(Y_t^{(q)} | N_t^{(q)}) \left| \frac{dY_t^{(q)}}{dN_{t+1}^{(q)}} \right|$$

Com $P(Y_t^{(q)} | N_t^{(q)})$ dado em (3.6) e $Y_t^{(q)} = \ln \left(\frac{N_{t+1}^{(q)}}{N_t^{(q)}} \right)$, temos que a densidade $P(N_{t+1}^{(q)} | N_t^{(q)})$ é dada por:

$$P(N_{t+1}^{(q)} | N_t^{(q)}) = \left(\frac{\tau_0}{2\pi} \right)^{1/2} \exp \left\{ - \frac{\tau_0}{2} (Y_t^{(q)} - \alpha_0 - \alpha_1 N_t^{(q)})^2 \right\} \frac{1}{N_{t+1}^q} \quad (3.8)$$

Finalmente, podemos escrever a densidade condicional dos dados originais $P(N_{t+1} | N_t)$ como:

$$P(N_{t+1} | N_t) = \left(\frac{\tau_0}{2\pi} \right)^{1/2} \exp \left\{ - \frac{\tau_0}{2} (Y_t^{(q)} - \alpha_0 - \alpha_1 N_t^{(q)})^2 \right\} \frac{1}{N_{t+1}^q} q N_{t+1}^{q-1} \quad (3.9)$$

Em notação matricial, definindo $Y_j = Y_{t_j}$ e $N_j = N_{t_j}$ para $j = 1 \dots n$, temos $\mathbf{Y}^{(q)} = (Y_1^{(q)}, \dots, Y_n^{(q)})$, $\mathbf{N}^{(q)} = (N_1^{(q)}, \dots, N_n^{(q)})$ e $\mathbf{X}^{(q)} = [\mathbf{1} \ \mathbf{N}^{(q)}]$ com $q > 0$ e $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$. Para simplificar, usamos a notação $\mathbf{D} = [\mathbf{Y}^{(q)}, \mathbf{N}^{(q)}]$. A função de verossimilhança é dada por:

$$\begin{aligned}
L(q, \boldsymbol{\alpha}, \tau_0 | \mathbf{D}) &= \left(\frac{\tau_0}{2\pi} \right)^{\frac{n-1}{2}} \exp \left\{ -\frac{\tau_0}{2} (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha})' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha}) \right\} \frac{q^{n-1}}{\prod_{t=1}^{n-1} N_{t+1}} \\
&= \left(\frac{q^2 \tau_0}{2\pi} \right)^{\frac{n-1}{2}} \exp \left\{ -\frac{\tau_0}{2} (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha})' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha}) \right\} \frac{1}{\prod_{t=1}^{n-1} N_{t+1}} \quad (3.10)
\end{aligned}$$

Os valores de $\boldsymbol{\alpha}$ e τ_0 , que maximizam a função (3.10), são os mesmos que maximizam o logaritmo dessa função. Então, denotando $l(q, \boldsymbol{\alpha}, \tau_0 | \mathbf{D}) = \ln[L(q, \boldsymbol{\alpha}, \tau_0 | \mathbf{D})]$, temos:

$$\begin{aligned}
l(q, \boldsymbol{\alpha}, \tau_0 | \mathbf{D}) &= \frac{(n-1)}{2} [\ln(\tau_0) - \ln(2\pi)] - \sum_{t=1}^{n-1} \ln(N_{t+1}) - \\
&\quad - \frac{\tau_0}{2q^2} (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha})' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha}) + (n-1) \ln(q) \quad (3.11)
\end{aligned}$$

Com valores fixos de q em um intervalo arbitrário $[a, b]$, calculamos os estimadores de máxima verossimilhança de τ_0 e $\boldsymbol{\alpha}$ com:

$$\hat{\boldsymbol{\alpha}}_q = [\mathbf{X}^{(q)'} \mathbf{X}^{(q)}]^{-1} \mathbf{X}^{(q)'} \mathbf{Y}^{(q)} = \begin{pmatrix} \hat{\alpha}_{0q} \\ \hat{\alpha}_{1q} \end{pmatrix} \quad (3.12)$$

$$\hat{\tau}_{0q} = \frac{(n-1)}{(\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \hat{\boldsymbol{\alpha}}_q)' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \hat{\boldsymbol{\alpha}}_q)} \quad (3.13)$$

O valor do EMVP de q deve ser calculado de forma a maximizar a função $l(q | \hat{\boldsymbol{\alpha}}_q, \hat{\tau}_{0q}, \mathbf{D})$ obtida substituindo $\hat{\boldsymbol{\alpha}}_q$ e $\hat{\tau}_{0q}$ em (3.11). Denotando por $\hat{q} = \max l(q | \hat{\boldsymbol{\alpha}}_q, \hat{\tau}_{0q}, \mathbf{D})$, obtemos as estimativas de máxima verossimilhança *profile* de $\boldsymbol{\alpha}$ e τ_0 , substituindo \hat{q} em (3.12) e (3.13).

Com as estimativas dos parâmetros \hat{q} , $\hat{\tau}_0$, $\hat{\alpha}_0$ e $\hat{\alpha}_1$, calculamos estimativas dos parâmetros τ , ρ e K , usando as relações:

$$\hat{\tau} = \hat{q}^2 \hat{\tau}_0, \quad \hat{\rho} = \hat{\alpha}_0 \quad \text{e} \quad \hat{K} = \left(-\frac{\hat{\alpha}_0}{\hat{\alpha}_1} \right)^{1/\hat{q}} \quad (3.14)$$

Além disso, podemos calcular um intervalo de confiança para q de acordo com a relação:

$$2[l_p(\widehat{q}) - l_p(q)] \rightarrow \chi_{\nu_q}^2 \quad (3.15)$$

onde l_p é a função log-verossimilhança *profile* e ν_q é o número de componentes independentes em q . Maiores detalhes sobre esse método podem ser encontrados em Barndoff et ali., 1994.

3.3 Cálculo de intervalos de confiança para os parâmetros ρ , K e τ

Considerar o vetor de parâmetros $\gamma = (\alpha_0, \alpha_1, \tau_0, q_0)$ e a relação com o vetor de parâmetros de interesse $\theta = (\rho, K, \tau, q)$ dada por $\alpha_0 = \rho/q$, $\alpha_1 = -\rho/qK^q$, $\tau_0 = \tau/q^2$, $q_0 = q$.

Podemos calcular os intervalos de confiança para $\theta_q = (\rho_q, K_q, \tau_q)$ usando a propriedade de normalidade assintótica dos estimadores de máxima verossimilhança *profile* $\widehat{\theta}_q \sim Normal[\theta_q, I_q^{-1}(\widehat{\theta}_q)]$ (Murphy e Van der Vaart, 2000), onde $I_q(\widehat{\theta}_q)$ é a matriz de informação de Fisher observada para θ dado q . Denotando por $I(\widehat{\gamma})$ a matriz de informação de Fisher observada para os parâmetros γ , e $I(\widehat{\theta})$ a matriz de informação de Fisher para θ , podemos calcular $I(\widehat{\theta})$ usando o jacobiano da transformação dado pela matriz $J_{\gamma\theta}$ e a relação (3.16):

$$J_{\gamma\theta} = \frac{\partial \gamma}{\partial \theta} = \begin{pmatrix} \frac{1}{q} & 0 & 0 & \frac{-\rho}{q^2} \\ \frac{-1}{qK^q} & \frac{\rho}{K^{q+1}} & 0 & \frac{\rho}{q^2K^q} + \frac{\rho \ln(K)}{qK^q} \\ 0 & 0 & \frac{1}{q^2} & \frac{-2\tau}{q^3} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$I(\widehat{\theta}) = J'_{\gamma\theta} I(\widehat{\gamma}) J_{\gamma\theta} = \begin{pmatrix} I_{\widehat{\theta}_q \widehat{\theta}_q} & I_{\widehat{\theta}_q q} \\ I_{\widehat{\theta}_q q} & I_{qq} \end{pmatrix} \quad (3.16)$$

Devido ao mal condicionamento da matriz, não é possível calcular $I^{-1}(\widehat{\theta})$ diretamente, mas podemos usar a relação $\widehat{V}(\widehat{\theta}_q) = [I_{\widehat{\theta}_q \widehat{\theta}_q} - I_{\widehat{\theta}_q q} I_{qq}^{-1} I'_{\widehat{\theta}_q q}]^{-1}$ para o cálculo aproximado da variância estimada para θ_q (Seber, 1989). Com isso, calculamos os intervalos de confiança usando o desvio padrão estimado. O intervalo com 95% de confiança para θ_q é dado por:

$$\left[\widehat{\theta}_q - 2\sqrt{\widehat{V}(\widehat{\theta}_q)}, \widehat{\theta}_q + 2\sqrt{\widehat{V}(\widehat{\theta}_q)} \right] \quad (3.17)$$

3.4 Técnicas de *bootstrap*

Na identificação de parâmetros de um dado modelo, temos que considerar a estimação pontual e a estimação por intervalo. No primeiro caso, simples valores (ou estimativas) dos parâmetros são calculados a partir de propriedades locais do modelo no espaço de parâmetros, para o conjunto de dados observado. Essas estimativas só podem ser determinadas se os parâmetros puderem ser identificados de forma única para um certo conjunto de dados. A identificabilidade dos parâmetros do modelo é, assim, uma condição necessária para essa análise. Se os parâmetros não são unicamente identificáveis, a estrutura do modelo deve ser simplificada até que a identificabilidade seja atingida.

O outro caso é o de estimação por intervalo. Esse procedimento consiste em fazer estimativas da região provável para os parâmetros e, portanto, não busca uma solução única. No caso de uma boa identificabilidade dos parâmetros, a estimativa por intervalo resulta em uma região estreita (ou um intervalo de confiança estreito). Quando a identificabilidade é pobre e se existe um ou mais parâmetros não identificáveis, temos uma região mais ampla para os parâmetros com identificabilidade pobre e para os parâmetros com eles correlacionados. Em modelos ecológicos, é comum uma pobre identificabilidade dos parâmetros porque os dados usados na estimação são escassos em relação à complexidade dos modelos (Homes, 2001).

Nessas situações, as estimativas por intervalo são mais informativas que as simples estimativas pontuais. Uma abordagem usada para obter estimativas por intervalo é a inferência bayesiana (Gelman et al., 1995) que é adequada porque usa a teoria de probabilidade e o conhecimento a priori sobre os parâmetros é expresso em uma densidade de probabilidade.

O uso de outras técnicas de estimativas por intervalo para modelos ecológicos e que possam tratar amostras pequenas torna-se necessário. Devido à correlação entre os parâmetros e o mal condicionamento da matriz de informação de Fisher, a simples aplicação da técnica de máxima verossimilhança leva a cálculos imprecisos dos intervalos de confiança dos parâmetros. Além disso, observamos que o cálculo com a teoria assintótica, em alguns casos, pode levar a intervalos que não contêm o verdadeiro valor do parâmetro. Isso se deve, em parte, à complexidade do modelo e se agrava devido à limitação do uso da teoria assintótica em função do tamanho da amostra.

Neste trabalho propomos a utilização do método de bootstrap (Wehrens et al., 2000 ; Efron e Tibshirane, 1993) para a estimativa por intervalo dos parâmetros do modelo de Richards. Resumidamente, podemos descrever essa técnica como sendo um procedimento de re-amostragem para o cálculo de medidas de precisão de estimativas como erro padrão, vício e intervalos de confiança. A vantagem desse procedimento é que ele pode ser usado para estimadores de qualquer parâmetro como média , mediana, coeficiente de correlação etc.

3.4.1 *Bootstrap* paramétrico

Para apresentar o procedimento padrão *bootstrap* consideramos uma variável aleatória X , com valores observados (x_1, \dots, x_n) e função distribuição de probabilidade acumulada $F(X, \theta)$, onde θ são os parâmetros da função. Denotamos, também, a estimativa de máxima verossimilhança de θ por $\hat{\theta} = g(X)$.

A estimativa do intervalo de confiança de θ é usualmente obtida do estimador $\hat{\theta}$, considerando-se a distribuição de probabilidade de $\hat{\theta} - \theta$. Se s_α denota o α -percentil da distribuição, então o intervalo de confiança para θ é baseado na afirmação probabilística de que $P(s_{\alpha/2} \leq \hat{\theta} - \theta \leq s_{1-\alpha/2}) = 1 - \alpha$, que, após ser reescrita, leva ao intervalo

$$\hat{\theta} - s_{1-\alpha/2} \leq \theta \leq \hat{\theta} - s_{\alpha/2}$$

Note-se que, para estimar esse intervalo precisamos conhecer ou ser capaz de aproximar a distribuição de $\hat{\theta} - \theta$. A hipótese de normalidade assintótica assume que $\hat{\theta} \sim N(\theta, I^{-1}(\theta))$, onde $I(\theta)$ é a matriz de informação de Fisher, ou considera a aproximação $(\hat{\theta} - \theta) \sim t_{n-p}$, onde t_{n-p} representa a distribuição *t*-Student com $n - p$ graus de liberdade, sendo p o total de parâmetros desconhecidos a serem estimados.

O método *bootstrap* é baseado na idéia que a variabilidade de $\hat{\theta}$ em torno de θ é semelhante à variabilidade do estimador *bootstrap* $\hat{\theta}^*$ em torno de $\hat{\theta}$. Em alguns casos, o estimador *bootstrap* pode ser calculado analiticamente, mas em geral são necessárias reamostragens. Nestes casos o procedimento, conhecido como *bootstrap* paramétrico, consiste em gerar amostras da distribuição $F(X, \hat{\theta})$.

Aqui denotamos as B amostras geradas por $X^*(b)$, $b = 1, \dots, B$, e as estimativas *bootstrap* de θ calculadas para cada uma dessas amostras são denotadas por $\theta^*(b)$, $b = 1, \dots, B$. Com as B estimativas *bootstrap*, podemos calcular a estimativa do erro padrão de $\hat{\theta}$ como:

$$\widehat{ep}^*(B) = \left\{ \frac{\sum_{b=1}^B (\theta^*(b) - \widehat{\theta}^*)^2}{B-1} \right\}^{1/2} \quad \text{com} \quad \widehat{\theta}^* = \frac{\sum_{b=1}^B \theta^*(b)}{B} \quad (3.18)$$

A estimativa ideal do erro padrão seria obtida com $B \rightarrow \infty$, mas o esforço computacional, que depende principalmente do cálculo das estimativas *bootstrap* $\theta^*(b)$, cresce linearmente com B . É claro que a precisão da estimativa do erro padrão deve ser tão boa quanto as estimativas de qualquer outro parâmetro de interesse e, portanto, esperamos que a estimativa tenha vício e desvio padrão pequenos. A estimativa *bootstrap* ideal $\widehat{ep}^*(\infty)$ tem o menor desvio padrão entre todas as estimativas não viciadas. Uma maneira de escolher o valor de B satisfatoriamente é calculando o coeficiente de variação de $\widehat{ep}^*(B)$. O aumento na variabilidade devido à escolha de um valor de B insuficiente é refletido em um aumento no coeficiente de variação. Segundo Efron e Tibshirani (1993), frequentemente são necessárias 200 replicações para estimar o desvio padrão e valores muito maiores para B no caso do cálculo de intervalos de confiança.

3.4.2 Intervalos de Confiança *Bootstrap*

Através do uso do *bootstrap*, podemos obter intervalos precisos sem usar a suposição de normalidade assintótica. Vários métodos têm sido propostos para calcular o intervalo de confiança com o método *bootstrap*. Uma revisão pode ser encontrada em Efron e Tibshirani, (1993) e outra revisão, mais recente, pode se encontrada em Wehrens et al, (2000).

Entre as técnicas conhecidas, as mais precisas são: 1) Percentil *bootstrap* 2) Percentil *bootstrap* corrigido BC_a (Bias-Corrected and accelerated) e 3) *ABC* (Approximate *Bootstrap* Confidence intervals).

Método dos percentis *bootstrap*:

O intervalo de confiança baseado na distribuição *bootstrap* parte da suposição de que a distribuição de $(\widehat{\theta} - \theta)$ é aproximada pela distribuição de $(\widehat{\theta}^* - \theta)$. Isso significa

que também os percentis da distribuição de $(\hat{\theta} - \theta)$ são aproximados pelos percentis da distribuição *bootstrap* de $(\hat{\theta}^* - \theta)$. Assim, os percentis $s_{\alpha/2}$ e $s_{1-\alpha/2}$ da distribuição de $(\hat{\theta} - \theta)$ que são usados no intervalo de confiança clássico dado por:

$$\hat{\theta} - s_{1-\alpha/2} \leq \theta \leq \hat{\theta} - s_{\alpha/2}$$

são substituídos pelos percentis da aproximação *bootstrap* $s_{(B+1)\alpha/2}^*$ e $s_{(B+1)(1-\alpha/2)}^*$, e o intervalo de confiança é dado por:

$$\hat{\theta} - s_{(B+1)(1-\alpha/2)}^* \leq \theta \leq \hat{\theta} - s_{(B+1)\alpha/2}^* \quad (3.19)$$

Podemos reescrever o intervalo de confiança (3.19) considerando os quantis da distribuição de $\hat{\theta}^*$ em lugar da versão centrada $(\hat{\theta}^* - \hat{\theta})$. Assim, se r_{α}^* e s_{α}^* são respectivamente os α -ésimos percentis das distribuições de $\hat{\theta}^*$ e $(\hat{\theta}^* - \hat{\theta})$, então:

$$r_{\alpha}^* = s_{\alpha}^* + \hat{\theta} \quad (3.20)$$

pela simples razão de as duas distribuições diferirem somente por um deslocamento $\hat{\theta}$.

Substituindo a relação (3.20) na equação (3.19), temos o intervalo de confiança baseado na distribuição *bootstrap*, dado por:

$$2\hat{\theta} - r_{(B+1)(1-\alpha/2)}^* \leq \theta \leq 2\hat{\theta} - r_{(B+1)\alpha/2}^* \quad (3.21)$$

Uma propriedade importante do intervalo baseado nos percentis é que ele respeita transformações nos parâmetros da distribuição. Supondo que adotamos uma transformação $\hat{\phi} = m(\hat{\theta})$ que normaliza a distribuição de $\hat{\theta}$, ou seja, que $\hat{\phi}$ tem distribuição gaussiana com média ϕ e variância c^2 , então o intervalo baseado em percentis para θ é:

$$m^{-1}(\hat{\phi} - cz_{(1-\alpha)}) \leq \theta \leq m^{-1}(\hat{\phi} - cz_{\alpha}) \quad (3.22)$$

A Essa propriedade é importante principalmente quando a distribuição de $\hat{\theta}$ é claramente assimétrica, e o intervalo normal padrão só é correto quando adotamos uma transformação que normaliza essa distribuição. O método baseado nos percentis já faz essa transformação automaticamente. A grande vantagem desse método, portanto, é não exigir o conhecimento explícito da transformação que normaliza a distribuição de $\hat{\theta}$. Entretanto, para que o método do percentil *bootstrap* tenha boa precisão, é necessário que essa transformação exista, mas isso nem sempre ocorre. Para tais situações, uma correção no método dos percentis foi proposto por Efron e Tibshirane (1993). O método

corrigido é denominado de BC_a (bias-corrected and accelerated). O método BC_a melhora o método dos percentis em ambos os sentidos, teoria e prática. No entanto, ele exige maior esforço computacional. Para contornar essa dificuldade, foi proposto, também por Efron e Tibshirane, o método ABC (Approximate *Bootstrap* Confidence intervals), que não abordamos nesta tese.

Método BC_a :

A construção do intervalo por este método é mais complicada que a construção dos intervalos percentis, mas seu uso é mais fácil. Seja $\hat{\theta}^{*(\alpha)}$ o α -ésimo percentil da distribuição dos B estimadores *bootstrap* denotados por $\hat{\theta}^*(1), \dots, \hat{\theta}^*(B)$.

O intervalo de confiança calculado pelo método BC_a também usa os percentis da distribuição *bootstrap*, mas não necessariamente os mesmos percentis do método anterior. Os percentis usados agora dependem de dois números a e z_0 , chamados de aceleração e correção de vício. Então o intervalo BC_a com $100(1 - \alpha)\%$ de confiança é dado por:

$$(\hat{\theta}_{inf}, \hat{\theta}_{sup}) = (\hat{\theta}^{*(\alpha_1)}, \hat{\theta}^{*(\alpha_2)}) \quad (3.23)$$

com

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(\alpha)})}\right) \text{ e } \alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(1-\alpha)}}{1 - \hat{a}(\hat{z}_0 + z^{(1-\alpha)})}\right) \quad (3.24)$$

Nas equações em (3.24), $\Phi(\cdot)$ representa a função distribuição acumulada normal padrão, e $z^{(\alpha)}$ é o $100\alpha - \text{ésimo}$ percentil da distribuição normal padrão. Note-se que, se z_0 e a são iguais a zero, então o intervalo de confiança BC_a coincide com o intervalo de confiança normal padrão. O valor da correção do vício z_0 é calculado diretamente da proporção entre o número de estimativas *bootstrap* menores que a estimativa original $\hat{\theta}$, ou seja:

$$\hat{z}_0 = \Phi^{-1}\left(\frac{\#\{\hat{\theta}^*(b) < \hat{\theta}\}}{B}\right) \quad (3.25)$$

onde $\Phi(\cdot)$ representa a função distribuição acumulada normal padrão. Rigorosamente falando, z_0 mede o vício mediano de $\hat{\theta}^*$ em relação a $\hat{\theta}$ em unidade normal. Obtemos $z_0 = 0$ se exatamente metade das estimativas $\hat{\theta}^*(b)$ é menor que $\hat{\theta}$.

A Existem várias formas para calcular a aceleração a . A mais simples é dada em termos do estimador jackknife da estatística $\hat{\theta} = s(x)$. Seja $x_{(i)}$ a amostra original com o i -ésimo ponto retirado e a estatística $\hat{\theta}_{(i)} = s(x_{(i)})$. Defina $\hat{\theta}(\cdot) = \sum_{i=1}^n \hat{\theta}_{(i)}/n$. Uma expressão para a aceleração é dada por:

$$\hat{a} = \frac{\sum_{i=1}^n (\hat{\theta}(\cdot) - \hat{\theta}_{(i)})^3}{6 \left\{ \sum_{i=1}^n (\hat{\theta}(\cdot) - \hat{\theta}_{(i)})^2 \right\}^{3/2}} \quad (3.26)$$

A A estimativa \hat{a} é chamada de aceleração porque se refere à taxa de variação do erro padrão de $\hat{\theta}$ com relação ao valor verdadeiro do parâmetro θ . A aproximação normal padrão $\hat{\theta} \sim N(\theta, ep^2)$ assume que o erro padrão de $\hat{\theta}$ é o mesmo para todo estimador θ . No entanto, isso frequentemente não é realista, e a aceleração corrige essa distorção.

A O método BC_a tem duas vantagens teóricas importantes. Primeiro, ele preserva transformações nos parâmetros, como no método dos percentis. Isso significa que os pontos extremos do intervalo BC_a são transformados corretamente, se o parâmetro de interesse θ é trocado por uma função de θ . Por exemplo, o intervalo BC_a para $\sqrt{\theta}$ é obtido tomando-se a raiz quadrada dos pontos extremos do intervalo BC_a para θ .

A segunda vantagem é quanto à sua precisão. Um intervalo com $1 - 2\alpha$ de confiança tem probabilidade α de não conter o verdadeiro valor do parâmetro.

$$P(\theta < \hat{\theta}_{\text{inf}}) = P(\theta > \hat{\theta}_{\text{sup}}) = \alpha \quad (3.27)$$

Os intervalos de confiança aproximados são classificados de acordo com a sua precisão em relação a esta probabilidade. O método BC_a tem precisão de segunda ordem, ou seja, os erros com relação a (3.27) vão para zero com taxa $1/n$, isto é

$$P(\theta < \hat{\theta}_{\text{inf}}) \doteq \alpha + \frac{c_{\text{inf}}}{n} \text{ e } P(\theta > \hat{\theta}_{\text{sup}}) \doteq \alpha + \frac{c_{\text{sup}}}{n} \quad (3.28)$$

onde n é o tamanho da amostra e c_{inf} e c_{sup} são duas constantes.

Os intervalos padrão e os calculados pelo método dos percentis têm precisão de primeira ordem, portanto os erros com relação a (3.27) têm magnitude maior, pois

$$P(\theta < \hat{\theta}_{\text{inf}}) \doteq \alpha + \frac{c_{\text{inf}}}{\sqrt{n}} \text{ e } P(\theta > \hat{\theta}_{\text{sup}}) \doteq \alpha + \frac{c_{\text{sup}}}{\sqrt{n}} \quad (3.29)$$

O método BC_a é recomendado para uso geral, já que é o único que respeita transformações e tem boa precisão.

3.5 Escolha de modelos - Teste da razão de Verossimilhança *Profile*

Como foi visto na seção 3 do capítulo 2, o modelo logístico é equivalente ao modelo de Richards com $q = 1$. Para construir o teste da razão de verossimilhança a fim de verificar qual é o modelo que se ajusta melhor aos dados, consideramos as hipóteses:

$$H_0 : q = 1 \times H_1 : q \neq 1 \quad (3.30)$$

Então, sob H_0 o modelo logístico é o mais adequado e, se rejeitamos essa hipótese, ficamos com o modelo de Richards com o valor de q dado pelo estimador de máxima verossimilhança *profile* \hat{q} . A estatística do teste é dada pela razão dos máximos das funções de verossimilhança de cada modelo.

O máximo da função de verossimilhança do modelo de Richards denotado por $\max L_R$ é obtido substituindo-se os estimadores de máxima verossimilhança *profile* dos parâmetros do modelo $\hat{\tau}_R$, $\hat{\alpha}_R$ e \hat{q} , e é dado por:

$$\max L_R = \left(\frac{\hat{\tau}_R}{2\pi\hat{q}^2} \right)^{\binom{n-1}{2}} \exp \left\{ - \frac{\hat{\tau}_R}{2\hat{q}^2} (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_R)' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_R) \right\} \frac{\hat{q}^{n-1}}{\prod_{t=1}^{n-1} N_t}$$

$$\max L_R \propto \hat{\tau}_R^{\binom{n-1}{2}} \exp \left\{ - \frac{\hat{\tau}_R}{2\hat{q}^2} (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_R)' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_R) \right\} \quad (3.31)$$

Para o modelo logístico denotamos o máximo da função de verossimilhança por $\max L_{lo}$ e, substituindo os estimadores $\hat{\alpha}_{lo}$ e $\hat{\tau}_{lo}$, temos:

$$\max L_{lo} = \left(\frac{\hat{\tau}_{lo}}{2\pi} \right)^{\binom{n-1}{2}} \exp \left\{ - \frac{\hat{\tau}_{lo}}{2} (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_{lo})' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_{lo}) \right\} \frac{1}{\prod_{t=1}^{n-1} N_t}$$

$$\max L_{lo} \propto \hat{\tau}_{lo}^{\binom{n-1}{2}} \exp \left\{ - \frac{\hat{\tau}_{lo}}{2} (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_{lo})' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)}\hat{\alpha}_{lo}) \right\} \quad (3.32)$$

Então, a razão de máxima verossimilhança, que é a estatística do teste, é dada por:

$$\hat{\Lambda} = \frac{\max L_{lo}}{\max L_R}$$

$$\hat{\Lambda} \propto \left(\frac{\hat{\tau}_{lo}}{\hat{\tau}_R} \right)^{\left(\frac{n-1}{2}\right)} \exp \left\{ \frac{\hat{\tau}_R}{2\hat{q}^2} \sum_t (Y_t^{(\hat{q})} - \hat{\alpha}_R X_t^{(\hat{q})})^2 - \frac{\hat{\tau}_{lo}}{2} \sum_t (Y_t - \hat{\alpha}_{lo} X_t)^2 \right\} \quad (3.33)$$

A hipótese nula é rejeitada se o máximo da verossimilhança sob o modelo logístico é menor que o máximo da verossimilhança sob o modelo de Richards. O valor crítico para esse teste não pode ser obtido diretamente porque a estatística $\hat{\Lambda}$ em (3.33) não tem uma distribuição conhecida. Podemos usar a distribuição assintótica de $\hat{\lambda} = -2\log\hat{\Lambda}$ que é Qui-quadrado com $\nu = k - k_0$ graus de liberdade, onde k é o número de parâmetros do modelo de Richards e k_0 , o número de parâmetros do modelo logístico (Johnson, Wichern, 1992).

Uma alternativa que propomos é gerar valores de $\hat{\lambda}$ substituindo os estimadores de máxima verossimilhança *profile* em (3.33) e tomar o percentil 95% da distribuição empírica como valor crítico para o teste. Para gerar amostras dos estimadores dos parâmetros dos dois modelos, novamente usamos a técnica *bootstrap* e então, obtivemos a distribuição de $\hat{\lambda}$.

3.6 Aplicação da inferência clássica

Para avaliar o desempenho das técnicas apresentadas (máxima verossimilhança *profile* e método *bootstrap*) com relação ao tamanho da amostra, geramos dois conjuntos de dados, sendo um de tamanho 500 e outro de tamanho 50. Foram gerados 5000 valores com o modelo de Richards com os parâmetros $q = 0,6$, $K = 9000$, $\rho = 0,7$ e $\tau = 2$. Desprezamos os 4000 valores iniciais e utilizamos os últimos 1000 valores para então selecionar as duas amostras, evitando assim a influência do valor inicial escolhido para o tamanho da população. As trajetórias do processo $\{N_t, 0 \leq t \leq n\}$ com $n = 500$ e $n = 50$ estão apresentadas graficamente nas Figuras 3.1 e 3.2, respectivamente. Chamamos de Amostra 1 aquela com 500 pontos e Amostra 2 aquela com 50 pontos.

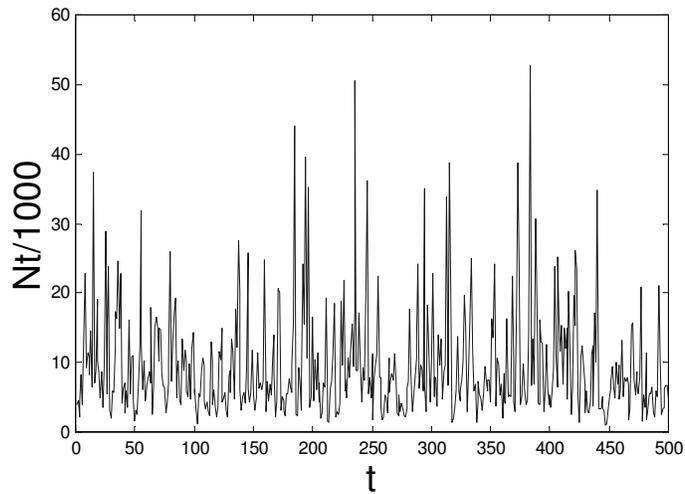


Figura 3.1: Amostra 1: n=500

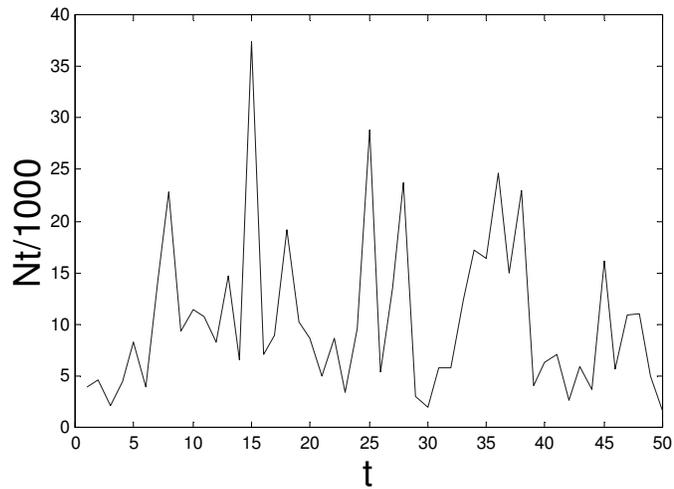


Figura 3.2: Amostra 2: n=50

Com essas duas amostras calculamos os estimadores de máxima verossimilhança *profile* dos parâmetros e os respectivos intervalos de confiança assintóticos. Também calculamos os intervalos bootstrap dos parâmetros com os métodos do percentil e BC_a (bias-corrected and accelerated). Na Figura 3.3 temos os gráficos do logaritmo das funções de verossimilhança *profile* nos quais podemos observar os intervalos de confiança 95% (IC) para q , para as Amostras 1 e 2, respectivamente.

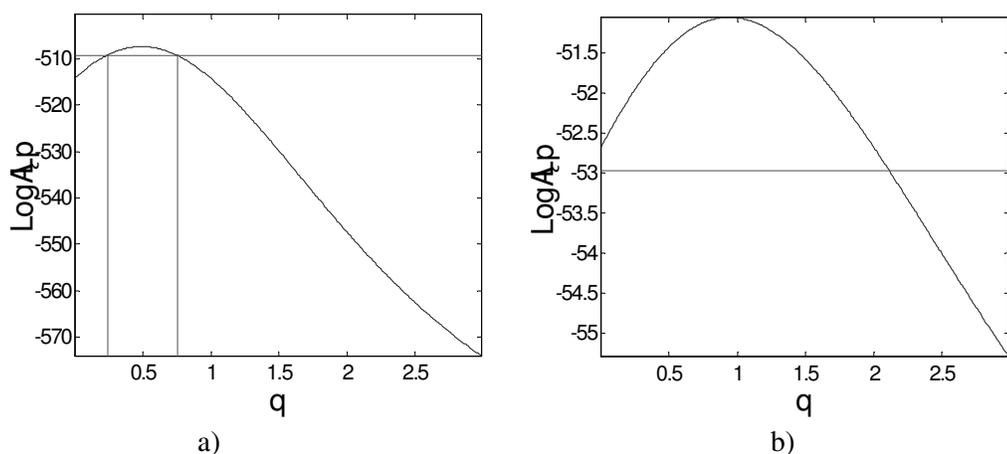


Figura 3.3 : Log Verossimilhança *profile*: a) Amostra 1 com IC para $q:(0,24 ; 0,75)$ e b) Amostra 2 com IC para $q:(0 ; 2,11)$

A seguir, nas Tabelas 3.1 e 3.2, apresentamos os resultados obtidos com o método da máxima verossimilhança *profile* (EMVP) e a comparação dos intervalos assintóticos com os intervalos obtidos com os dois métodos do *bootstrap* para as Amostras 1 e 2, respectivamente.

Tabela 3.1: Amostra 1 ($n=500$) $\tau = 2$, $q = 0,6$, $K = 9000$ e $\rho = 0,7$

	EMVP	IC assintótico	IC - percentil	IC - BCa
τ	2,2306	[2,2160 ; 2,2452]*	[2 ; 2,5588]	[1,9821 ; 2,5373]
q	0,49	[0,24 ; 0,75]	[0,24 ; 0,72]	[0,2423 ; 0,72]
K	8197,1	[8196,3 ; 8197,9]*	[7367,3 ; 9094,8]	[7382,9 ; 9124,7]
ρ	0,6704	[0,6309 ; 0,7099]	[0,5910 ; 0,7574]	[0,5769 ; 0,7454]

* Intervalos que não contêm o valor verdadeiro

Tabela 3.2: Amostra 2 ($n=50$) $\tau = 2$, $q = 0,6$, $K = 9000$ e $\rho = 0,7$

	EMVP	IC assintótico	IC - percentil	IC - BCa
τ	2,0831	[1,3369 ; 2,8293]	[1,4314 ; 3,4636]	[1,3519 ; 2,9820]
q	0,94	[0 ; 2,11]	[0,1 ; 1]	[0,1 ; 1]
K	10164,8	[10138 ; 10191]*	[7119,3 ; 12715,9]	[7737,7 ; 13637,6]
ρ	0,7626	[0,4162 ; 1,1090]	[0,4847 ; 1,1357]	[0,3893 ; 1,0260]

* Intervalo que não contém o valor verdadeiro

Analisando as Tabelas 3.1 e 3.2, verificamos que o método *bootstrap* é mais adequado para calcular intervalos de confiança para os parâmetros do modelo de

Richards. Todos os intervalos *bootstrap* contêm os valores verdadeiros dos parâmetros, fato que não se verifica com os intervalos assintóticos, mesmo no caso da amostra com 500 observações. Isso ocorre porque, como é previsto com a teoria assintótica, com amostras maiores a amplitude dos intervalos de confiança diminui. No entanto a propriedade de normalidade assintótica leva a uma imprecisão fazendo com que os intervalos de confiança não incluam os valores verdadeiros dos parâmetros.

Nas Figuras 3.4 e 3.5, apresentamos os histogramas das estimativas *bootstrap* para as Amostras 1 e 2 respectivamente. Esses histogramas se aproximam das distribuições verdadeiras dos EMVP e neles podemos observar a deficiência da técnica padrão ao estimar a variabilidade das estimativas com base na teoria assintótica usando a matriz de informação de Fisher. O fato é que as distribuições verdadeiras dos estimadores podem ser assimétricas, como vemos no caso do parâmetro q , e, assim, a aproximação normal não é boa porque atribui mais massa de probabilidade aos valores maiores do que a distribuição verdadeira. Além disso, a dependência do processo $\{\epsilon_t, t \geq 0\}$ agrava a imprecisão, uma vez que a convergência é mais rápida com independência na amostra.

A aproximação assintótica para o estimador de q utiliza o fato de que $-2[l_p(\hat{q}) - l_p(q)] \sim \chi_1^2$. Essa aproximação é adequada para a Amostra 1 ($n = 500$).

No caso da Amostra 2, em que $n = 50$, a estimativa do parâmetro de forma apresenta valores próximos de $q = 1$ com uma elevada frequência, como mostra o histograma de q na Figura 3.5. Esse resultado indica que, mesmo considerando-se uma técnica robusta para estimar a variabilidade deste parâmetro, a estimativa pode levar a um erro na identificação do modelo optando por um modelo logístico ($q = 1$), quando na verdade a população foi gerada por um modelo de Richards ($q = 0,6$).

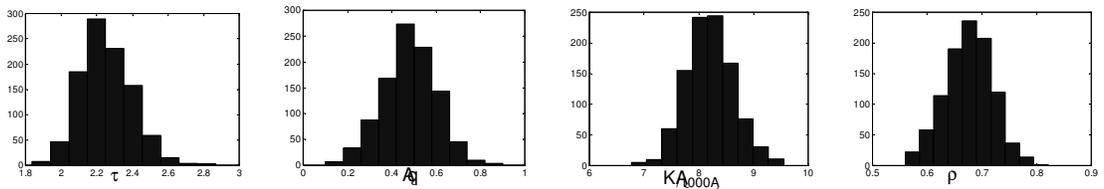


Figura 3.4: Distribuições *bootstrap* dos parâmetros τ , q , K e ρ para a Amostra 1 ($n = 500$)

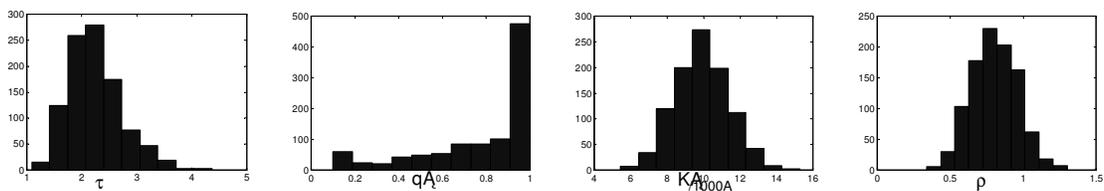


Figura 3.5: Distribuições *bootstrap* dos parâmetros τ , q , K e ρ para a Amostra 2 ($n = 50$)

Resultado do Teste da Razão de Verossimilhança *Profile*

As estimativas dos intervalos de confiança para o parâmetro q apontam para um modelo logístico no caso da Amostra 2 e para um modelo de Richards para a Amostra 1. Para selecionar o modelo mais adequado para cada uma das amostras, consideramos o teste da razão de verossimilhança comparando o modelo logístico com o modelo de Richards. As hipóteses do teste são: $H_0 : q = 1 \times H_1 : q \neq 1$.

Rejeitamos H_0 se $\hat{\lambda} > \lambda_c$; usando a teoria assintótica, temos $\hat{\lambda} = -2\log(\hat{\Lambda}) \sim \chi^2_{\alpha,1}$, assim com nível de significância $\alpha=0,05$, $\lambda_c = 3,841$. A imprecisão da teoria assintótica, que já se apresentou no cálculo dos intervalos de confiança dos parâmetros, também está presente aqui no cálculo de $\hat{\lambda}$. Assim, adotamos o procedimento de *bootstrap* para estimar o valor crítico do teste λ_c com mais precisão.

Aplicamos o teste para oito amostras com tamanhos diferentes, para verificar a partir de que tamanho esse resultado é válido. Todas essas amostras foram geradas com o modelo de Richards, com os parâmetros $q = 0,6$, $\rho = 0,7$, $\tau = 2$ e $K = 9000$. A Tabela 3.3 mostra os resultados do teste. Os valores de λ_c foram obtidos com o método *bootstrap* considerando-se no numerador o modelo logístico. Como ilustração apresentamos na Figura 3.6 os gráficos das distribuições empíricas de $\hat{\lambda}$ para as amostras 1 ($n = 500$) e 2 ($n = 50$).

Tabela 3.3: Resultados - Teste da razão de verossimilhança

n	$\hat{\lambda}$	λ_c	Conclusão
50	0,0099	3,1578	Não rejeito H_0 – escolher o modelo logístico
70	0,0563	3,3795	Não rejeito H_0 – escolher o modelo logístico
100	0,3273	2,8703	Não rejeito H_0 – escolher o modelo logístico
150	1,2929	3,4522	Não rejeito H_0 – escolher o modelo logístico
200	3,0489	2,9596	Rejeito H_0 – escolher o modelo de Richards
250	4,2855	2,8946	Rejeito H_0 – escolher o modelo de Richards
300	7,3773	3,1323	Rejeito H_0 – escolher o modelo de Richards
500	16,6817	3,0714	Rejeito H_0 – escolher o modelo de Richards

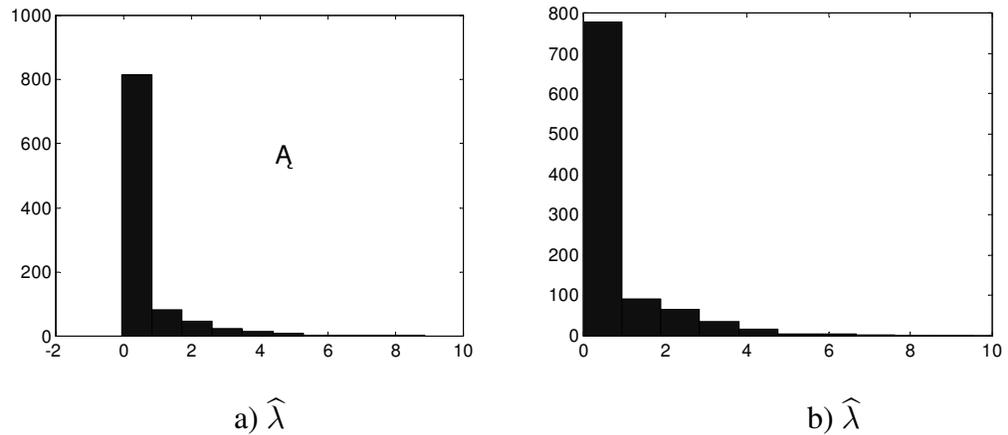


Figura 3.6: Distribuições empíricas da estatística do teste - a) $n = 500$, b) $n = 50$

A O método do teste de razão de verossimilhança se mostrou inadequado para amostras com menos de 200 observações. Nesses casos o modelo escolhido foi o logístico, quando os dados foram gerados pelo modelo de Richards. Esse método, portanto, não é recomendável para amostras pequenas.

A técnica de *bootstrap* se mostrou precisa, mas não o suficiente para evitar a superestimação da variabilidade de alguns parâmetros quando a amostra é pequena ($n = 50$), e poderíamos, erroneamente, escolher o modelo logístico ($q = 1$).

A coleta de dados sobre a população de uma espécie por longos períodos de tempo é algo difícil e muito raro pelas razões comentadas anteriormente. Para estes casos, o método apresentado tem limitações que ficam evidentes no estudo apresentado acima, e propomos como alternativa a abordagem bayesiana.

Para a identificação dos modelos de crescimento populacional nos casos com poucos dados amostrais e com dependência entre as observações, é interessante levar em conta as informações de especialistas tais como biólogos, ecologistas e técnicos de áreas relacionadas ao problema para incorporar algum conhecimento prévio ao procedimento da inferência. Essas considerações e pontos de vista podem ser incorporados no modelo quando fazemos a inferência bayesiana.

Capítulo 4

Abordagem bayesiana para o modelo de Richards

Neste capítulo apresentamos as técnicas bayesianas de inferência e escolha de modelos, fazendo um paralelo com as técnicas clássicas apresentadas no capítulo 3. Para comparar essas técnicas do ponto de vista prático, no final do capítulo apresentamos os resultados numéricos com as técnicas bayesianas para as mesmas amostras utilizadas na aplicação do capítulo 3.

4.1 Discussão sobre o uso de métodos clássicos e bayesianos

Os métodos bayesianos constituem uma maneira diferente de fazer ciência, sendo a estatística bayesiana muito mais que uma nova ferramenta para o repertório de métodos estatísticos. A inferência bayesiana rejeita vários "dogmas" da estatística e dos métodos científicos que são correntemente aceitos em ecologia e outras ciências. A abordagem bayesiana tem dividido o mundo da estatística em facções e podemos dizer que sua influência tem crescido rapidamente.

A controvérsia com relação à estatística bayesiana não está no teorema de Bayes mas sim no seu uso. Sejam X e Y variáveis aleatórias, com funções densidade de probabilidade $p_X(x)$ e $p_Y(y)$ respectivamente. O teorema de Bayes, quando expresso para variáveis aleatórias contínuas, estabelece que:

$$p_Y(y|x) = \frac{p_X(x|y)p_Y(y)}{p_X(x)} \quad (4.1)$$

com $p_X(x) = \int p_X(x|y)p_Y(y)dy$ e no qual $p_Y(y|x)$ é a função densidade de probabilidade condicional de Y dado $X = x$, e $p_X(x|y)$ é a função densidade de probabilidade de X dado $Y = y$.

Na estatística bayesiana, $p_Y(y)$ representa a "crença" subjetiva do pesquisador a priori sobre um parâmetro Y . O parâmetro é uma variável aleatória, e a função densidade de probabilidade a priori modela probabilisticamente a "crença" sobre possíveis valores do parâmetro. O pesquisador escolhe a forma da função densidade de probabilidade a priori que quantifica sua melhor opinião sobre o parâmetro. Na estatística bayesiana, x é

a saída observada do experimento, e $p_X(x|y)$ é a função de verossimilhança, que é familiar para os estatísticos frequentistas (função densidade de probabilidade dos dados X , conhecido o valor do parâmetro y). As conclusões sobre o parâmetro são sumarizadas na função densidade de probabilidade a posteriori $p_Y(y|x)$, que expressa a combinação entre a crença subjetiva do pesquisador e as informações contidas nos dados x . Em outras palavras, a inferência bayesiana combina a função de verossimilhança com uma função densidade de probabilidade conjunta dos parâmetros, que chamamos densidade conjunta a priori. Essa densidade é baseada em informações fornecidas, a priori, por especialistas da área em questão (no caso ecologistas, biólogos, técnicos em bovinocultura etc).

Na abordagem bayesiana, todas as conclusões sobre o valor do parâmetro são tiradas a partir da função densidade de probabilidade conjunta a posteriori. Os dados apenas entram na função de verossimilhança. Em particular, nenhuma probabilidade sobre o espaço amostral, além do valor real da função de verossimilhança, é admitida nas conclusões. As probabilidades sobre o espaço amostral, tal como a probabilidade que uma estatística de um teste exceda um valor crítico, envolvem "dados que não ocorreram" e são excluídas da análise. O princípio de incluir somente os dados reais na análise e excluir considerações sobre as possibilidades do espaço amostral é conhecido como princípio da verossimilhança.

Portanto, há dois elementos-chave na abordagem bayesiana: o primeiro é a quantificação de "crenças" a priori sobre o parâmetro na forma de uma distribuição de probabilidade e a incorporação dessas "crenças" na análise; o segundo é a aceitação do princípio da verossimilhança e a concomitante rejeição de todas as probabilidades sobre o espaço amostral das conclusões inferenciais sobre o parâmetro.

Nos testes estatísticos frequentistas, o $P - value$ é a probabilidade de rejeitar a hipótese nula quando esta é verdadeira (erro tipo I). Se esta probabilidade é menor que um dado nível de significância, então rejeita-se a hipótese nula. Para os bayesianos, o uso de $P - values$ para tirar conclusões viola o princípio da verossimilhança porque envolve uma probabilidade sobre o espaço amostral. Nos testes bayesianos cada hipótese recebe uma probabilidade subjetiva de estar correta. Os dados modificam estas "crenças" nas hipóteses através da função de verossimilhança. A hipótese que tiver maior probabilidade a posteriori é considerada pelo pesquisador como a verdadeira. Um par de hipóteses pode ser comparado com a razão de suas probabilidades a posteriori, que é proporcional à razão de verossimilhança ou fator de Bayes. Os conceitos de nível de

significância e poder do teste são probabilidades sobre o espaço amostral e não são considerados na estatística bayesiana.

A. O conceito de intervalo de confiança é central para a estatística frequentista. No entanto, a interpretação desse conceito, muitas vezes, sugere mais do que ele realmente informa. O intervalo com 95 % de confiança nos informa que, se tivéssemos 100 amostras aleatórias do mesmo tamanho e para cada uma delas calculássemos um intervalo de confiança, o parâmetro estaria contido em 95 dos intervalos. Ou seja, na estatística frequentista, os dados são aleatórios e portanto podemos considerar a probabilidade do intervalo conter o parâmetro, que é fixo. Os intervalos de confiança são baseados em aproximações assintóticas e, muitas vezes, a amostra disponível não tem tamanho suficiente para isso. Pode ocorrer um intervalo de confiança com extensão maior que o domínio do parâmetro. Por essas razões, os pesquisadores bayesianos criticam o método. Com o desenvolvimento de abordagens computacionais, como a verossimilhança profile e as técnicas de bootstrap, essas críticas tornam-se discutíveis.

Do ponto de vista bayesiano, a amostra é fixa e, portanto, não calculamos probabilidades no espaço amostral, e o parâmetro é uma variável aleatória. O intervalo de credibilidade $(1 - \alpha)$ é construído com esse conceito, tomando como pontos extremos do intervalo os percentis $\alpha/2$ e $1 - \alpha/2$ da função densidade de probabilidade a posteriori. No intervalo de credibilidade $(1 - \alpha)$ é a probabilidade do parâmetro estar entre esses dois números fixos (Dennis, 1996).

Uma das vantagens do método bayesiano é a possibilidade de combinar dados de diversos estudos através do uso da densidade a priori e, se isso é possível, não há a necessidade de amostras grandes já que a abordagem bayesiana não se baseia em teoria assintótica. Por outro lado, nem sempre temos informação a priori disponível e por isso pode ser usada uma densidade a priori não informativa. Nesse caso o tamanho da amostra é tão importante quanto na inferência clássica.

A grande desvantagem da inferência bayesiana é que, com uma densidade a priori imprópria, podemos obter uma densidade a posteriori imprópria, que não serve para tirar as conclusões de interesse. Além disso, pode ser difícil verificar se a densidade a posteriori é própria. Uma função densidade f é própria se satisfaz as propriedades básicas de função densidade de probabilidade dadas por: $f(x) \geq 0$, $-\infty < x < \infty$ e $\int_{-\infty}^{\infty} f(x)dx = 1$.

No problema de crescimento populacional com o modelo de Richards, a abordagem bayesiana é adequada, uma vez que os parâmetros de interesse ρ e K são

variáveis aleatórias e também pelo fato de que esses estudos contam com poucas observações. Mesmo com densidade a priori imprópria, com os modelos que estamos estudando verificamos que a densidade a posteriori é própria.

Para destacar que nesta abordagem os parâmetros são variáveis aleatórias, substituímos a notação do vetor de parâmetros α por β .

4.2 Inferência bayesiana para os modelos logístico e de Gompertz

No caso dos modelos logístico e de Gompertz, é possível calcular diretamente as densidades marginais a posteriori de β e τ e, assim, obter os estimadores bayesianos correspondentes de ρ e K através das relações dadas na seção 3 do capítulo 2. Para verificar essa particularidade desses modelos, consideramos a função de verossimilhança para o modelo geral de regressão linear $Y_t = \beta_0 + \beta_1 X_t + z_t$ e os resíduos $z_t = Y_t - \beta_0 - \beta_1 X_t$ i.i.d. $NA(0, \tau^{-1})$. No modelo logístico $X_t = N_t$ e no modelo de Gompertz $X_t = \ln(N_t)$. Denotando-se por \mathbf{D} o vetor dos dados $\mathbf{D} = (Y_1, \dots, Y_n, N_1, \dots, N_n)$, temos a função de verossimilhança em notação matricial dada por:

$$L(\beta, \tau | \mathbf{D}) = \left(\frac{\tau}{2\pi} \right)^{\frac{n-1}{2}} \exp \left\{ -\frac{\tau}{2} \left[C + (\beta - \hat{\beta})(\mathbf{X}'\mathbf{X})(\beta - \hat{\beta}) \right] \right\} \quad (4.2)$$

onde $C = (\mathbf{Y} - \mathbf{X}\hat{\beta})'(\mathbf{Y} - \mathbf{X}\hat{\beta})$ e $\hat{\beta}$ é o estimador de máxima verossimilhança de β . Considera-se a densidade a priori não informativa Uniforme (Box e Tiao, 1974) dada por

$$\pi_0(\beta, \tau) \propto \frac{1}{\tau} \quad (4.3)$$

A densidade conjunta a posteriori é obtida combinando-se a função de verossimilhança (4.2) e a densidade conjunta a priori dada em (4.3). Esta é a densidade Normal-Gama, que é própria. Assim temos que:

$$\begin{aligned} \pi(\beta, \tau | \mathbf{D}) &\propto \pi_0(\beta, \tau) L(\beta, \tau | \mathbf{D}) \propto \\ &\tau^{\frac{n-1}{2}-1} \exp \left\{ -\frac{\tau}{2} \left[C + (\beta - \hat{\beta})(\mathbf{X}'\mathbf{X})(\beta - \hat{\beta}) \right] \right\} \end{aligned} \quad (4.4)$$

Para calcular a densidade marginal a posteriori de τ , basta integrar a densidade conjunta com relação a β e temos:

$$\begin{aligned}
\pi(\tau|\mathbf{D}) &= \int_{-\infty}^{\infty} \pi(\boldsymbol{\beta}, \tau|\mathbf{D})d\boldsymbol{\beta} = \\
&\int_{-\infty}^{\infty} \tau^{\frac{n-1}{2}-1} \exp\left\{-\frac{\tau}{2}\left[C + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})(\mathbf{X}'\mathbf{X})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right]\right\}d\boldsymbol{\beta} \\
\pi(\tau|\mathbf{D}) &= \tau^{\frac{n-1}{2}-1} \exp\left\{-\frac{\tau}{2}C\right\} \frac{\sqrt{\tau}}{\sqrt{2\pi}|\mathbf{X}'\mathbf{X}|^{-1/2}} \\
\pi(\tau|\mathbf{D}) &\propto \tau^{\frac{n-1}{2}-1} \exp\left\{-\frac{\tau}{2}C\right\} \tag{4.5}
\end{aligned}$$

Da mesma forma, para calcular a densidade marginal a posteriori de $\boldsymbol{\beta}$, integrando a densidade conjunta a posteriori (4.4) com relação a τ , temos que:

$$\begin{aligned}
\pi(\boldsymbol{\beta}|\mathbf{D}) &= \int_0^{\infty} \pi(\boldsymbol{\beta}, \tau|\mathbf{D})d\tau = \\
&\int_0^{\infty} \tau^{\frac{n-1}{2}-1} \exp\left\{-\frac{\tau}{2}\left[C + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})(\mathbf{X}'\mathbf{X})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right]\right\}d\tau \\
\pi(\boldsymbol{\beta}|\mathbf{D}) &\propto \frac{2^{\frac{n-1}{2}}\Gamma\left(\frac{n-1}{2}\right)}{\left[C + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})(\mathbf{X}'\mathbf{X})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right]^{\frac{n-1}{2}}} \propto \\
&\frac{1}{\left[1 + \frac{1}{C}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})(\mathbf{X}'\mathbf{X})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})\right]^{\frac{n-1}{2}}} \tag{4.6}
\end{aligned}$$

Portanto, a densidade marginal a posteriori de τ é a densidade *Gama* $\left(\frac{n}{2}, \frac{C}{2}\right)$, e a densidade marginal a posteriori para $\boldsymbol{\beta} = (\beta_0, \beta_1)$ é a densidade *t-Student* bivariada com $(n - 2)$ graus de liberdade. Com essas densidades marginais a posteriori, podemos calcular estimativas dos parâmetros dos modelos lineares como o Gompertz e o logístico com as relações $\rho = \beta_0$ e $K = -\beta_0/\beta_1$.

4.3 Inferência bayesiana para o modelo de Richards

Para fazer a inferência bayesiana dos parâmetros do modelo de Richards, propomos a utilização do método de simulação de Monte Carlo em cadeias de Markov (Gilks et al, 1996) e utilizamos especificamente o algoritmo amostrador de Gibbs (Gelfand e Smith, 1990) em conjunto com o algoritmo Metropolis-Hasting (Chib e Greenberg, 1995). Esses dois algoritmos são apresentados no final desta seção.

Para encontrar a densidade a posteriori conjunta dos parâmetros do modelo, vamos reescrever a função de verossimilhança dada no capítulo 3 em (3.10) lembrando que, para esse modelo, $\mathbf{Y}^{(q)} = (Y_1^{(q)}, \dots, Y_n^{(q)})$, $\mathbf{N}^{(q)} = (N_1^{(q)}, \dots, N_n^{(q)})$, $\mathbf{X}^{(q)} = [\mathbf{1} \ \mathbf{N}^{(q)}]$ e $\mathbf{D} = [\mathbf{X}^{(q)} \ \mathbf{N}^{(q)}]$.

$$L(q, \boldsymbol{\beta}, \tau_0 | \mathbf{D}) = \left(\frac{\tau_0}{2\pi} \right)^{\frac{n-1}{2}} \exp \left\{ - \frac{\tau_0}{2} [B_q(\boldsymbol{\beta}) + C_q] \right\} \frac{q^{n-1}}{\prod_{t=1}^{n-1} N_{t+1}} \quad (4.7)$$

com

$$B_q(\boldsymbol{\beta}) = (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_q)(\mathbf{X}^{(q)'} \mathbf{X}^{(q)})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_q), C_q = (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \hat{\boldsymbol{\beta}}_q)' (\mathbf{X}^{(q)} - \mathbf{X}^{(q)} \hat{\boldsymbol{\beta}}_q)$$

e $\hat{\boldsymbol{\beta}}_q$ é o estimador de mínimos quadrados de $\boldsymbol{\beta}$ para q fixo.

O uso da família de transformações de potência nos dados para obter um modelo linear gaussiano, que aplicamos aqui para o modelo de Richards, foi estudado por Box e Cox, que desenvolveram as abordagens clássica e bayesiana paralelamente. Existem dois problemas nesta análise. Em primeiro lugar, os resultados clássicos e bayesianos podem levar a inferências contraditórias. O segundo ponto é que as densidades a priori não informativas impróprias propostas dependem da saída do experimento, que é o vetor de observações. Esse aspecto foi criticado por vários autores incluindo-se Nelder (1964) e Fraser (1967,1968). Box e Cox também alertam repetidamente em seu artigo que existe uma maneira melhor para construir a densidade a priori. Sobre isso Lindley (1972, p48) diz que a análise de Box e Cox é enigmática porque usa uma densidade sobre o espaço paramétrico que depende dos dados (apud Pericchi, 1981).

Segundo Pericchi (1981), a densidade a priori deve ser obtida sem a suposição de independência a priori dos parâmetros. Essa suposição é feita no cálculo da densidade a priori proposta por Box e Cox, que é dada por

$$\pi_0(q, \boldsymbol{\beta}, \sigma) = \pi_1(\boldsymbol{\beta}, \sigma | q) \pi_2(q) \propto \frac{\pi_2(q)}{\sigma}$$

onde $\pi_2(q)$ é uma constante com $0 < a \leq q \leq b$.

Para calcular a densidade a priori não informativa para β_q e τ_0 , consideramos o jacobiano da transformação $\sigma = \tau_0^{-1/2}$, então temos:

$$\pi_1(\beta, \tau_0|q) = \pi_1(\beta, \sigma|q) \left| \frac{d\tau_0^{-1/2}}{d\tau_0} \right| = \tau_0^{1/2} \frac{1}{2} \tau_0^{-3/2} = \tau_0^{-1}$$

Portanto, a densidade não informativa para q , β_q e τ_0 é dada por:

$$\pi_0(q, \beta, \tau_0) \propto \pi_1(\beta, \tau_0|q) \pi_2(q) \propto \frac{1}{\tau_0} \quad (4.8)$$

Sem a suposição de independência dos parâmetros (Pericchi, 1981 ; Box e Tiao, 1974), obtemos:

$$\pi_0(q, \beta, \tau_0) \propto \frac{1}{2} \quad (4.9)$$

Observamos que (4.9) nem sequer é uma densidade e optamos pelo uso da densidade a priori não informativa proposta por Box e Cox.

Podemos escrever a densidade conjunta a posteriori, combinando a função de verossimilhança (4.7) com a densidade conjunta a priori (4.8):

$$\pi(q, \beta, \tau_0|\mathbf{D}) \propto \tau_0^{\frac{n-1}{2}-1} \exp\left\{-\frac{\tau_0}{2}[B_q(\beta) + C_q]\right\} \quad (4.10)$$

O estimador bayesiano do vetor de parâmetros $\gamma = (q, \beta, \tau_0)$ é aquele que minimiza o valor esperado a posteriori de uma função perda denotado por $E_\pi[u(\gamma, \hat{\gamma})]$. Se considerarmos a função de perda quadrática, temos $\hat{\gamma} = E_\pi[\gamma]$. O cálculo deste valor esperado com relação à densidade a posteriori (4.10) só pode ser feito por aproximação numérica e, portanto, usamos os algoritmos amostrador de Gibbs e Metropolis-Hastings. Para a aplicação desses algoritmos, é necessário o cálculo das densidades condicionais a posteriori $\pi(\tau_0|q, \beta, \mathbf{D})$ e $\pi(\beta|q, \tau_0, \mathbf{D})$, que podem ser obtidas diretamente de (4.10) e são dadas por:

$$\tau_0|q, \beta, \mathbf{D} \sim Gama\left[\frac{n-1}{2}, \frac{1}{2}[B_q(\beta) + C_q]\right] \quad (4.11)$$

$$\beta|q, \tau_0, \mathbf{D} \sim Normal\left[\hat{\beta}_q, \tau_0^{-1}(\mathbf{X}^{(q)'} \mathbf{X}^{(q)})^{-1}\right] \quad (4.12)$$

A falta de informação a priori sobre o parâmetro q dificulta a escolha de um núcleo de transição para gerar amostras desse parâmetro no algoritmo Metropolis-

Hastings. Supondo-se que $0 < a \leq q \leq b$, uma alternativa é fazer a reparametrização a seguir:

$$\phi = \log\left(\frac{q-a}{b-q}\right) \quad (4.13)$$

onde temos que $\phi = -\infty$ se $q = a$ e $\phi = +\infty$ se $q = b$.

Considerando ϕ com uma densidade a priori gaussiana com média $(a+b)/2$ e variância ξ^2 , podemos usar esta densidade como núcleo de transição, no algoritmo Metropolis-Hastings, para gerar amostras de ϕ . Então, podemos escrever a densidade condicional de ϕ como:

$$\pi(\phi|\tau_0, \boldsymbol{\beta}, \mathbf{D}) \propto N((a+b)/2, \xi^2)\psi(\tau_0, \boldsymbol{\beta}, q, \mathbf{D}) \quad (4.14)$$

com $\psi(\tau_0, \boldsymbol{\beta}, q, \mathbf{D}) = \exp\left\{-\frac{\tau_0}{2}[B_q(\boldsymbol{\beta}) + C_q]\right\}$.

A amostra do parâmetro q pode ser obtida com os valores gerados de ϕ , usando-se a inversa da transformação (4.13) dada por $q = (a + be^\phi)/(1 + e^\phi)$. As amostras geradas com essas densidades condicionais através dos algoritmos de simulação de Monte Carlo (amostrador de Gibbs e Metropolis-Hastings) podem ser usadas para construir amostras das densidades marginais a posteriori dos parâmetros de interesse τ , ρ e K , usando-se as relações $\tau = q^2\tau_0$, $\rho = \beta_0$ e $K = (-\beta_0/\beta_1)^{1/q}$.

Com essas amostras, temos condições de avaliar o comportamento dos parâmetros com estimativas pontuais, intervalos de credibilidade, histogramas e outras estatísticas de interesse como a moda a posteriori, por exemplo. A convergência dos algoritmos pode ser avaliada com o critério de convergência de Gelman e Rubin (Gelman e Rubin, 1992).

4.3.1 Algoritmo Amostrador de Gibbs em conjunto com o Algoritmo Metropolis-Hastings

A seguir apresentamos os passos dos algoritmos de simulação de Monte Carlo em cadeias de Markov denominados amostrador de Gibbs e Metropolis-Hastings.

1. Atribuir valores iniciais para as variáveis $\tau_0^{(0)}$, $\boldsymbol{\beta}^{(0)}$. Escolher os valores a e b , gerar $\phi \sim N(\frac{a+b}{2}, \xi^2)$ onde a variância ξ^2 também é escolhida arbitrariamente e calcular $q^{(0)} = (a + be^{\phi^{(0)}})/(1 + e^{\phi^{(0)}})$;

Faça $k = 0$

2. Gerar valores das distribuições condicionais (4.11) e (4.12) de forma que:

$$\tau_0^{(k+1)} \text{ é gerado de } P(\tau_0|q^{(k)}, \boldsymbol{\beta}^{(k)}, \mathbf{D})$$

$$\boldsymbol{\beta}^{(k+1)} \text{ é gerado de } P(\boldsymbol{\beta}|q^{(k)}, \tau_0^{(k+1)}, \mathbf{D})$$

3. Gerar um valor candidato para ϕ com o núcleo da densidade condicional (4.14) de forma que

$$\phi^{(c)} \text{ é gerado de } N((a + b)/2, \xi^2);$$

4. Testar o valor candidato $\phi^{(c)}$ de acordo com :

Calcular $q^{(c)} = (a + be^{\phi^{(c)}})/(1 + e^{\phi^{(c)}})$, $\widehat{B}_{q^{(c)}}$, $C_{q^{(c)}}$ e a razão:

$$\alpha = \min \left\{ 1, \frac{\psi(\tau_0^{(k+1)}, \boldsymbol{\beta}^{(k+1)}, q^{(c)}, \mathbf{D})}{\psi(\tau_0^{(k+1)}, \boldsymbol{\beta}^{(k+1)}, q^{(k)}, \mathbf{D})} \right\}$$

$$\text{com } \psi(\tau, \boldsymbol{\beta}, q, \mathbf{D}) = \exp \left\{ -\frac{\tau_0}{2} [B_q(\boldsymbol{\beta}) + C_q] \right\} q^{n-1}$$

$$B_q(\boldsymbol{\beta}) = (\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_q)(\mathbf{X}^{(q)'} \mathbf{X}^{(q)})(\boldsymbol{\beta} - \widehat{\boldsymbol{\beta}}_q), C_q = (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \widehat{\boldsymbol{\beta}}_q)' (\mathbf{Y}^{(q)} - \mathbf{X}^{(q)} \widehat{\boldsymbol{\beta}}_q)$$

e $\widehat{\boldsymbol{\beta}}_q$ é o estimador de mínimos quadrados de $\boldsymbol{\beta}$ para dado q .

Gerar $u \sim U[0, 1]$, se $u \leq \alpha$, o valor candidato $q^{(c)}$ é aceito e, então, $q^{(k+1)} = q^{(c)}$; caso contrário, $q^{(k+1)} = q^{(k)}$ e a cadeia não se move.

5. Fazer $k \leftarrow k + 1$ repetir os passos 2, 3 e 4. O valor de M (número de iterações) é determinado pela porcentagem de aceitação da variável ϕ (que deve ser no mínimo 25%) e também pelo critério de convergência Gelman e Rubin.

4.4 Escolha de modelos bayesiana - Critério das Densidades Preditivas Ordenadas

As densidades preditivas ordenadas (DPO) podem ser usadas para a escolha entre dois modelos. Com esse método, verificamos qual é o modelo que melhor se ajusta aos dados. Seja $\mathbf{D}_t = (N_0, N_1, \dots, N_t)$, $t \leq T$. Assumindo-se o modelo de Richards, a densidade de $N_{t+1}|N_t$ é dada por:

$$f(N_{t+1}|N_t, \boldsymbol{\beta}, \tau_0, q) = \frac{q}{N_{t+1}} \left(\frac{\tau_0}{2\pi} \right)^{1/2} \exp \left\{ -\frac{\tau_0}{2} (Y_t^{(q)} - \beta_0 - \beta_1 N_t^{(q)})^2 \right\} \quad (4.15)$$

Dada a densidade conjunta a posteriori $\pi(\boldsymbol{\beta}, \tau_0, q|\mathbf{D}_t)$, temos que

$$f(N_{t+1}, \boldsymbol{\beta}, \tau_0, q|\mathbf{D}_t) = f(N_{t+1}|\boldsymbol{\beta}, \tau_0, q, \mathbf{D}_t) \pi(\boldsymbol{\beta}, \tau_0, q|\mathbf{D}_t) \quad (4.16)$$

Então a densidade preditiva de $N_{t+1}|\mathbf{D}_t$ é calculada integrando-se a densidade conjunta dada por (4.16), de forma que:

$$\begin{aligned} f(N_{t+1}|\mathbf{D}_t) &= \int_{\boldsymbol{\beta}} \int_{\tau_0} \int_q f(N_{t+1}, \boldsymbol{\beta}, \tau_0, q|\mathbf{D}_t) dq d\tau_0 d\boldsymbol{\beta} = \\ &= \int_{\boldsymbol{\beta}} \int_{\tau_0} \int_q f(N_{t+1}|\boldsymbol{\beta}, \tau_0, q, \mathbf{D}_t) \pi(\boldsymbol{\beta}, \tau_0, q|\mathbf{D}_t) dq d\tau_0 d\boldsymbol{\beta} \end{aligned} \quad (4.17)$$

Portanto, a densidade preditiva é o valor esperado de $f(N_{t+1}|\boldsymbol{\beta}, \tau_0, q, \mathbf{D}_t)$ com relação à densidade a posteriori:

$$f(N_{t+1}|\mathbf{D}_t) = E_{\pi}[f(N_{t+1}|\boldsymbol{\beta}, \tau_0, q, \mathbf{D}_t)]$$

A solução da integral (4.17) com o modelo de Richards não pode ser obtida analiticamente, e, por isso, usamos a aproximação de Monte Carlo dada por:

$$\hat{f}(N_{t+1}|\mathbf{D}_t) = \frac{1}{J} \sum_{j=1}^J f(N_{t+1}|\boldsymbol{\beta}^{(j)}, \tau_0^{(j)}, q^{(j)}, \mathbf{D}_t) \quad (4.18)$$

$\boldsymbol{\beta}^{(j)}, \tau_0^{(j)}, q^{(j)}, j = 1, \dots, J$ são as amostras dos parâmetros obtidas com os algoritmos de Monte Carlo em Cadeias de Markov (MCMC). O método das densidades preditivas ordenadas consiste em calcular:

$$\hat{c}_h = \hat{f}(N_h|\mathbf{D}_t) \text{ para } h = t+1, \dots, T$$

Assim o critério é baseado na comparação dos valores do produto de \hat{c}_h para cada modelo. Se temos dois modelos, M_1 e M_2 , usados para explicar o comportamento dos dados \mathbf{D}_t , calculamos $\hat{c}(M_i) = \prod_{h=t+1}^T \hat{c}_h$ para $i = 1, 2$ e o modelo que tiver o valor de \hat{c} maior é o que se ajusta melhor aos dados.

4.5 Aplicação da inferência bayesiana para as Amostras 1 e 2

Para fazer a comparação entre as técnicas clássicas e bayesianas, aplicamos a inferência bayesiana aos dois conjuntos de dados apresentados no capítulo 3. Com a

abordagem bayesiana encontramos as densidades a posteriori utilizando a simulação de Monte Carlo em cadeias de Markov, mais especificamente os algoritmos Amostrador de Gibbs e Metropolis-Hastings em conjunto. Nessa simulação, geramos uma cadeia de Markov com 10000 valores para cada parâmetro e descartamos os 1000 valores iniciais como período de aquecimento. Para atingir uma taxa de aceitação de pelo menos 25% no Metropoli-Hastings, dos 9000 valores restantes, selecionamos as amostras finais tomando valores gerados a cada 5 iterações. A convergência dos algoritmos foi diagnosticada usando-se o critério de Gelman e Rubin, com R próximo de 1 para todos os parâmetros. As taxas de aceitação refletem a adequação do algoritmo Metropolis-Hastings (Besag, 1995). Com as amostras selecionadas, construímos os histogramas das densidades a posteriori apresentadas nas Figuras 4.1 e 4.2, para as Amostras 1 e 2, respectivamente.

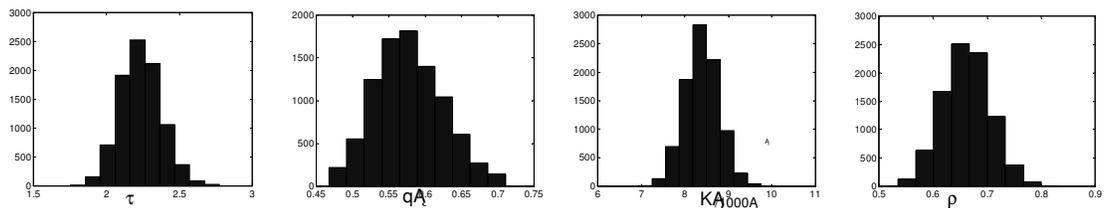


Figura 4.1: Distribuições a posteriori dos parâmetros τ , q , K e ρ para a Amostra 1 ($n = 500$)

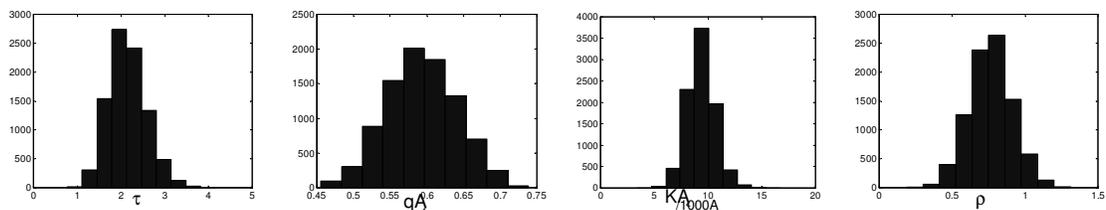


Figura 4.2: Distribuições a posteriori dos parâmetros τ , q , K e ρ para a Amostra 2 ($n = 50$)

A Se observarmos na Figura 4.2 o histograma do parâmetro q para a Amostra 2, vemos que a moda da distribuição está bem próxima do valor verdadeiro, o que não ocorre com a distribuição bootstrap das estimativas de q (Figura 3.5), na qual a moda é 1 e poderia nos levar a escolher o modelo logístico.

Nas Tabelas 4.1 e 4.2, apresentamos os resumos das densidades a posteriori (médias a posteriori e intervalos de credibilidade) para cada parâmetro, o resultado do critério de convergência de Gelman e Rubin (R) e a porcentagem de aceitação do algoritmo Metropolis-Hastings, para as Amostras 1 e 2, respectivamente.

Observando essas tabelas, vemos que as estimativas bayesianas do parâmetro q se aproximam mais do valor verdadeiro tanto na Amostra 1 como na Amostra 2, que é

pequena. Além disso, todos os intervalos de credibilidade contêm os valores verdadeiros dos parâmetros.

Tabela 4.1: Amostra 1 ($n = 500$) $\tau = 2$, $q = 0,6$, $K = 9000$ e $\rho = 0,7$

	Ebayes	Int. Cred.	R	% de aceitação no M-H
τ	2,2382	[1,9656 ; 2,5274]	0,9999	-
q	0,5948	[0,5510 ; 0,6387]	1,0068	37 %
K	8434,6	[7688,6 ; 9206,7]	1,0004	-
ρ	0,6610	[0,5737 ; 0,7478]	1,0002	-

Tabela 4.2: Amostra 2 ($n = 50$) $\tau = 2$, $q = 0,6$, $K = 9000$ e $\rho = 0,7$

	Ebayes	Int. Cred	R	% de aceitação no M-H
τ	2,0163	[1,2935 ; 2,9299]	1,0001	-
q	0,6006	[0,5544 ; 0,6454]	0,9999	32 %
K	9421,4	[6828,5 ; 12197,3]	1,0004	-
ρ	0,7736	[0,4608 ; 1,0879]	1,0005	-

Resultado do Critério das preditivas ordenadas

Para comparar com o método clássico de seleção de modelos, aplicamos o método bayesiano das preditivas ordenadas para as mesmas amostras da Tabela 3.3 do Capítulo 3, página 46. Na Tabela 4.3, temos os resultados com esse critério. Denotando-se por M_1 o modelo logístico e por M_2 o modelo de Richards, escolhemos o modelo de Richards se $\hat{c}(M_1) < \hat{c}(M_2)$ ou seja, se $r_{12} = \frac{\hat{c}(M_1)}{\hat{c}(M_2)} < 1$.

Tabela 4.3: Resultados do critério das preditivas ordenadas

n	r_{12}	Conclusão
50	1,3901	Escolher o modelo logístico
70	1,2642	Escolher o modelo logístico
100	0,9698	Escolher o modelo de Richards
150	0,5524	Escolher o modelo de Richards
200	0,2392	Escolher o modelo de Richards
250	0,1324	Escolher o modelo de Richards
300	0,0308	Escolher o modelo de Richards
500	0,00027	Escolher o modelo de Richards

Esse método se mostrou mais eficiente que o método clássico apresentado no Capítulo 3 para amostras menores. O critério selecionou o modelo correto a partir de amostras de tamanho 100, mas para amostras menores, ele também não é eficiente, basta notar que para a amostra de tamanho $n = 50$, ele seleciona o modelo logístico e os resultados da tabela 4.2 nos levariam a escolher o modelo de Richards. Para ilustrar, temos na Figura 4.3 o gráfico das preditivas ordenadas para $n = 500$.

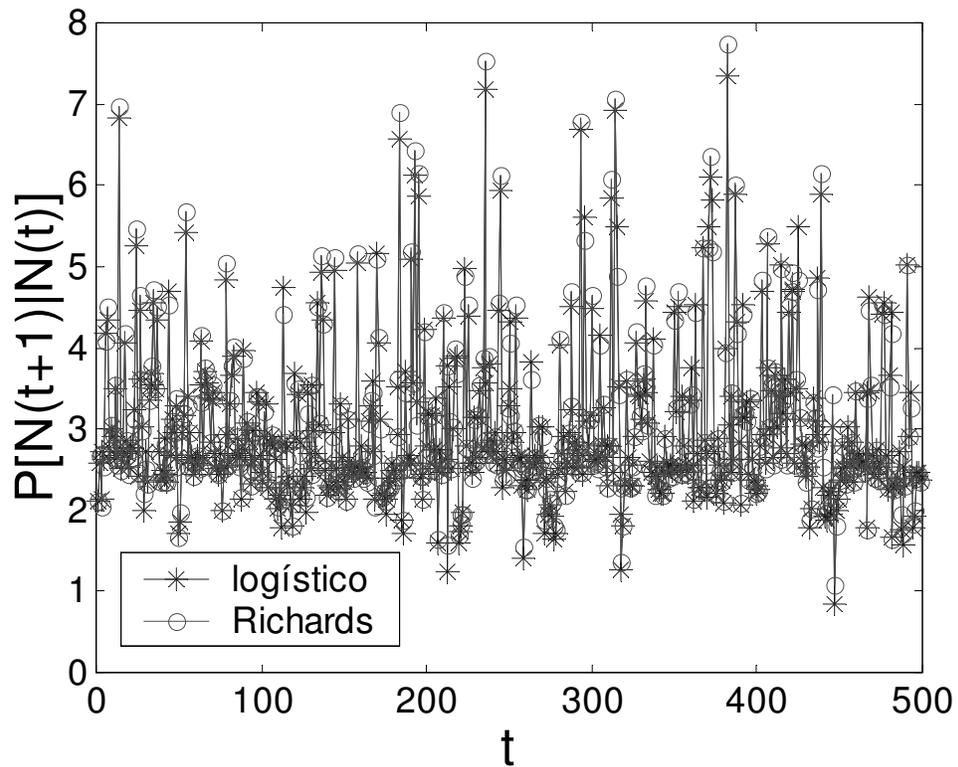


Figura 4.3 : Gráfico das Densidades preditivas ordenadas

No gráfico da Figura 4.3 observamos que os pontos calculados com a densidade preditiva do modelo de Richards são maiores que os do modelo logístico então concluímos que o modelo de Richards é o modelo mais adequado aos dados.

Capítulo 5

Inferência e cálculo de indicadores de viabilidade da população com modelo de Richards com intervenção

Neste capítulo, abordamos o problema de inferência para uma população com crescimento de Richards sujeito a intervenção de retirada de indivíduos numa população. Nesse caso estamos interessados em estimar os parâmetros do modelo conhecendo não somente a trajetória do processo $\{N_t, t \geq 0\}$ mas também a intervenção $\{h_t, t \geq 0\}$, que corresponde a retirada de uma parte dos animais. A intervenção pode ser modelada como um processo estocástico ou como uma função determinística.

Assim como no modelo de Richards sem intervenção, a inferência clássica com este modelo só é possível usando a técnica da máxima verossimilhança *profile*, com a qual estimamos os parâmetros de interesse fixando um deles (vide capítulo 3). A inferência bayesiana para este modelo também só é possível usando simulação de Monte Carlo em cadeias de Markov (vide capítulo 4). Além da inferência apresentamos o cálculo de dois indicadores de viabilidade de uma população. O primeiro indicador é a probabilidade da população atingir uma determinada meta em um determinado horizonte e o segundo é o tempo até uma população atingir uma meta.

5.1 O modelo de Richards com intervenção estocástica

Em muitos casos, a intervenção h_t deve ser modelada como um processo estocástico. Se consideramos que a taxa de abate depende da população disponível, e assumindo-se $\{h_t(N_t), t \geq 0\}$, uma função estocástica, temos o modelo de Richards com intervenção estocástica dado por:

$$N_{t+1} = N_t \exp\left\{\frac{\rho}{q} \left[1 - \left(\frac{N_t}{K}\right)^q\right]\right\} \exp\{\epsilon_t\} - h_t(N_t)N_t \quad (5.1)$$

Um modelo possível para $h_t(N_t)$ é:

$$h_t(N_t) = \phi_0 + \phi_1 N_t + \omega_t \quad (5.2)$$

A equação (5.1) pode ser escrita como:

$$\left[\frac{N_{t+1}}{N_t} + h_t(N_t) \right] = \exp \left\{ \rho \left[1 - \left(\frac{N_t}{K} \right)^q \right] \right\}^{1/q} \exp \{ q \epsilon_t \}^{1/q} \quad (5.3)$$

Elevando ambos os lados da equação à potência q , temos:

$$\left[\frac{N_{t+1}}{N_t} + h_t(N_t) \right]^q = \exp \left\{ \rho \left[1 - \left(\frac{N_t}{K} \right)^q \right] \right\} \exp \{ q \epsilon_t \} \quad (5.4)$$

5.1.1 Inferência clássica para o modelo de Richards com intervenção estocástica

Para o uso do método de máxima verossimilhança *profile*, fixamos o parâmetro q e denotamos $Y_{h,t}^{(q)} = \ln \left[\frac{N_{t+1}}{N_t} + h_t(N_t) \right]^q$, $N_t^{(q)} = N_t^q$ e $z_t = q \epsilon_t$. Então, podemos reescrever a equação (5.3) como uma relação linear:

$$Y_{h,t}^{(q)} = \alpha_0 + \alpha_1 N_t^{(q)} + z_t \quad (5.5)$$

com $\alpha_0 = \rho$ e $\alpha_1 = -\frac{\rho}{K^q}$.

Assumindo que cada ϵ_t é uma variável aleatória gaussiana com média zero e variância τ^{-1} , temos que z_t tem densidade gaussiana com média zero e variância $\tau_0^{-1} = q^2 \tau^{-1}$. Com isso temos, da equação (5.5), que a densidade de $Y_{h,t}^{(q)} | N_t^{(q)}$ é gaussiana com média $\alpha_0 + \alpha_1 N_t^{(q)}$ e variância $\tau_0^{-1} = q^2 \tau^{-1}$.

A inferência para o modelo de Richards com intervenção estocástica requer o conhecimento da densidade conjunta $P(N_{t+1}, h_t(N_t) | N_t)$. Uma vez que $h_t(N_t)$ não é função de N_{t+1} , temos que:

$$P(N_{t+1}, h_t(N_t) | N_t) = P(Y_{h,t}^{(q)} | N_t, h_t(N_t)) P(h_t(N_t) | N_t) J_{YN} \quad (5.6)$$

O jacobiano da transformação $Y_{h,t}^{(q)} = \ln \left[\frac{N_{t+1}}{N_t} + h_t(N_t) \right]^q$ é dado por:

$$J_{YN} = \left| \frac{dY_{h,t}^{(q)}}{dN_{t+1}} \right| = \frac{q}{N_{t+1} + h_t(N_t) N_t}$$

Então

$$P(N_{t+1}, h_t(N_t) | N_t) \propto \tau_0^{\frac{1}{2}} \exp \left\{ -\frac{\tau_0}{2} (Y_{h,t}^{(q)} - \alpha_0 - \alpha_1 N_t^{(q)})^2 \right\}$$

$$\times \frac{qP(h_t(N_t)|N_t)}{[N_{t+1} + h_t(N_t)N_t]} \quad (5.7)$$

Assumindo que ω_t é *i.i.d.* $N(0, \eta^{-1})$, com o modelo (5.2) temos $P(h_t(N_t)|N_t) \sim N(\phi_0 + \phi_1 N_t, \eta^{-1})$ ou seja:

$$P(h_t(N_t)|N_t) \propto \eta^{\frac{1}{2}} \exp\left\{-\frac{\eta}{2}(h_t(N_t) - \phi_0 - \phi_1 N_t)^2\right\} \quad (5.8)$$

Substituindo (5.8) em (5.7), temos:

$$\begin{aligned} P(N_{t+1}, h_t(N_t)|N_t) &\propto \frac{q \tau_0^{\frac{1}{2}} \eta^{\frac{1}{2}}}{[N_{t+1} + h_t(N_t)N_t]} \exp\left\{-\frac{\tau_0}{2}(Y_{h,t}^{(q)} - \alpha_0 - \alpha_1 N_t^{(q)})^2\right\} \\ &\times \exp\left\{-\frac{\eta}{2}(h_t(N_t) - \phi_0 - \phi_1 N_t)^2\right\} \end{aligned} \quad (5.9)$$

Função de verossimilhança *profile*

Em notação matricial temos $\mathbf{h} = (h_1(N_1), \dots, h_n(N_n))$, $\mathbf{Y}_{\mathbf{h}}^{(q)} = (Y_{h,1}^{(q)}, \dots, Y_{h,n}^{(q)})$, $\mathbf{N}^{(q)} = (N_1^{(q)}, \dots, N_n^{(q)})$ e $\mathbf{X}^{(q)} = [\mathbf{1} \ \mathbf{N}^{(q)}]$ com $q > 0$, $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$, $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \tau_0)$, $\boldsymbol{\phi} = (\phi_0, \phi_1)$ e $\boldsymbol{\gamma} = (\boldsymbol{\phi}, \eta)$. Simplificamos a notação com $\mathbf{D} = [\mathbf{N}^{(q)}, \mathbf{h}]$, então a função de verossimilhança é dada por:

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\gamma}, q | \mathbf{D}) &= \prod_{t=1}^{n-1} P(N_{t+1}, h_t(N_t)|N_t) \\ L(\boldsymbol{\theta}, \boldsymbol{\gamma}, q | \mathbf{D}) &\propto q^{n-1} \tau_0^{\frac{n-1}{2}} \eta^{\frac{n-1}{2}} \exp\left\{-\frac{\tau_0}{2} (\mathbf{Y}_{\mathbf{h}}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha})' (\mathbf{Y}_{\mathbf{h}}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha})\right\} \times \\ &\times \exp\left\{-\frac{\eta}{2} (\mathbf{h} - \mathbf{N} \boldsymbol{\phi})' (\mathbf{h} - \mathbf{N} \boldsymbol{\phi})\right\} \times \frac{1}{\prod_{t=1}^{n-1} [N_{t+1} + h_t N_t]} \end{aligned} \quad (5.10)$$

$$\text{onde } \mathbf{N} = \begin{pmatrix} 1 & N_1 \\ 1 & N_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & N_n \end{pmatrix}.$$

Devido a separabilidade de $L(\theta, \gamma, q | \mathbf{D})$ com relação aos parâmetros θ e γ , podemos escrever a função log verossimilhança como:

$$l(\theta, \gamma, q | \mathbf{D}) = l(\theta, q | \mathbf{D}) + l(\gamma | \mathbf{D})$$

onde

$$\begin{aligned} l(\theta, q | \mathbf{D}) = & \frac{n-1}{2} \ln(\tau_0) - \frac{\tau_0}{2} (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha})' (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha}) + \\ & + (n-1) \ln(q) - \sum_{t=1}^{n-1} \ln \left[N_{t+1} + h_t N_t \right] \end{aligned} \quad (5.11)$$

$$l(\gamma | \mathbf{D}) = \frac{n-1}{2} \ln(\eta) - \frac{\eta}{2} (\mathbf{h} - \mathbf{N}\boldsymbol{\phi})' (\mathbf{h} - \mathbf{N}\boldsymbol{\phi}) \quad (5.12)$$

A matriz de informação de Fisher para $\boldsymbol{\theta}$ e $\boldsymbol{\alpha}$, para um dado q , tem a forma:

$$I_q = \begin{pmatrix} I_{\boldsymbol{\theta}\boldsymbol{\theta}}^{(q)} & 0 \\ 0 & I_{\boldsymbol{\alpha}\boldsymbol{\alpha}} \end{pmatrix} \quad (5.13)$$

A partir desse ponto, o procedimento para a obtenção de estimadores de máxima verossimilhança *profile* e de intervalos de confiança para τ_0 e $\boldsymbol{\alpha}$ dado q é análogo ao feito no capítulo 3, usando agora, como função log verossimilhança *profile*, a função dada em (5.11). Lembrando que aqui temos $Y_{h,t}^{(q)} = \ln \left[\frac{N_{t+1}}{N_t} + h_t(N_t) \right]^q$ e $\mathbf{Y}_h^{(q)} = (Y_{h,1}^{(q)}, \dots, Y_{h,n}^{(q)})$, os estimadores de $\boldsymbol{\alpha}$ e τ_0 com q fixo são dados por:

$$\hat{\boldsymbol{\alpha}}_q = [\mathbf{X}^{(q)'} \mathbf{X}^{(q)}]^{-1} \mathbf{X}^{(q)'} \mathbf{Y}_h^{(q)} \quad (5.14)$$

$$\hat{\tau}_{0q} = \frac{(n-1)}{(\mathbf{Y}_h^{(q)} - \hat{\boldsymbol{\alpha}}_q \mathbf{X}^{(q)})' (\mathbf{Y}_h^{(q)} - \hat{\boldsymbol{\alpha}}_q \mathbf{X}^{(q)})} \quad (5.15)$$

Com $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$ e $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \tau_0)$ os intervalos de confiança assintóticos para $\alpha_0, \alpha_1, \tau_0$ e q são calculados com base nos resultados:

$$(n-1)(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \sim N(0, I_{\boldsymbol{\theta}\boldsymbol{\theta}}^{(q)-1})$$

$$-2[l(q | \mathbf{D}) - l(\hat{q} | \mathbf{D})] \sim \chi_1^2$$

Os estimadores de máxima verossimilhança para ϕ e η são calculados por:

$$\hat{\phi} = (N'N)^{-1}N'h \quad (5.16)$$

$$\hat{\eta} = \frac{n-1}{(h - N\hat{\phi})'(h - N\hat{\phi})} \quad (5.17)$$

Com $\phi = (\phi_0, \phi_1)$ e $\gamma = (\phi, \eta)$ os intervalos de confiança para ϕ_0 , ϕ_1 e η podem ser calculados usando o resultado da teoria assintótica:

$$(n-1)(\gamma - \hat{\gamma}) \sim N(0, I_{\gamma\gamma}^{-1})$$

5.1.2 Inferência bayesiana para o modelo de Richards com intervenção estocástica

Para destacar que, nesta abordagem, os parâmetros são variáveis aleatórias, substituímos a notação do vetor de parâmetros α por β . Iniciamos definindo a densidade a priori conjunta não informativa para $\theta = (\beta, \tau_0)$, $\gamma = (\phi, \eta)$ e q :

$$\pi_0(\theta, \gamma, q) \propto \frac{1}{\tau_0} \frac{1}{\eta} \quad (5.18)$$

Combinando a função de verossimilhança (5.10) com a densidade a priori conjunta (5.18) usando o teorema de Bayes, temos a densidade conjunta a posteriori, dada por:

$$\begin{aligned} \pi(\theta, \gamma, q | D) &\propto L(\theta, \gamma, q | D) \pi_0(\theta, \gamma, q) \\ \pi(\theta, \gamma, q | D) &\propto \tau_0^{\frac{n-1}{2}-1} \exp\left\{-\frac{\tau_0}{2}(\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)}\beta)'(\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)}\beta)\right\} \times \\ & q^{n-1} \prod_{t=1}^{n-1} (N_{t+1} - h_t N_t)^{-1} \eta^{\frac{n-1}{2}-1} \exp\left\{-\frac{\eta}{2}(h - N\phi)'(h - N\phi)\right\} \end{aligned} \quad (5.19)$$

Nesse caso, com densidade a priori conjunta não informativa e assumindo independência entre os parâmetros $\theta = (\beta, \tau_0)$ e $\gamma = (\phi, \eta)$, temos a densidade a posteriori conjunta escrita como:

$$\pi(\theta, \gamma, q | D) = \pi_1(\theta, q | D) \pi_2(\gamma | D) \quad (5.20)$$

com

$$\begin{aligned} \pi_1(\boldsymbol{\theta}, q | \mathbf{D}) &= \tau_0^{\frac{n-1}{2}-1} \exp\left\{ -\frac{\tau_0}{2} (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta})' (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta}) \right\} \\ &\quad \times q^{n-1} \prod_{t=1}^{n-1} (N_{t+1} - h_t N_t)^{-1} \end{aligned}$$

$$\pi_2(\boldsymbol{\gamma} | \mathbf{D}) \propto \eta^{\frac{n-1}{2}-1} \exp\left\{ -\frac{\eta}{2} (h - \mathbf{N}\boldsymbol{\phi})' (h - \mathbf{N}\boldsymbol{\phi}) \right\}$$

Como consequência da independência, os estimadores bayesianos, da mesma forma que na abordagem clássica, podem ser calculados separadamente. Além disso, a densidade a posteriori para $\boldsymbol{\gamma}$ pode ser facilmente integrada para obter as densidades marginais de $\boldsymbol{\phi}$ e η . A densidade marginal de η é uma Gama com parâmetros $(n-1)/2$ e $(h - \mathbf{N}\boldsymbol{\phi})' (h - \mathbf{N}\boldsymbol{\phi})/2$ e a densidade marginal de $\boldsymbol{\phi}$ é uma t -Student bivariada com $n-2$ graus de liberdade.

Assim, os estimadores bayesianos para $\boldsymbol{\phi}$ e η coincidem com os estimadores clássicos dados por (5.16) e (5.17) respectivamente. Podemos combinar essas estimativas com um algoritmo MCMC para calcular os estimadores bayesianos para $\boldsymbol{\beta} = (\beta_0, \beta_1)$, τ_0 e q , considerando as densidades condicionais:

$$\pi(\boldsymbol{\beta} | \tau_0, q, \mathbf{D}) \propto \exp\left\{ -\frac{\tau_0}{2} (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta})' (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta}) \right\} \quad (5.21)$$

$$\pi(\tau_0 | \boldsymbol{\beta}, q, \mathbf{D}) \propto \tau_0^{\frac{n-1}{2}-1} \exp\left\{ -\frac{\tau_0}{2} (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta})' (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta}) \right\} \quad (5.22)$$

$$\pi(q | \boldsymbol{\beta}, \tau_0, \mathbf{D}) \propto \exp\left\{ -\frac{\tau_0}{2} (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta})' (\mathbf{Y}_h^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\beta}) \right\} q^{n-1} \quad (5.23)$$

As densidades condicionais (5.21) e (5.22) são conhecidas e portanto podemos gerar diretamente delas as amostras de $\boldsymbol{\beta}$ e τ_0 , respectivamente. A densidade condicional (5.23) não tem uma forma conhecida e assim como fizemos para o modelo sem intervenção, podemos usar a reparametrização dada no capítulo 4, em (4.13) para gerar as amostras de q usando o algoritmo Metropolis-Hastings.

5.1.3 Cálculo de indicadores da viabilidade com o modelo de Richards com intervenção estocástica

Considerando o modelo de Richards com intervenção estocástica, podemos simular o comportamento da população por um longo período de tempo, gerando possíveis trajetórias para os processos $\{N_t^{(j)}, h_t^{(j)}(N_t^{(j)}), j = 1, \dots, J\}$. Estimamos com as equações (2.22) e (2.23), a probabilidade da população atingir uma determinada meta e o tempo até atingir essa meta, respectivamente.

O cálculo bayesiano de indicadores de viabilidade com intervenção estocástica exige um esforço computacional muito grande. Denotando cada valor gerado com MCMC por $q^{(i)}, \tau^{(i)}, \rho^{(i)}$ e $K^{(i)}, i = 1, \dots, M$, então para cada i , teríamos que gerar uma quantidade grande (por exemplo, 100 mil) de trajetórias da população e calcular uma estimativa do tempo esperado e da probabilidade de atingir a meta. Com essas estimativas, teríamos então as densidades a posteriori para os indicadores, mas para isso seriam rodadas $M \times 100$ mil simulações.

5.2 O modelo de Richards com intervenção determinística

Um exemplo real no qual a intervenção $h_t(N_t)$ pode ser considerada uma função determinística, é o caso em que o governo estabelece as quantidades máximas de abate (por exemplo a pesca de animais de espécies ameaçadas de extinção como a lagosta). Nesse caso, as estimativas dos parâmetros e do risco de extinção são obtidos para uma política de intervenção pré-determinada, de forma mais simples do que nas seções anteriores.

Assumindo-se $\{h_t(N_t), t \geq 0\}$, uma função determinística de N_t , que representa a proporção de animais "retirados" em t , podemos escrever o modelo de Richards com intervenção da mesma forma como em (5.1):

$$N_{t+1} = N_t \exp\left\{\frac{\rho}{q} \left[1 - \left(\frac{N_t}{K}\right)^q\right]\right\} \exp\{\epsilon_t\} - h_t(N_t)N_t$$

A densidade dos dados originais $N_{t+1}|N_t$ é dada por:

$$\begin{aligned} P(N_{t+1}|N_t) &= P(Y_{h,t}^{(q)}|N_t^{(q)})J_{YN} = \\ &= \left(\frac{\tau_0}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau_0}{2} \left(Y_{h,t}^{(q)} - \alpha_0 - \alpha_1 N_t^{(q)}\right)^2\right\} \frac{q}{N_{t+1} + h_t(N_t)N_t} \end{aligned} \quad (5.24)$$

5.2.1 Inferência clássica para o modelo de Richards com intervenção determinística

Escrevendo a função de verossimilhança para o modelo (5.24), podemos calcular os estimadores de máxima verossimilhança *profile* (EMVP) dos parâmetros α_0, α_1 e τ_0 com q fixo.

Em notação matricial temos $\mathbf{h} = (h_1, \dots, h_n)$ $\mathbf{Y}_{\mathbf{h}}^{(q)} = (Y_{h,1}^{(q)}, \dots, Y_{h,n}^{(q)})$, $\mathbf{N}^{(q)} = (N_1^{(q)}, \dots, N_n^{(q)})$ e $\mathbf{X}^{(q)} = [\mathbf{1} \mathbf{N}^{(q)}]$ com $q > 0$ e $\boldsymbol{\alpha} = (\alpha_0, \alpha_1)$. Novamente, simplificamos a notação com $\mathbf{D} = [\mathbf{N}^{(q)}, \mathbf{h}]$. A função de verossimilhança é dada por:

$$L(q, \boldsymbol{\alpha}, \tau_0 | \mathbf{D}) = \left(\frac{\tau_0}{2\pi} \right)^{\frac{n-1}{2}} \exp \left\{ - \frac{\tau_0}{2} (\mathbf{Y}_{\mathbf{h}}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha})' (\mathbf{Y}_{\mathbf{h}}^{(q)} - \mathbf{X}^{(q)} \boldsymbol{\alpha}) \right\} \\ \times \frac{q^{n-1}}{\prod_{t=1}^{n-1} [N_{t+1} + h_t(N_t)N_t]} \quad (5.25)$$

A partir desse ponto, o procedimento para a obtenção de estimadores de máxima verossimilhança *profile* e de intervalos de confiança é análogo ao feito no capítulo 3. Lembrando que aqui temos $Y_{h,t}^{(q)} = \ln \left[\frac{N_{t+1}}{N_t} + h_t \right]^q$ e, portanto, os estimadores de $\boldsymbol{\alpha}$ e τ_0 com q fixo são dados por:

$$\widehat{\boldsymbol{\alpha}}_q = [\mathbf{X}^{(q)'} \mathbf{X}^{(q)}]^{-1} \mathbf{X}^{(q)'} \mathbf{Y}_{\mathbf{h}}^{(q)} \quad (5.26)$$

$$\widehat{\tau}_{0q} = \frac{(n-1)}{(\mathbf{Y}_{\mathbf{h}}^{(q)} - \widehat{\boldsymbol{\alpha}}_q \mathbf{X}^{(q)})' (\mathbf{Y}_{\mathbf{h}}^{(q)} - \widehat{\boldsymbol{\alpha}}_q \mathbf{X}^{(q)})} \quad (5.27)$$

5.2.2 Inferência bayesiana para o modelo de Richards com intervenção determinística

A inferência bayesiana para o modelo com intervenção determinística é análoga àquela feita com o modelo de Richards sem intervenção, uma vez que h_t somente altera o cálculo do vetor $\mathbf{Y}_{\mathbf{h}}^{(q)}$. A distribuição $P(N_{t+1}|N_t)$ é alterada somente no seu parâmetro de escala (média).

5.2.3 Cálculo de indicadores da viabilidade com o modelo de Richards com intervenção determinística

Para o cálculo de indicadores de viabilidade da população com o modelo de Richards, temos que calcular a matriz de probabilidade de transição como mencionado no Capítulo 2.

Cálculo da matriz de transição - Abordagem clássica

Temos o modelo de Richards com intervenção:

$$N_{t+1} = N_t \exp\left\{\frac{\rho}{q} \left[1 - \left(\frac{N_t}{K}\right)^q\right] + \epsilon_t\right\} - h_t(N_t)N_t$$

Denotamos $F(N_t) = \frac{\rho}{q} \left[1 - \left(\frac{N_t}{K}\right)^q\right]$ e reescrevemos o modelo como

$$N_{t+1} = N_t \exp\{F(N_t) + \epsilon_t\} - h_t(N_t)N_t \quad (5.28)$$

$$\frac{N_{t+1}}{N_t} + h_t(N_t) = \exp\{F(N_t) + \epsilon_t\}$$

$$\ln \left[\frac{N_{t+1}}{N_t} + h_t(N_t) \right] - F(N_t) = \epsilon_t$$

Se I_i e I_j são dois intervalos quaisquer do espaço de estados da cadeia E e assumindo-se que $I_j = [A_j, B_j]$, então existe um intervalo $[a_j, b_j]$ associado ao ruído ϵ_t , tal que podemos calcular a matriz de transição de um passo com entradas $P_{ij} = P(N_{t+1} \in I_j | N_t \in I_i)$ com:

$$P_{ij} = P(a_j \leq \epsilon_t \leq b_j) = P\left(a_j \leq \ln \left[\frac{N_{t+1}}{N_t} + h_t(N_t) \right] - F(N_t) \leq b_j\right) =$$

$$P\left(e^{a_j + F(N_t)} \leq \frac{N_{t+1}}{N_t} + h_t(N_t) \leq e^{b_j + F(N_t)}\right) =$$

$$P\left(N_t[e^{a_j + F(N_t)} - h_t(N_t)] \leq N_{t+1} \leq N_t[e^{b_j + F(N_t)} - h_t(N_t)]\right) \quad (5.29)$$

De onde vem as identidades $A_j = N_t[e^{a_j + F(N_t)} - h_t(N_t)]$ e $B_j = N_t[e^{b_j + F(N_t)} - h_t(N_t)]$ dado que $N_t \in I_i$.

Assumindo que ϵ_t é um processo com distribuição $N(0, \tau^{-1})$, podemos calcular P_{ij} com a função de distribuição acumulada gaussiana padrão com:

$$P_{ij} = P(N_{t+1} \in I_j | N_t \in I_i) = \Phi(\tau^{1/2} B_j) - \Phi(\tau^{1/2} A_j) \quad (5.30)$$

Essa forma de calcular a matriz de probabilidades de transição é uma aproximação discreta do processo original, uma vez que consideramos as probabilidades de transição entre os intervalos I_i e I_j , não importando os valores que a população pode assumir dentro dos intervalos. A precisão desse método depende do refinamento da discretização do espaço de estados.

Na presença de intervenção, a matriz de probabilidade de transição depende de h_t , assim esse método só pode ser aplicado com h_t fixo, ou determinístico. Quando adotamos h_t estocástico, o cálculo dos indicadores de viabilidade só pode ser feito por simulação como foi mostrado na seção 1 deste capítulo.

Cálculo da matriz de transição - Abordagem bayesiana

Na abordagem bayesiana as probabilidades P_{ij} dependem dos parâmetros $\theta = (\rho, K, \tau, q)$ que são variáveis aleatórias, assim $P_{ij}(\theta)$ também é uma variável aleatória definida por $P_{ij}(\theta) = P(a_j \leq \epsilon_t \leq b_j | \theta)$ onde a_j e b_j são funções de θ . Para calcular a matriz com as entradas P_{ij} , temos que:

$$P(\epsilon_t \in [a_j, b_j], \theta | \mathbf{D}) \propto P(a_j \leq \epsilon_t \leq b_j | \theta) \pi(\theta | \mathbf{D}) \quad (5.43)$$

onde $\pi(\theta | \mathbf{D})$ é a densidade a posteriori conjunta dos parâmetros. Portanto,

$$P(\epsilon_t \in [a_j, b_j] | \mathbf{D}) = \int_{\theta} P(\epsilon_t \in [a_j, b_j], \theta | \mathbf{D}) d\theta$$

$$P(\epsilon_t \in [a_j, b_j] | \mathbf{D}) = \int_{\theta} P(a_j \leq \epsilon_t \leq b_j | \theta) \pi(\theta | \mathbf{D}) d\theta \quad (5.44)$$

Uma vez que a densidade a posteriori conjunta foi obtida por MCMC gerando-se as cadeias denotadas por $\theta^{(k)}$, $k = 1, 2, \dots, M$, então uma estimativa de P_{ij} , obtida por Monte Carlo, é dada por:

$$\hat{P}_{ij} = \frac{1}{M} \sum_{k=1}^M P(a_j \leq \epsilon_t \leq b_j | \theta^{(k)}) \quad (5.45)$$

A densidade a posteriori para os P_{ij} pode ser construída usando as cadeias geradas para os parâmetros θ . Escrevendo

$$P_{ij}(\theta^{(k)}) = P(a_j \leq \epsilon_t \leq b_j | \theta^{(k)}) \quad (5.46)$$

Assim, temos a cadeia para $P_{ij}(\theta^{(k)})$ substituindo cada elemento da cadeia $\{\theta^{(k)}, k = 1, \dots, M\}$ na equação (5.46).

O valor esperado do tempo de primeira passagem é calculado da mesma maneira. Usando $P_{ij}(\theta)$ temos $T_{ij}(\theta)$ e o estimador bayesiano para o tempo é:

$$\mu_{ij} = E_\theta[E(T_{ij}(\theta) | \theta)] \quad (5.47)$$

Uma estimativa para μ_{ij} usando a cadeia gerada por MCMC é dada por:

$$\hat{\mu}_{ij}(\theta^{(k)}) = \frac{1}{M} \sum_{k=1}^M E[T_{ij}(\theta^{(k)}) | \theta^{(k)}] \quad (5.48)$$

Da mesma forma como foi feito para $P_{ij}(\theta^{(k)})$, teremos uma cadeia para $\mu_{ij}(\theta^{(k)}) = E[T_{ij}(\theta^{(k)}) | \theta^{(k)}]$ substituindo cada elemento da cadeia $\{\theta^{(k)}, k = 1, \dots, M\}$ na equação (5.48).

Capítulo 6

Aplicação utilizando dados reais: Avaliação do crescimento do rebanho bovino brasileiro

A população bovina mundial tem uma tendência de estabilização principalmente nos grandes países produtores de carne. No entanto, o rebanho bovino brasileiro apresentou um franco crescimento nos últimos anos, pois com posição privilegiada para este crescimento no hemisfério ocidental, o país apresenta condições ótimas para a exploração de uma pecuária ecológica em pastagens naturais ou cultivadas. Nossas dimensões territoriais e os impressionantes números do rebanho brasileiro tornam-se ainda mais importantes quando pensamos no potencial socioeconômico da cadeia produtiva da carne bovina. Em termos de produção de carne, o Brasil fechou o ano 1999 como segundo produtor mundial, sendo superado somente pelos EUA (Pineda, 2000).

Um estudo para saber a chance (probabilidade) da população bovina ultrapassar determinado patamar, dada uma política de abate predeterminada, e em quanto tempo essa meta seria atingida, é de extremo interesse para a bovinocultura e para a economia brasileira como um todo.

Nesta aplicação, propomos uma avaliação do crescimento do rebanho bovino brasileiro com base nos dados do número efetivo de cabeças e abate anual no período de 1983 até 2002 de acordo com os dados em (ANUALPEC,2003). Podemos avaliar este crescimento por meio do cálculo de alguns indicadores da viabilidade da população bovina brasileira.

Na figura 6.1 temos o gráfico do número de cabeças de gado bovino brasileiro de 1983 a 2002 e na Tabela 6.1, apresentamos os dados do número de cabeças, abate e taxas de abate anuais.

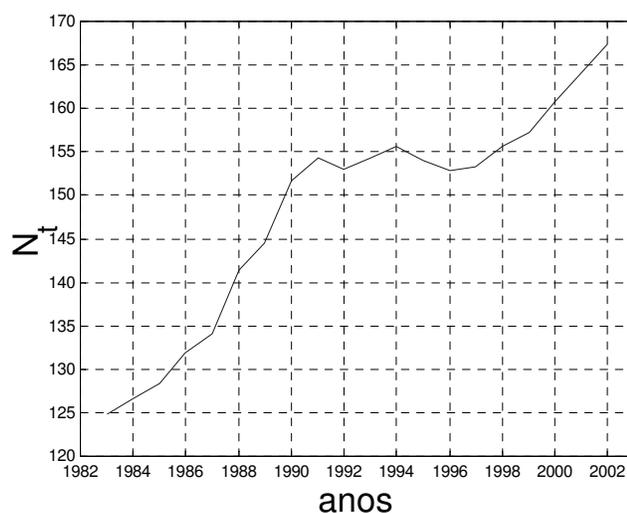


Figura 6.1 : Rebanho bovino brasileiro (milhões)

Tabela 6.1 : Dados ANUALPEC 2003

Ano	População (número de cabeças)	Abate (número de cabeças)	Taxa de abate
1983	124.798.475	20.019.453	0,1604
1984	126.528.255	20.662.727	0,1633
1985	128.362.982	19.514.118	0,1520
1986	131.870.251	21.596.101	0,1479
1987	134.084.435	24.137.350	0,1800
1988	141.360.698	23.947.724	0,1694
1989	144.442.495	24.160.035	0,1673
1990	151.719.702	28.750.883	0,1895
1991	154.297.163	29.856.501	0,1935
1992	153.033.849	29.530.016	0,1929
1993	154.261.655	33.320.517	0,2160
1994	155.650.830	32.686.674	0,2100
1995	154.058.176	36.819.904	0,2390
1996	152.835.009	36.680.402	0,2400
1997	153.236.729	34.171.790	0,2230
1998	155.623.621	34.548.443	0,2220
1999	157.259.739	34.754.402	0,2210
2000	160.764.649	35.207.458	0,2190
2001	164.099.008	36.265.880	0,2210
2002	167.412.089	37.835.132	0,2260

6.1 Resultados com inferência clássica

Como podemos observar na Figura 6.1, o rebanho brasileiro está em franco crescimento, o que significa que ainda não temos dados suficientes para estimar a capacidade de suporte desta população. Optamos, pois, por estipular um valor alto para este nível e, então, estimar os outros parâmetros de interesse. Há uma simplificação nos cálculos para inferência com esse parâmetro fixo como pode ser visto no apêndice, no final da tese. O valor $K = 400$ milhões foi sugerido por um especialista da EMBRAPA - Pecuária Sudeste. Na Tabela 6.2 apresentamos as estimativas de máxima verossimilhança de q e ρ para alguns valores de K . Como essas estimativas não variam muito com K , fixamos o valor de K em 400 milhões e na Tabela 6.3 apresentamos as estimativas pontuais e por intervalo desses parâmetros e de τ .

Tabela 6.2: Estimativas pontuais de q e ρ variando o valor de K

Parâmetro	$K = 170$ milhões	$K = 250$ milhões	$K = 350$ milhões
q	37,5	36,8	36,1
ρ	7,1976	7,0633	6,9865

Tabela 6.3: Estimativas clássicas de τ , q e ρ com $K = 400$ milhões (fixo)

Parâmetro	EMVP	Int. de Confiança assintótico
τ	2118,8	[374,6405 ; 3862,9595]
q	35,9	[2,7 ; 37,5]
ρ	6,9	[6,8874 ; 6,9126]

Na Figura 6.2, temos a simulação de uma população com o modelo de Richards com abate real (Tabela 6.1) e as estimativas dos parâmetros apresentados na Tabela 6.3, assumindo o valor inicial (ano de 1983) para o ajuste. Podemos, então, comparar com o crescimento real e verificar se o modelo está adequado para esse tipo de população. Os valores ajustados com o modelo de Richards estão na Tabela 6.4, utilizando a previsão de um passo. Com essas estimativas, obtivemos que a taxa média de crescimento do rebanho bovino brasileiro é dada por $\bar{r} = \frac{1}{20} \sum_{t=1}^{20} r_t$, $r_t = \ln \left[\frac{N_{t+1} + N_t h_t}{N_t} \right] = 0,1931$.

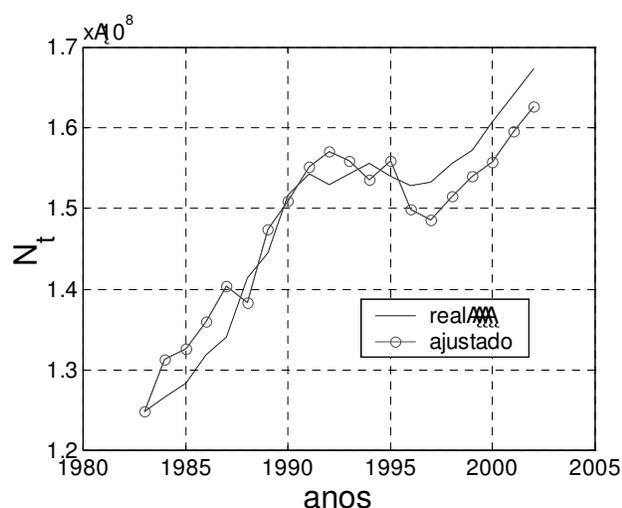


Figura 6.2: Comparação dos dados reais com a simulação usando modelo de Richards

Os resultados da Tabela 6.3 e a simulação do crescimento do rebanho bovino com o modelo de Richards mostram que esse modelo é adequado. Como o parâmetro de forma q está muito longe de 1, se ajustássemos o modelo logístico, estaríamos cometendo um erro grande.

Tabela 6.4: Valores ajustados com o modelo de Richards com a taxa de abate real

Ano	Valores Reais	Valores Ajustados	Ano	Valores Reais	Valores Ajustados
1983	124.798.475	124.798.475	1993	154.261.655	155.940.000
1984	126.528.255	131.230.000	1994	155.650.830	153.630.000
1985	128.362.982	132.680.000	1995	154.058.176	155.950.000
1986	131.870.251	136.050.000	1996	152.835.009	149.890.000
1987	134.084.435	140.310.000	1997	153.236.729	148.540.000
1988	141.360.698	138.360.000	1998	155.623.621	151.540.000
1989	144.442.495	147.370.000	1999	157.259.739	154.050.000
1990	151.719.702	150.890.000	2000	160.764.649	155.830.000
1991	154.297.163	155.120.000	2001	164.099.008	159.630.000
1992	153.033.849	157.140.000	2002	167.412.089	162.610.000

Considerando que a meta desejada para o rebanho bovino é de 180 milhões de cabeças em 2015, o cálculo da probabilidade de atingir essa meta foi feito com um modelo de cadeia de Markov usando-se as estimativas pontuais clássicas da seguinte forma: Primeiro, discretizamos o espaço de estados em 40 intervalos, dados na Tabela 6.5. Calculamos a probabilidade de, em 13 anos, o rebanho atingir os estados

$N_t \geq 180.000.000$ (o que corresponde a atingir o conjunto $\bigcup_{j=24}^{40} I_j$), a partir do estado inicial $N_t = 167.412.089$ (o que corresponde ao intervalo I_{19}).

O estado inicial é a população no ano de 2002 e como queremos calcular a probabilidade de atingir a meta em 13 anos (ano de 2015), temos que calcular a matriz P^{13} . Essa matriz foi calculada para várias taxas de abate com a equação recursiva de Chapman-Kolmogorov. Para obter a probabilidade de atingir a meta, basta tomar o somatório das colunas 24 até 40 na linha 19 da matriz P^{13} :

$$P(N_{13} \in \bigcup_{j=24}^{40} I_j | N_1 \in I_{19}) = \sum_{j=24}^{40} P(N_{13} \in I_j | N_1 \in I_{19}).$$

Tabela 6.5 : Intervalos do espaço de estados

I_1	$N_t \leq 124.800.000$	I_{21}	[172.400.000 ; 174.900.000]
I_2	[124.800.000 ; 127.300.000]	I_{22}	[174.900.000 ; 177.410.000]
I_3	[127.300.000 ; 129.810.000]	I_{23}	[177.410.000 ; 179.920.000]
I_4	[129.810.000 ; 132.310.000]	I_{24}	[179.920.000 ; 182.420.000]
I_5	[132.310.000 ; 134.820.000]	I_{25}	[182.420.000 ; 184.930.000]
I_6	[134.820.000 ; 137.320.000]	I_{26}	[184.930.000 ; 187.430.000]
I_7	[137.320.000 ; 139.830.000]	I_{27}	[187.430.000 ; 189.940.000]
I_8	[139.830.000 ; 142.340.000]	I_{28}	[189.940.000 ; 192.440.000]
I_9	[142.340.000 ; 144.840.000]	I_{29}	[192.440.000 ; 194.950.000]
I_{10}	[144.840.000 ; 147.350.000]	I_{30}	[194.950.000 ; 197.450.000]
I_{11}	[147.350.000 ; 149.850.000]	I_{31}	[197.450.000 ; 199.960.000]
I_{12}	[149.850.000 ; 152.360.000]	I_{32}	[199.960.000 ; 202.460.000]
I_{13}	[152.360.000 ; 154.860.000]	I_{33}	[202.460.000 ; 204.970.000]
I_{14}	[154.860.000 ; 157.370.000]	I_{34}	[204.970.000 ; 207.470.000]
I_{15}	[157.370.000 ; 159.870.000]	I_{35}	[207.470.000 ; 209.980.000]
I_{16}	[159.870.000 ; 162.380.000]	I_{36}	[209.980.000 ; 212.480.000]
I_{17}	[162.380.000 ; 164.880.000]	I_{37}	[212.480.000 ; 214.990.000]
I_{18}	[164.880.000 ; 167.390.000]	I_{38}	[214.990.000 ; 217.490.000]
I_{19}	[167.390.000 ; 169.890.000]	I_{39}	[217.490.000 ; 220.000.000]
I_{20}	[169.890.000 ; 172.400.000]	I_{40}	$N_t > 220.000.000$

Além do cálculo da probabilidade de atingir a meta de 180 milhões de cabeças em 2015, também calculamos qual é o tempo esperado até atingir essa meta para vários valores de taxa de abate.

Segundo Pineda (2000), a taxa de abate do rebanho brasileiro no ano 1980 era de 15% e, no ano 2000 passou a ser de aproximadamente 22% e, se o Brasil continuar com este crescimento, o país poderá atingir um rebanho bovino de 200 milhões de cabeças com abate de 40% anual em 2015. Como a nossa estimativa de crescimento médio anual é de 19,31%, concluímos que uma taxa de abate fixa acima desse valor resultaria em uma diminuição e não em um aumento do rebanho, portanto a taxa de abate de 40%, citada por Pineda, não é realista. Com essa taxa de abate o rebanho se extingiria.

O cálculo da probabilidade de atingir a meta de 180 milhões de cabeças de gado em 2015 e do tempo esperado até atingir essa meta foi feito para alguns valores de taxa de abate fixos, usando o método de cadeia de Markov e também usando simulação de Monte Carlo. Na Figura 6.3a), apresentamos o gráfico das probabilidades e na Figura 6.3b) temos os tempos esperados, em função da taxa de abate. Os resultados numéricos estão na Tabela 6.6.

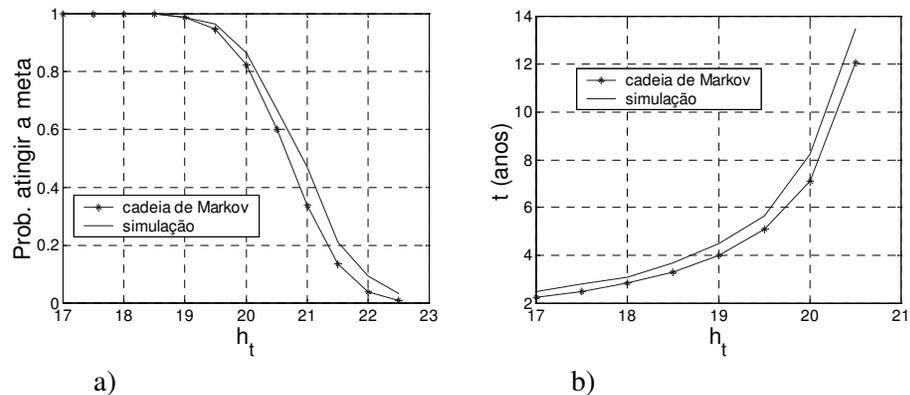


Figura 6.3: a) Probabilidades de atingir a meta, b) Tempos esperados até atingir a meta (em função da taxa de abate)

Tabela 6.6: Probabilidades e tempos com várias taxas de abate

$h_t(N_t)$	Prob.- C.M.	Prob.- S.	Tempo - C.M.	Tempo - S.
17%	1,0000	1,0000	2,2545	2,4820
17,5%	1,0000	1,0000	2,5099	2,7930
18%	0,9999	1,0000	2,8465	3,1040
18,5%	0,9986	0,9990	3,3107	3,6930
19%	0,9893	0,9880	3,9909	4,4850
19,5%	0,9468	0,9640	5,0798	5,6510
20%	0,8245	0,8640	7,0954	8,2130
20,5%	0,6000	0,6690	12,0587	13,4830
21%	0,3360	0,4680	37,4638	47,6550
21,5%	0,1362	0,2110	T_∞	T_∞
22%	0,0382	0,0930	T_∞	T_∞
22,5%	0,0074	0,0320	T_∞	T_∞

Aç.M. - cadeia de Markov e S. - Simulação de Monte Carlo

Comparando os resultados dos dois métodos (simulação de Monte Carlo e cadeia de Markov), assumindo intervenção determinística, verificamos que os valores estão próximos, o que significa que a discretização em 40 intervalos para o cálculo usando cadeia de Markov parece ser suficiente. Para taxas de abate acima de 21%, observamos que o tempo até atingir a meta seria sub-estimado pelos dois métodos e portanto consideramos esses tempos como T_∞ .

Como podemos observar na Figura 6.3a e na Tabela 6.6, a probabilidade do rebanho bovino brasileiro atingir as 180 milhões de cabeças em 2015 com abate anual acima de 22% é praticamente nula. Isso ocorre porque estamos assumindo a taxa de abate constante e, na verdade, essa taxa poderia começar menor e, com o crescimento da população, ela poderia aumentar também. Com abate fixo de 21% o rebanho brasileiro levaria aproximadamente 37 anos para atingir a meta. Com a taxa de abate fixa de 17%, temos a maior probabilidade e o menor tempo para alcançar a meta, mas sabemos que seria um abate muito pequeno, menor do que vem sendo praticado hoje. Se optarmos por um abate anual fixo, esse não deve exceder 20%.

A Observando a taxa de abate real entre os anos de 1983 e 2002, verificamos que ela não é só crescente. Na verdade, a taxa fica oscilando, como podemos ver na Figura 6.4.

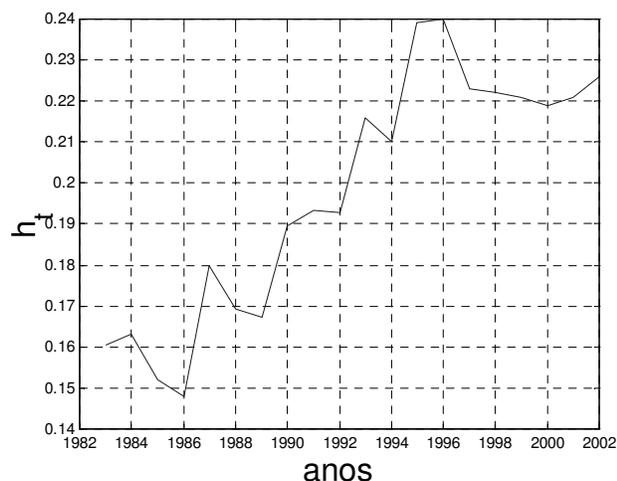


Figura 6.4: Gráfico da taxa de abate real (Tabela 8)

Os resultados anteriores nos mostram que é importante encontrar a taxa de abate ideal, ou seja: Qual deve ser a taxa de abate para que se mantenha a produção desejada e, ao mesmo tempo, se atinja a meta para o rebanho com maior probabilidade? Como foi visto no capítulo 5, podemos considerar a taxa de abate estocástica como uma função linear de N_t , dada por:

$$h_t = \phi_0 + \phi_1 N_t + \omega_t, \text{ onde } \omega_t \text{ é } i.i.d. N(0, \eta^{-1}).$$

Na tabela 6.7 apresentamos as estimativas pontuais e por intervalo dos parâmetros η e $\phi = (\phi_0, \phi_1)$. Na figura 6.5 temos o gráfico da taxa de abate real e simulada com as estimativas da tabela 6.7.

Tabela 6.7: Estimativas clássicas de η , ϕ_0 e ϕ_1

Parâmetro	EMVP	IC assintótico
η	4124,9	[1448,30 ; 6801,44]
ϕ_0	- 0,0969	[- 0,1039 ; - 0,0899]
ϕ_1	0,0020	[0,00194 ; 0,00203]

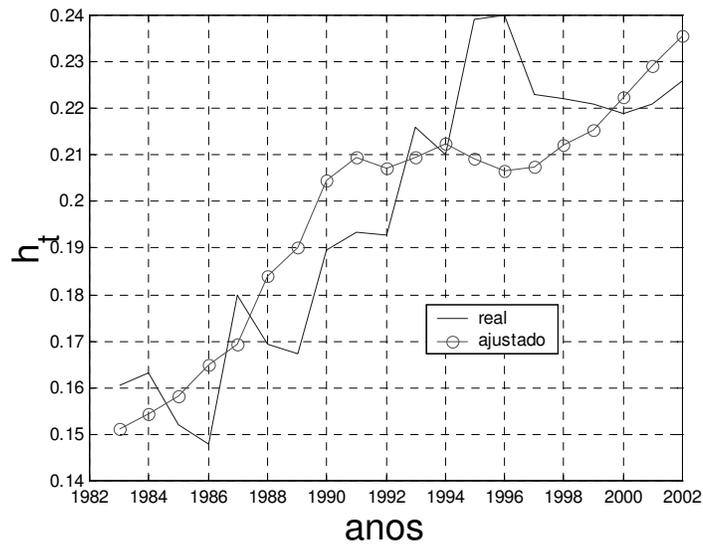


Figura 6.5: Taxas de abate - real e ajustada

Considerando o abate estocástico, calculamos as probabilidades de atingir 180 milhões de cabeças e os tempos para isso baseados em 100 mil trajetórias simuladas. Na figura 6.6 apresentamos o gráfico com algumas trajetórias simuladas para a população.

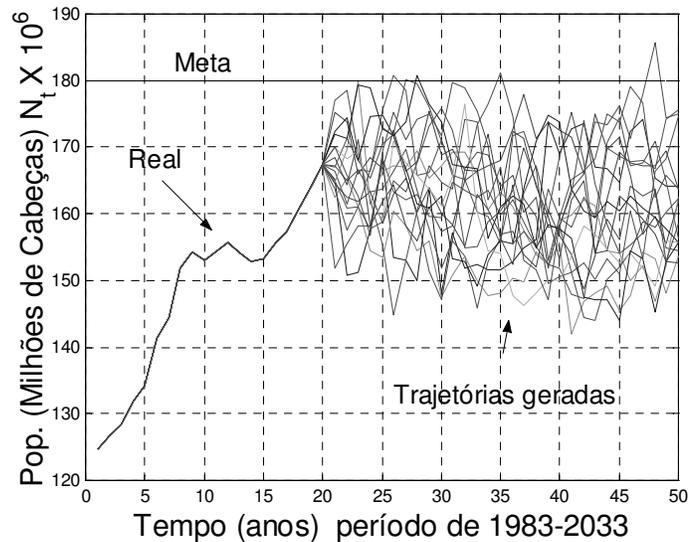


Figura 6.6: Gráfico com algumas trajetórias simuladas da população

Na tabela 6.8 apresentamos as probabilidades e os tempos para atingir a meta, considerando-se cenários onde a taxa de abate foi multiplicada por um fator de redução que varia entre 80% e 100% e, na Figura 6.7 temos os gráficos das probabilidades de atingir a meta e dos tempos esperados em função do fator de redução.

Tabela 6.8: Probabilidades e tempos para atingir a meta

Cenários	Fator de redução	$P(N_t \geq 180 \times 10^6)$	Tempo (anos)
1	80	0,9470	6,6760
2	82	0,8410	8,7670
3	84	0,7020	11,870
4	86	0,4860	17,480
5	88	0,3480	26,730
6	90	0,2130	39,128
7	92	0,1270	55,638
8	94	0,0560	67,654
9	96	0,0170	75,720
10	98	0,0120	77,579
11	100	0,0080	79,106

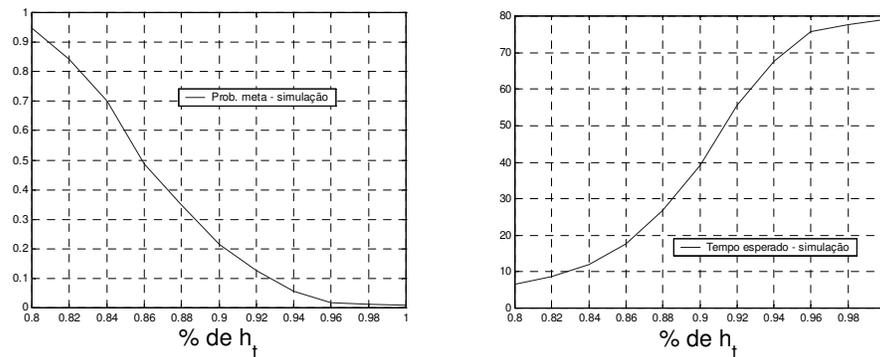


Figura 6.7: Probabilidades de atingir a meta e tempos esperados em função do fator de redução

Considerando 100 mil trajetórias simuladas para $h(N_t)$, a taxa de abate média é próxima de 0,21, na Figura 6.8 apresentamos 5 trajetórias geradas com alguns cenários da tabela 6.8.

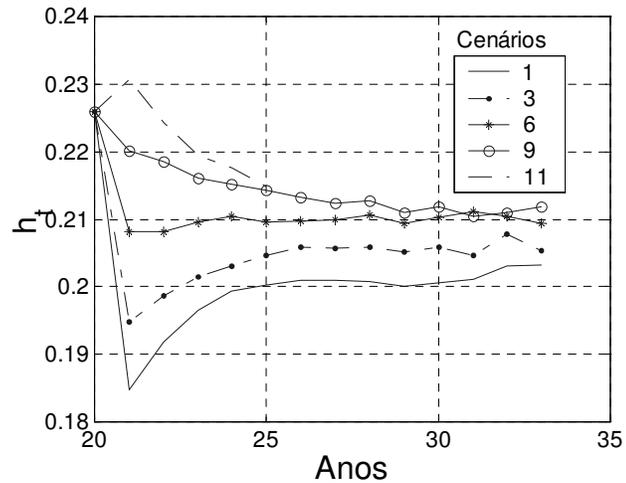


Figura 6.8: Trajetórias simuladas para o abate

As estimativas apresentadas na Tabela 6.8 mostram que ao considerar uma taxa de abate média superior a 0,21 (cenários de 7 a 11) a probabilidade da população atingir a meta estabelecida é mínima, enquanto que considerando uma taxa inferior (cenários de 1 a 4) possibilitamos o crescimento da população e, conseqüentemente, o crescimento da produção de carne bovina para atender a demanda.

A análise desses resultados mostra que, o aumento na taxa de abate, que seria usar 100% de h_t , pode tornar inviável a meta estabelecida, ou seja, essa política de abate não é a ideal.

6.2 Resultados com inferência bayesiana

Com a simulação em cadeias de Markov, geramos 5 cadeias com 10000 valores para cada parâmetro (cada cadeia gerada com um valor inicial diferente), desprezamos os 1000 valores iniciais de cada cadeia, e com os 45000 valores restantes selecionamos 3000 valores (tomando de 15 em 15) como amostras das densidades marginais a posteriori. Asseguramos a convergência com o critério de Gelman e Rubin. Os histogramas dessas densidades estão apresentados na Figura 6.9.

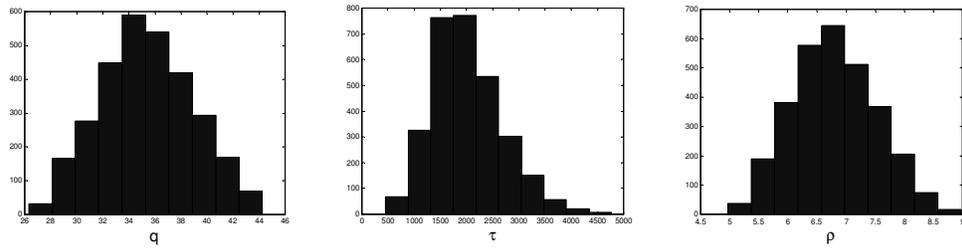


Figura 6.9: Histogramas das densidades a posteriori de q, τ e ρ

Os resultados da inferência bayesiana estão apresentados na Tabela 6.9. Observamos que as estimativas pontuais estão próximas daquelas da Tabela 6.3, mas as estimativas por intervalo podem ser consideradas melhores, devido ao problema do tamanho da amostra ($n = 20$). A convergência foi verificada com o critério de Gelman e Rubin ($R \approx 1$).

Tabela 6.9: Estimativas bayesianas de τ, q e ρ com $K = 400$ milhões (fixo)

Parâmetro	Média	Int. Credibilidade	R	% aceitação (M-H)
τ	2011,5068	[916,1499 ; 3524,1845]	0,9999	-
q	34,9699	[28,7948 ; 42,4359]	1,0072	21,9
ρ	6,7999	[5,4865 ; 8,2251]	1,0075	-

A Com as densidades a posteriori dos parâmetros, construímos as densidades a posteriori para probabilidade de atingir a meta em 2015 e para o tempo até atingir essa meta para várias taxas de abate fixas. Nas figuras 6.10 e 6.12, apresentamos os histogramas das densidades a posteriori desses dois indicadores, respectivamente. Na tabela 6.10, temos os resumos a posteriori para a probabilidade de atingir a meta, em função da taxa de abate e na Figura 6.11 apresentamos esses resultados em um gráfico.

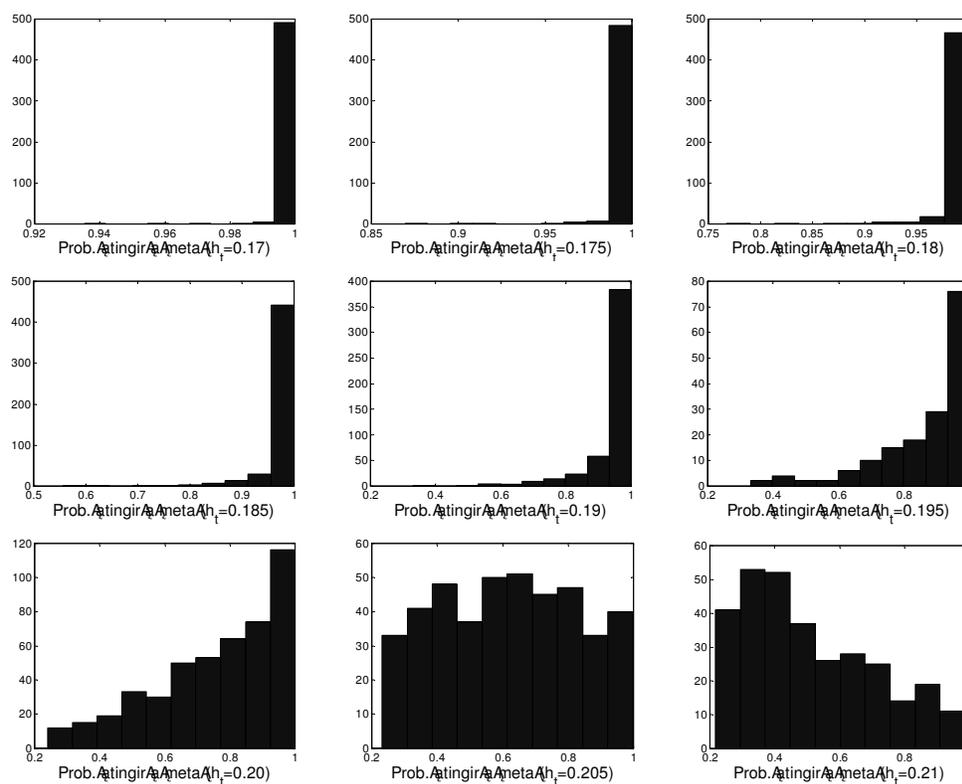


Figura 6.10: Histogramas das densidades a posteriori da probabilidade de atingir a meta em 2015

As densidades na Figura 6.10 são assimétricas, o que indica que a média não é a medida ideal como resumo a posteriori, podemos adotar a moda das densidades como medida.

Tabela 6.10: Indicador bayesiano para a probabilidade de atingir a meta

Taxa de abate	Médias das Prob.	Modas das Prob.	Intervalo de Credibilidade
0,17	0,9994	0,9966	[0,9956 ; 1,0000]
0,175	0,9981	0,9939	[0,9842 ; 1,0000]
0,18	0,9939	0,9873	[0,9469 ; 1,0000]
0,185	0,9813	0,9762	[0,8592 ; 1,0000]
0,19	0,9481	0,9694	[0,8592 ; 1,0000]
0,195	0,8699	0,9645	[0,6803 ; 1,0000]
0,20	0,7576	0,9594	[0,4396 ; 0,9992]
0,205	0,6158	0,6018	[0,3100 ; 0,9935]
0,21	0,5108	0,3309	[0,2576 ; 0,9683]
0,215	0,3354	0,2120	[0,1219 ; 0,8072]
0,22	0,2913	0,1784	[0,1052 ; 0,7511]
0,225	0,2349	0,1180	[0,0697 ; 0,6783]

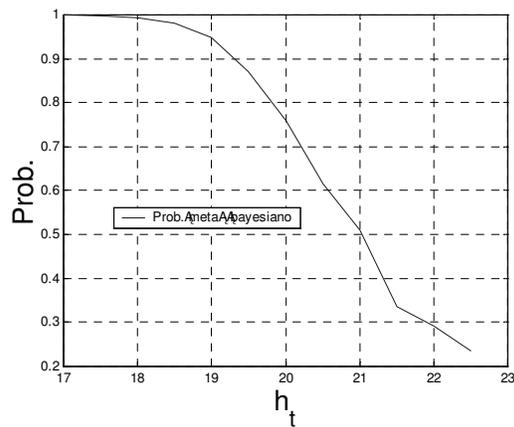


Figura 6.11: Médias das Probabilidades de atingir meta em 2015

Uma taxa de abate fixa de $h_t \geq 21\%$ leva a população a diminuir, assim o tempo esperado para atingir 180 milhões de cabeças tende ao infinito (numericamente encontramos a matriz $(I - Q)$ perto da singularidade e valores de tempo esperado muito grandes: acima de 500 anos).

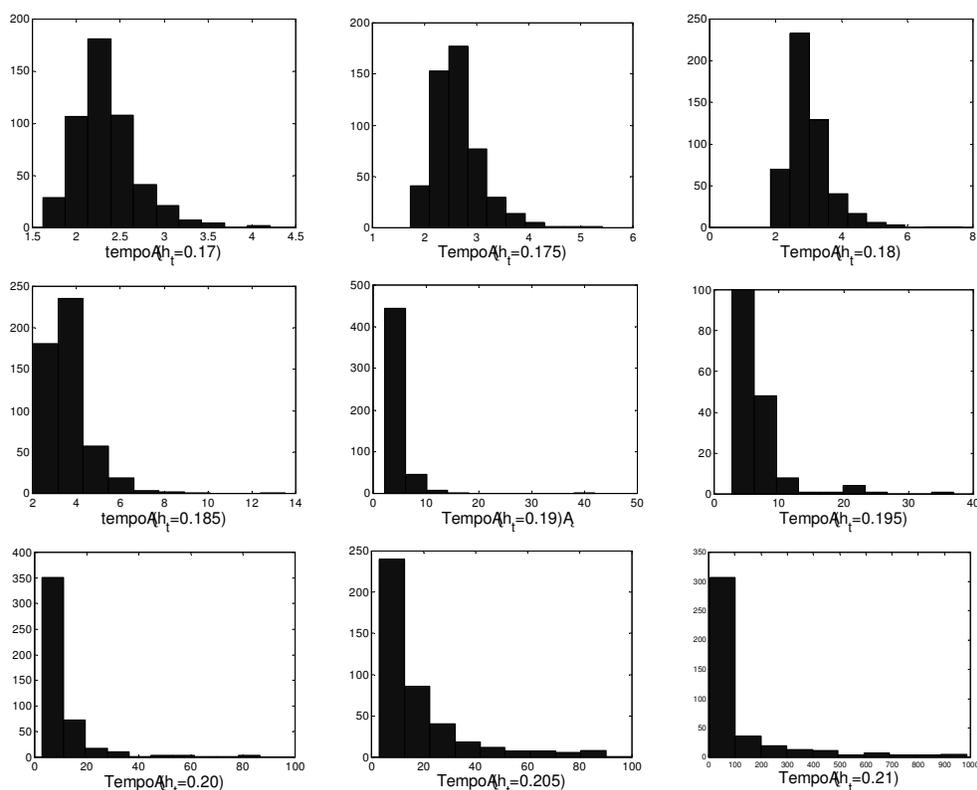


Figura 6.12: Histogramas das densidades a posteriori do tempo até atingir a meta de 180 milhões

Observando a assimetria nas densidades da Figura 6.12, podemos adotar a moda das densidades como medida. Na tabela 6.11, apresentamos os resumos a posteriori para o tempo até atingir a meta de 180 milhões de cabeças em 15 anos, em função da taxa de abate e na figura 6.13 temos um gráfico com esses resultados.

Tabela 6.11: Indicador bayesiano para o tempo até atingir a meta (em anos)

Taxa de abate	Média dos Tempos	Moda dos Tempos	Intervalo de Credibilidade
0,17	2,3394	2,2535	[1,8091 ; 3,1923]
0,175	2,6289	2,6521	[1,9537 ; 3,7788]
0,18	3,0252	2,7419	[2,1301 ; 4,6661]
0,185	3,6094	3,7143	[2,3484 ; 6,2472]
0,19	4,6053	4,0438	[2,6275 ; 9,6666]
0,195	6,6118	4,4240	[3,1087 ; 21,5028]
0,20	10,7090	7,0507	[3,5284 ; 45,3263]
0,205	18,5242	14,9770	[4,2910 ; 74,5184]
0,21	109,7351	50,6912	[5,5740 ; 718,6009]
0,215	T_{∞}	T_{∞}	—
0,22	T_{∞}	T_{∞}	—
0,225	T_{∞}	T_{∞}	—

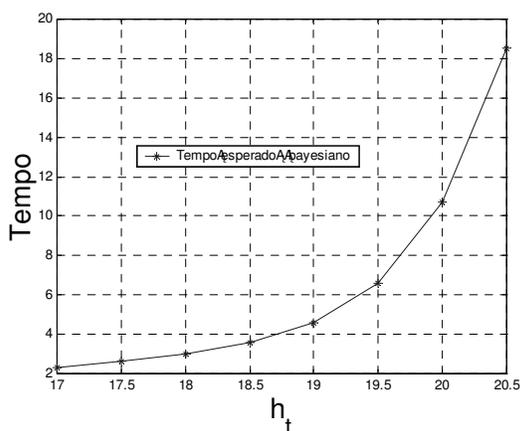


Figura 6.13: Médias dos Tempos em função do abate

A taxa de crescimento da população é aproximadamente dada por $tc = \rho/q$. No caso bayesiano temos uma distribuição a posteriori da taxa de crescimento, o histograma dessa distribuição está na figura 6.14 e na tabela 6.12 apresentamos o estimador de máxima verossimilhança dessa taxa, a média a posteriori e os percentis 90% e 95% para verificar que, realmente uma taxa de abate fixa não deve exceder 20%.

Tabela 6.12: Estimativas para a taxa de crescimento

	EMVP	EBayesiano	percentil 90%	percentil 95%
$tc = \rho/q$	0,1931	0,1922	0,1992	0,2013

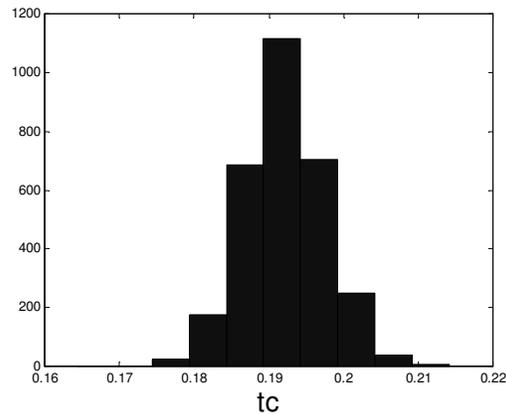


Figura 6.14: Histograma da densidade de ρ/q

A

A Pudemos através do ajuste do modelo com intervenção estocástica e com simulações de Monte Carlo, estimar os indicadores de viabilidade da população.

Considerando a intervenção determinística, calculamos os indicadores assumindo que a população pode ser modelada por uma cadeia de Markov e também usamos simulação de Monte Carlo para comparar os resultados desses dois métodos. A modelagem por cadeia de Markov seria a maneira mais natural para obter estimativas dos indicadores de viabilidade de uma população, entretanto o resultado depende da discretização adotada no espaço de estados. Por outro lado o método usando simulação de Monte Carlo exige que se fixe um horizonte, que neste trabalho chamamos de T_∞ . Essa limitação impede o cálculo preciso do tempo esperado até atingir a meta.

A análise dos resultados numéricos mostra que uma taxa de abate superior a 20% pode dificultar o crescimento da população, e no futuro levar a uma diminuição da produção de carne bovina brasileira. Para atingir uma meta mais elevada, como 200 milhões de cabeças e manter a produção desejada, em um período inferior a 20 anos, torna-se fundamental uma política ótima de abate de forma a se maximizar um certo retorno esperado em um horizonte finito.

Pensando em uma intervenção fixa próxima à taxa média de crescimento $r(N_t) \approx 0,20$ concluímos que a meta de 180 milhões de cabeças seria atingida aproximadamente em 8 anos. Considerando a intervenção estocástica, simulamos 11 cenários variando o fator de redução da taxa de abate $h_t(N_t)$ e concluímos que a meta seria atingida entre 7 e 17 anos (cenários de 1 a 4).

Isso nos leva a crer que utilizar uma intervenção fixa igual 22,6% (abate do ano 2002), não garante a produção sustentável. No cenário atual o que se pode concluir é que o rebanho bovino brasileiro não deve atingir uma meta de 180 milhões de cabeças.

As estimativas pontuais obtidas usando a abordagem bayesiana são próximas às obtidas com a abordagem clássica. A abordagem bayesiana é recomendável para a inferência dos parâmetros pois a informação contida nas densidades a posteriori é maior do que a informação contida nas estimativas pontuais e por intervalo obtidos com os resultados de teoria assintótica. No entanto, para o cálculo dos indicadores de viabilidade da população, encontramos uma limitação na abordagem bayesiana quando consideramos o abate estocástico, que é mais realista que o abate fixo. Como foi visto na sub-seção 5.1.3 do capítulo 5, o cálculo dos indicadores com a abordagem bayesiana e com abate estocástico necessita de grande esforço computacional. Diante deste fato, recomenda-se o uso das estimativas bayesianas pontuais para calcular os indicadores de viabilidade usando simulação de Monte Carlo.

Capítulo 7

Conclusões e propostas futuras

7.1 Conclusões

O problema de crescimento de uma população animal no Brasil ainda é pouco estudado devido a diversas dificuldades. A falta de dados confiáveis e as dificuldades na implementação efetiva de programas de censo e controle são fatores que levam muitos pesquisadores da área a estudar populações nativas de outros países.

Se assumimos que temos dados confiáveis, o problema do cálculo de indicadores da viabilidade de uma população, como a probabilidade de extinção da espécie ou o tempo esperado até a extinção de uma espécie ameaçada, encontra outros tipos de dificuldades, como por exemplo, a escolha de um modelo adequado para os dados e a melhor técnica de inferência para os parâmetros do modelo.

Para essa tese, escolheu-se modelos estocásticos discretos para uma única espécie pela sua aplicabilidade e simplicidade, frente às várias incertezas que envolvem a obtenção de dados sobre populações animais. Existem modelos mais complexos que envolvem mais espécies e, também, que consideram o crescimento a tempo contínuo, mas as dificuldades, do ponto de vista da inferência, são maiores nesse caso. No caso em que o objetivo é o cálculo de políticas de controle da população, também é recomendável ponderar a complexidade do modelo com as dificuldades que este pode apresentar para esses cálculos.

O problema de identificação é fundamental para estudos de risco de extinção, manejo de populações e controle e avaliação da viabilidade da população. No entanto, esse problema se torna complicado quando as amostras são pequenas produzindo estimativas imprecisas. Os pesquisadores da área têm usado os modelos logístico e de Gompertz, por sua simplicidade. Nesta tese, mostramos que, com a transformação de Box e Cox e a técnica de verossimilhança profile, além da técnica de bootstrap para estimativas por intervalo, a inferência clássica é viável para modelos mais complexos como o modelo de Richards.

A idéia é obter ferramentas melhores para a avaliação do crescimento populacional, inferindo com mais precisão a taxa intrínseca de crescimento (ρ) e a

capacidade de suporte da população (K), que são importantes na determinação do risco de extinção.

O uso da teoria assintótica e a abordagem clássica do problema de inferência dos parâmetros do modelo de Richards apresentam algumas dificuldades. O mal condicionamento da matriz de informação de Fisher associada ao modelo acarreta a falta de identificabilidade do parâmetro de forma e assim dificulta muito o cálculo de uma estimativa única. O uso de métodos numéricos como os métodos de Newton e quase-Newton, para maximizar a função de verossimilhança não é possível, pois a matriz hessiana para esse modelo é mal condicionada e próxima da singularidade, o que prejudica a convergência dos métodos.

Para contornar essa dificuldade, interpretamos o modelo de Richards como um modelo logístico quando aplicamos aos dados a transformação de Box e Cox. Neste caso, propomos o cálculo dos estimadores de máxima verossimilhança profile para os parâmetros do modelo de Richards, onde o parâmetro de forma q coincide com o parâmetro da transformação de potência de Box e Cox.

As estimativas por intervalo calculadas com base na teoria assintótica podem ser imprecisas e, portanto, propomos os métodos bootstrap para o cálculo de estimativas por intervalo. A técnica de bootstrap é recomendável, uma vez que, em situações reais, podemos encontrar conjuntos de dados favoráveis para o cálculo da variabilidade das estimativas, e ajustar um modelo mais realista. Isso é mais razoável do que simplesmente adotar o modelo logístico a priori sem uma investigação mais aprofundada e precisa.

No problema de crescimento populacional, a abordagem bayesiana é adequada, uma vez que os parâmetros de interesse ρ e K são variáveis aleatórias e, também, pelo fato que esses estudos contam com poucas observações. Uma das vantagens do método bayesiano é a possibilidade de combinar dados de diversos estudos por meio do uso da densidade a priori e, se isso é possível, não há a necessidade de amostras grandes já que a abordagem bayesiana não se baseia em teoria assintótica.

A abordagem bayesiana pode ser usada com densidade a priori não informativa. Mesmo com densidade a priori imprópria (seção 1, Capítulo 4), com o modelo de Richards verificamos que a densidade a posteriori é própria para um parâmetro de forma finito.

As conclusões baseadas na inferência bayesiana são tomadas com base nas densidades marginais a posteriori para os parâmetros, e isso nos dá muito mais informação do que as estimativas pontuais ou intervalos de confiança da inferência

clássica. Quando não é possível a obtenção das densidades marginais a posteriori a partir da integração da densidade conjunta a posteriori, como é o caso do modelo de Richards, utilizamos os métodos de simulação de Monte Carlo em cadeias de Markov.

Analisando os resultados numéricos obtidos com os dados gerados, vimos que, com a amostra 2 ($n = 50$), obtivemos intervalos de confiança com amplitudes muito grandes e com a amostra 1 ($n = 500$), os intervalos têm amplitudes menores mas, em alguns casos, estão fechados em torno de um valor médio longe dos valores verdadeiros dos parâmetros, de forma que esses valores ficaram fora do intervalo. Isso ocorreu devido à suposição de distribuição simétrica (gaussiana) da teoria assintótica. Esse resultado pode ser melhorado com o uso da técnica bootstrap. Esse mesmo problema aparece no critério de seleção de modelos com base na razão de máxima verossimilhança profile. Com a inferência bayesiana, todos esses problemas são superados, uma vez que se usa a densidade a posteriori dos parâmetros para calcular estimativas pontuais e intervalos de credibilidade.

Nesta tese estendemos o modelo de Richards para um modelo com intervenção, que pode ser uma função determinística ou estocástica do tamanho da população. No caso em que consideramos a intervenção determinística, os indicadores de viabilidade da população podem ser calculados através de uma aproximação por cadeia de Markov, discretizando o espaço de estados. A precisão das estimativas depende dessa discretização, ou seja, o número de intervalos escolhido para dividir o espaço de estados precisa ser suficiente para garantir boa precisão.

No caso de intervenção estocástica, os indicadores só são obtidos com a simulação de um número grande de trajetórias da população. Nesse caso temos maior precisão nas estimativas, mas para o cálculo do tempo esperado até atingir uma meta é necessário fixar um tempo máximo (T_{∞}), de forma que se a população não atingir a meta nesse tempo, considera-se que a meta não foi atingida. Essa forma de calcular o tempo esperado não é muito precisa porque incluindo esse tempo máximo no cálculo da média dos tempos, estamos subestimando o tempo até atingir a meta.

A inferência bayesiana é, portanto, uma alternativa eficiente para o cálculo de estimativas para os indicadores de viabilidade de uma população animal. Sabendo-se que o tempo até atingir uma meta e a probabilidade de atingir tal meta em tempo determinado são variáveis aleatórias, é muito mais razoável ter como estimativa desses indicadores as densidades a posteriori e seus resumos do que um valor médio, que é a única estimativa possível com a inferência clássica.

A Esses procedimentos podem levar a uma mudança de decisão quanto à política de preservação de uma espécie ou prever de forma mais precisa o tempo médio para uma população ser extinta.

7.2 Propostas Futuras

A Nessa seção apresentamos três linhas para pesquisas futuras com modelos de crescimento de populações. Propomos o uso de programação dinâmica estocástica para o cálculo de uma política ótima de abate que garanta a máxima produção sustentável. Uma segunda proposta é o uso de modelos de crescimento heterocedásticos, uma vez que é razoável supor que a variabilidade no crescimento da população depende do tamanho desta. Na abordagem bayesiana, propomos o uso de densidades a priori informativas para a taxa intrínseca de crescimento (ρ) e para a capacidade de suporte (K), pois

sabemos que esses parâmetros são de fácil interpretação biológica o que facilita a obtenção de informações de especialistas da área.

7.2.1 O Problema de controle ótimo de uma população - Formulação por programação dinâmica estocástica

Propomos desenvolver um modelo de otimização que determine a política ótima de abate, de forma que possamos garantir uma taxa de abate crescente com uma probabilidade de crescimento do rebanho aceitável. A solução desse problema pode ser obtida com técnicas de programação dinâmica estocástica, o que permite calcular a política ótima de abate levando-se em consideração o modelo estocástico de crescimento da população. Para que essa modelagem se torne viável, é necessário um grande esforço de pesquisa para elaborar funções-objetivo que retratem realisticamente os objetivos econômicos de abate (lucros) e as metas populacionais desejadas.

Vamos considerar o problema de extração ou abate de indivíduos de uma população em cativeiro. O objetivo é encontrar uma política de extração que atenda uma certa demanda do mercado consumidor, resultando assim, em lucro para o produtor. Mas essa política não pode visar somente o atendimento da demanda porque essa visão pode tornar inviável o crescimento da população ou até mesmo levar a extinção da população. Por isso o produtor deve ter sempre em mente o crescimento da sua criação, estabelecendo metas a serem atingidas em determinado horizonte de planejamento. Por exemplo, um produtor pretende atender a demanda do mercado mas deseja que, no horizonte de 15 anos, sua criação cresça a ponto de atingir um determinado tamanho.

Para formular esse objetivo vamos supor a demanda d_t , $t = 0, 1, \dots, T - 1$, conhecida (ou prevista) e a meta desejada pelo criador dada por $N_T = M$ indivíduos no final do horizonte de planejamento T . Conhecendo a dinâmica de crescimento dessa população, denotada por:

$$N_{t+1} = f(N_t, u_t, \epsilon_t), t = 0, \dots, T - 1$$

onde N_t é o número de indivíduos da população, u_t é o número de indivíduos a ser extraídos da população e ϵ_t é uma variável aleatória *i.i.d.* com distribuição conhecida.

A função que reflete o objetivo do criador em atender a demanda d_t e atingir a meta M ao final do horizonte pode ser escrita como:

$$F_0(N_0) = \sum_{t=0}^{T-1} \left[(d_t - u_t)^2 + (N_t - M)^2 \right] + (N_T - M)^2$$

de forma que encontrando $u_t, t = 0, \dots, T - 1$ que minimiza $F_0(N_0)$ estaremos encontrando a política que se aproxima da demanda e da meta. Ou seja, podemos interpretar o primeiro termo do somatório como uma espécie de erro quadrático entre a demanda e a produção, o segundo termo é o erro quadrático entre o tamanho da população e a meta e o termo $(N_T - M)^2$ é o erro quadrático entre o estado final e a meta.

A solução do problema de programação dinâmica certamente é um desafio para as próximas etapas das pesquisas do problema de controle ótimo de populações. Outras funções objetivo podem ser formuladas de acordo com o problema, por exemplo o estudo de populações selvagens e ameaçadas por intervenções humanas como a exploração imobiliária, o desmatamento, a mineração, a caça e a pesca.

7.2.2 Modelos heterocedásticos para crescimento populacional

Modelos contínuos

Muitos modelos de crescimento podem apresentar além da taxa de crescimento, também a caça (ou abate) como uma função não linear de variáveis aleatórias. O componente do erro também pode ter um comportamento heterocedástico, com variância dependendo do tamanho da população e de outras variáveis aleatórias. Um exemplo usado em Hertzler, 1997 sugere que a caça seja uma função do tamanho da população (N_t), da "catabilidade" (c) e do esforço de caça (E), todos podendo ser tratados como variáveis aleatórias.

$$h_t(c, E, N_t) = cE^a N_t^b$$

onde a e b , são constantes, assim como a taxa de crescimento é função de N_t, ρ, K e q :

$$r(\rho, K, q, N_t) = \frac{\rho}{q} \left(1 - \frac{N_t^q}{K^q} \right)$$

Nesse contexto considerando a equação de crescimento como:

$$N_t = N_0 + \int_0^t r(\rho, K, q, N_t) dt - \int_0^t h_t(c, E, N_t) dt + \int_0^t \sigma(N_t, h_t) d\omega_t$$

Podemos escrever essa equação na forma diferencial como:

$$dN_t = r(\rho, K, q, N_t) dt - h_t(c, E, N_t) dt + \sigma(N_t, h_t) d\omega_t \quad (7.1)$$

Para simplificar, vamos denotar por $r(\beta)$ a mudança esperada na equação (7.1) como função do vetor $\beta = (\rho, K, q, N_t)$ e $\sigma^2(\alpha)$ a variância na equação (7.1), função do vetor $\alpha = (N_t, h_t)$.

Modelos Discretos

A Um modelo de crescimento discreto pode ser escrito como:

$$N_{t+1} = N_t F(N_t, u_t) e^{\epsilon_t}$$

onde $F(N_t, u_t)$ é a taxa de crescimento per capita na presença de um abate u_t e ϵ_t é um processo estocástico independente.

Considerando $Y_t = \log(N_t)$ podemos escrever:

$$Y_{t+1} = Y_t + r(N_t, u_t) + \epsilon_t$$

onde $r(N_t, u_t) = \log[F(N_t, u_t)]$.

Em muitas situações práticas ϵ_t é um processo não correlacionado, $E(\epsilon_t \epsilon_{t+k}) = 0, k \neq 0$, mas heterocedástico $E(\epsilon_t^2) = \sigma_t^{1/2} a_t$ com a_t *i.i.d.* Normal (0, 1).

Um modelo muito presente na literatura é o modelo ARCH (auto-regressivo heterocedástico), neste caso

$$\sigma_t = a_0 + a_1 N_{t-1}^2 \text{ com } a_0 > 0 \text{ e } a_1 > 0$$

Uma generalização desse modelos é o modelo GARCH (ARCH generalizado), dado por:

$$\sigma_t = a_0 + a_1 N_{t-1}^2 + b_1 \sigma_{t-1} \text{ com } a_0 > 0, a_1 > 0 \text{ e } |b_1| < 1$$

Modelos com ordens maiores podem ser encontrados mas em geral os modelos ARCH(1) e GARCH(1,1) são os mais usados.

7.2.3 Abordagem Bayesiana com Densidades A Priori Informativas

Considerando o modelo de Richards, temos:

$$N_{t+1} = N_t \exp\{r(N_t) + \epsilon_t\}$$

com $r(N_t) = \frac{\rho}{q} \left[1 - \frac{N_t^q}{K^q} \right]$, ϵ_t é *i.i.d.* Normal(0, σ^2).

Podemos considerar como densidade a priori informativa para ρ e K , as densidades:

$$\pi(\rho) \sim \text{Gama}(a_0, b_0)$$

$$\pi(K) \sim \text{Normal}(\mu_0, \sigma_0^2)$$

onde a_0 , b_0 , μ_0 e σ_0^2 são hiperparâmetros conhecidos.

Escolhemos a distribuição Gama para ρ por ser um parâmetro positivo, mas não muito maior que zero, em alguns casos (quando $q \approx 1$), temos que $0 < \rho < 1$ e assim uma distribuição Beta pode ser usada. O uso da densidade a priori Normal para ρ é

inapropriado, porque para evitar valores negativos, teríamos que ter uma variância a priori muito pequena.

Para o parâmetro K , a densidade a priori Normal é viável e a partir de informações de especialistas, podemos encontrar os hiperparâmetros μ_0 e σ_0^2 .

Uma densidade a priori informativa para $\tau = 1/\sigma^2$, que é amplamente usada na abordagem bayesiana de modelos lineares, é a densidade Gama, então:

$$\pi(\tau) \sim \text{Gama}(c_0, d_0)$$

onde c_0 e d_0 são hiperparâmetros conhecidos.

O uso de densidades a priori informativas permite que adotemos um critério de seleção de modelos bayesiano, baseado na razão das densidades a posteriori, que é equivalente ao critério clássico da razão de máxima verossimilhança.

A razão de densidades a posteriori é escrita como:

$$B_{12} = \frac{\pi(\theta_1|D, M_1) P(M_1)}{\pi(\theta_2|D, M_2) P(M_2)}$$

onde θ_1 e θ_2 são os vetores de parâmetros dos modelos M_1 e M_2 , respectivamente. D representa os dados. Outro critério que pode ser usados é calcular:

$$P(N|M_i) = \int_{\theta} P(\theta_i, N|M_i) d\theta$$

onde $P(\theta_i, N|M_i) = P(N|\theta_i, M_i)\pi(\theta_i|M_i)$, $P(N|\theta_i, M_i)$ é a função de verossimilhança e $\pi(\theta_i|M_i)$ é a densidade a priori do vetor de parâmetros θ_i do modelo M_i . O modelo M_1 é preferível ao modelo M_2 se:

$$P(N|M_1) > P(N|M_2)$$

A consideração de modelos heterocedásticos torna-se mais difícil porque esses modelos têm muito mais parâmetros e a relação entre eles é muito mais complexa. No entanto, a abordagem bayesiana pode ser mais viável que a abordagem clássica porque podemos utilizar técnicas de simulação de Monte Carlo na inferência e na seleção de modelos.

Referências Bibliográficas

- Arnold, L. (1974). Stochastic Differential Equations: Theory and Applications, Wiley, New York.
- Barndorff-Nielsen, O. E. ; Cox, D. R. (1994). Inference and Asymptotics, Chapman & Hall.
- Besag, J. ; Green, P. ; Higdon, D.; Mengersen, K.(1995). Bayesian Computation and Stochastic Systems. *Statistical Science*, Vol. 10, 1, 3-66.
- Bickel, P. J. ; Doksum, K. A. (1977). Mathematical Statistics. Holden-Day, Inc.
- Bolen, E. G. ; Robinson, W. L. (1999). Wildlife Ecology and Management, Fourth Edition, Prentice Hall, Upper Saddle River, New Jersey.
- Box, G. E.; Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society, série B*, 6,211-252.
- Box, G. E.; Tiao, G. C. (1973). Bayesian Inference in Statistical Analysis, New York: Addison-Wesley.
- Boyce, S. (1992). Population Viability Analysis, *Annu. Rev. Ecol. Syst.* 23 , 481-506B.
- Cahales, L. B. ; Ciria, H. C. ; Peña, O. G. ; Jarque, M. V. ; Rodrigues, R. R. (2001) Efecto antitumoral de la corriente eléctrica directa: estudio en tumores de Ehrlich y fibrosarcoma SA-37. *Revista Cubana Oncología* 7(3), 176-179.
- Çinlar, E. (1975). Introduction to Stochastic Processes, Prentice-Hall, Inc. Englewood Cliffs, New Jersey.
- Chib, S.; Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *American Statistician*, 49,4,327-335.
- Costa, T. M. (2001). <http://www.geocities.com/TimesSquare/Labyrinth/2976>
- Dennis, B.; Munholland, P. L. ; Scott, J. M. (1991). Estimation of Growth and Extinction Parameters for Endangered Species, *Ecological Monographs*, 61(2), 115-143.
- Dennis, B. (1996) Discussion: Should Ecologists become Bayesians?. *Ecological Applications*, 6(4), 1095-1103.

Efron, B. & Tibshirani, R. (1993). An Introduction to the Bootstrap. Chapman & Hall, New York.

Fitzhugh, R. A. H. (1974). Analysis of growth curves and strategies for altering their shapes. *Journal of Animal Science*, 42, 4, pp. 1036-1051

Fraser, D. A. S. (1967). Data transformation and the linear model. *Annals of Math. Statist.*, 38, 4, 456-465.

Fraser, D. A. S. (1968). The Structure of Inference. Wiley, New York.

Gelfand, A. E.; Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85, 398-409.

Gelman, A. E.; Carlin, J.; Stern, H.; Rubin, D. (1995). Bayesian Data Analysis, Chapman & Hall, London.

Gelman, A. E.; Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7, 457-472.

Gilks, W. R.; Richardson, S.; Spiegelhalter, D. A. (1996). Markov Chain Monte Carlo in Practice, Chapman & Hill/CRC.

Hertzler, G.; Herman, J. and Lindner, R. K. (1997). Estimating a stochastic model of population dynamics with an application to kangaroos. *Natural Resource Modeling* 10, 303-343.

Heyman, D. P.; Sobel, M. J. (1982) Stochastic Models in Operations Research, Volume 1: Stochastic Processes and Operating Characteristics, McGraw-Hill Book Company.

Holmes, B. (2001). Estimating risk in declining populations with poor data. *Proceedings of the National Academy of Sciences for the United States of America* 98 (2001) 5072--5077.

Johnson, R. A.; Wichern, D. W., (1992). Applied Multivariate Statistical Analysis, Third Edition, Prentice Hall, Upper Saddle River, New Jersey.

Kass, R. E.; Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association* 90, 430, 773-795.

- Kroll, L. B. ; Tornero, M. T. T. (1994). Modelo de Gompertz com sazonalidade e autocorrelação nos erros para ajuste do crescimento ponderal em vaca leiteira. *Revista Pesquisa Agropecuária Brasileira*, 29(5).
- Lindley, D. V. (1972). Bayesian Statistic, a review. SIAM, Philadelphia, *Regional Conference Series in Applied Mathematics*.
- Ludwig, D. (1996). Uncertainty and the Assessment of Extinction Probabilities. *Ecological Applications* 6(4): 1067-1076.
- Ludwig, D. (1999). Is it meaningful to estimate a probability of extinction? *Ecology* 80:298-310.
- Malthus, T. R. (1798). An Essay on the principle of population, as it effects the future improvement of society with remarks on the speculations of Mr. Godwin, Mr. Condorcet and other writers. J. Johnson, V IX, London.
- Murphy, S. A. ; Van der Vaart, A. W. (2000). On Profile Likelihood, *Journal of the American Statistical Association*, 95, 450, 449-465.
- Nelder, J. A. (1964). Contribution to discussion os paper by G.E.P. Box and D.R. Cox, *J. R. Statistical Society B*, 26, 244-252.
- Odum, E. P. (1988) Ecologia, Editora Guanabara, Rio de Janeiro.
- Omlin, P.; Reichert, A. (1999). Comparison of Techniques for the Estimation of Model Prediction Uncertainty, *Ecological Modelling* 115, 45-59.
- Petriere, M. (1992). A ecologia quantitativa. 10° *Simpósio Nacional de Probabilidade e Estatística - UFRJ*.
- Pericchi, L. R., (1981). A Bayesian approach to transformation to normality, *Biometrika*, 68,1,35-43.
- Pineda, N. R. (2000). Influência do Nelore na produção de carne no Brasil. *Anais do Simpósio Nelore 2000*. 3-13.
- Reichert and M. Omlin, On the Usefulness of Over-parameterized Ecological Model, *Ecological Modelling* 95 (1997) 289--299.
- Richards, J. F. (1959). A flexible growth function for empirical use. *J. Exp. Bot.* 10:290.
- Seber, G. A. F. ; Wild, C. J. (1989). Nonlinear Regression. John Wiley & Sons.

Verdade, L. M. (2004). A exploração da Fauna Silvestre no Brasil: Jacarés, Sistemas e Recursos Humanos (Manuscrito - Comunicação pessoal).

Wehrens, H. Putter and L.M.C. Buydens, The bootstrap: a tutorial; *Chemometrics and Intelligent Laboratory Systems* (Elsevier, 2000).

Publicações resultantes da tese:

Loibel, S. M. C. , Andrade, M. G. , Do Val, J. B. R. Inferência Bayesiana para um Modelo de Crescimento Populacional. V EBEB (Encontro Brasileiro de Estatística Bayesiana), Campinas-SP, 1999.

Loibel, S. M. C. , Do Val, J. B. R., Andrade, M. G. Abordagens Bayesiana e Clássica para o Problema de Inferência com um Modelo de Crescimento Populacional. I COBAL (Congresso Bayesiano da América Latina), Ubatuba-SP, 2002 ; p.87.

Loibel, S. M. C., Do Val, J. B. R., Andrade, M. G. Identificação do Modelo de Crescimento Populacional de Richards. XIV CBA (Congresso Brasileiro de Automática), Natal-RN, 2002; p. 745-750.

Freitas, A. R., Loibel, S. M. C. , Andrade, M. G., Do Val, J. B. R. Crescimento do Rebanho Bovino Brasileiro. Reunião Anual da Sociedade Brasileira de Zootecnia , Santa Maria-RS, 2003; p.40. (SBZ, 2003. CD7p)

Freitas, A. R., Loibel, S. M. C. , Andrade, M. G., Do Val, J. B. R. Crescimento Populacional do Rebanho Bovino Brasileiro. Revista da Sociedade Brasileira Zootecnia, Viçosa-MG ,2003.

Loibel, S. M. C. , Andrade, M. G., Do Val, J. B. R. Critérios de Seleção de Modelos de Crescimento Populacional: Uma abordagem por simulação de Monte Carlo. VII EBEB (Encontro Brasileiro de Estatística Bayesiana), São Carlos-SP, 2004 ; p.33.

Artigo aceito para publicação:

Loibel, S. M. C. , Do Val, J. B. R., Andrade, M. G. Identificação de Modelo de Crescimento Populacional com Intervenção e Cálculo de Indicadores de Viabilidade da População. XV CBA (Congresso Brasileiro de Automática), Gramado-RS, 2004.

Artigos submetidos:

Loibel, S. M. C., Do Val, A. B. R., Andrade, M. G. Inference for the Richards Growth Model using Box-Cox Transformation and Bootstrap Techniques. *Ecological Modelling*, Amsterdam, Netherlands, 2003.

Loibel, S. M. C., Do Val, A. B. R., Andrade, M. G. Inference for the Richards Growth Model subject to interventions and a Population Viability Analysis. *Theoretical Population Biology*, 2004.

Artigos em preparação:

Abordagem Bayesiana para o Modelo de Crescimento de Richards.

Apêndice

Caso Particular : Inferência para o Modelo de Richards com intervenção determinística e K conhecido

Uma aplicação desse caso particular surge em conjunto com h_t determinístico quando a capacidade de suporte é fixada pelo criador, em função do espaço físico disponível para os animais e, a intervenção é determinada de acordo com os interesses pessoais do criador.

Em alguns estudos de populações nos quais temos poucas observações, não é possível estimar a capacidade de suporte (K), simplesmente porque a população não atingiu esse nível no período observado. Quando isso ocorre, podemos obter informações sobre a capacidade de suporte da população em livros, artigos ou com especialistas da área, que muitas vezes fornecem uma boa estimativa. Fixamos então o K nesse valor e estimamos os outros parâmetros.

Inferência clássica com K conhecido

Mesmo com K fixo, é necessário o uso do método de verossimilhança profile. Então fixamos o parâmetro q e denotamos $Y_{h,t}^{(q)} = \ln \left[\frac{N_{t+1}}{N_t} + h_t \right]^q$, $X_t^{(q)} = \left(\frac{N_t}{K} \right)^q$ e $z_t = q\epsilon_t$. Reescrevemos a equação (5.2) como:

$$Y_{h,t}^{(q)} = \rho[1 - X_t^{(q)}] + z_t$$

Novamente assumimos que z_t tem densidade Normal com média zero e variância $\tau_0^{-1} = q^2\tau^{-1}$.

Com isso, temos que a densidade de $Y_{h,t}^{(q)}|N_t^{(q)}$ é Normal com média $\rho[1 - X_t^{(q)}]$ e variância τ_0^{-1} . Então, a densidade dos dados originais $N_{t+1}|N_t$ é dada por:

$$P(N_{t+1}|N_t) = \left(\frac{\tau_0}{2\pi} \right)^{1/2} \exp \left\{ - \frac{\tau_0}{2} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2 \right\} \frac{q}{N_{t+1} + h_t N_t}$$

A função de verossimilhança profile neste caso é dada por:

$$L_p(\rho, \tau | \mathbf{D}) = \left(\frac{\tau_0}{2\pi}\right)^{(n-1)/2} \exp\left\{-\frac{\tau_0}{2} \sum_{t=1}^{n-1} (y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}])^2\right\} \frac{q^{n-1}}{\prod_{t=1}^{n-1} [N_{t+1} + h_t N_t]}$$

Para facilitar os cálculos, trabalhamos com o logaritmo da função de verossimilhança, dado por:

$$l_p = \left(\frac{n-1}{2}\right) [\ln(\tau_0) - \ln(2\pi)] - \frac{\tau_0}{2} \sum_{t=1}^{n-1} (y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}])^2 + \\ + (n-1) \ln(q) - \sum_{t=1}^{n-1} \ln[N_{t+1} + h_t N_t]$$

O procedimento para a obtenção de estimadores de máxima verossimilhança profile é análogo ao feito no capítulo 3, usando-se como função log verossimilhança profile, a função l_p e os estimadores para α e τ_0 pra q fixo dados por:

$$\hat{\tau}_{0q} = \frac{(n-1)}{\sum_{t=1}^{n-1} (y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}])^2}$$

$$\hat{\rho}_q = \frac{\sum_{t=1}^{n-1} y_{h,t}^{(q)} (1 - X_t^{(q)})}{\sum_{t=1}^{n-1} (1 - X_t^{(q)})^2}$$

O cálculo de intervalos de confiança para os parâmetros ρ e τ_0 é feito com base na propriedade de normalidade assintótica dos estimadores, dada por $\hat{\theta} \sim N[\theta, I^{-1}(\hat{\theta})]$, onde $\theta = [\rho, \tau_0]$ e $I(\hat{\theta})$ é a matriz de informação de Fisher observada. Denotando-se os elementos da diagonal da matriz $I^{-1}(\hat{\rho}, \hat{\tau}_0)$ por $v(\hat{\rho})$ e $v(\hat{\tau}_0)$, os intervalos assintóticos com $100(1-\alpha)\%$ de confiança para ρ e τ_0 são dados por:

$$\hat{\rho} \pm z_\alpha \sqrt{v(\hat{\rho})}, \quad \hat{\tau}_0 \pm z_\alpha \sqrt{v(\hat{\tau}_0)}$$

onde z_α é o α -ésimo percentil da distribuição gaussiana padrão.

Inferência bayesiana com K conhecido

Para encontrar a densidade a posteriori conjunta dos parâmetros do modelo com intervenção, reescrevemos a função de verossimilhança:

$$L(q, \rho, \tau_0 | \mathbf{D}) = \left(\frac{\tau_0}{2\pi} \right)^{\frac{n-1}{2}} \exp \left\{ - \frac{\tau_0}{2} \sum_{t=1}^{n-1} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2 \right\} \frac{q^{n-1}}{\prod_{t=1}^{n-1} [N_{t+1} + h_t N_t]}$$

Novamente, optamos pelo uso da densidade a priori não informativa de Jeffreys dada por:

$$\pi_0(q, \rho, \tau_0) \propto \pi_1(\rho, \tau_0 | q) \pi_2(q) = \frac{1}{\tau_0}$$

com $\pi_2(q)$ constante, $0 < a \leq q \leq b$.

A densidade conjunta a posteriori, combinando-se a função de verossimilhança com a densidade a priori, é:

$$\pi(q, \rho, \tau_0 | \mathbf{D}) = L(q, \rho, \tau_0 | \mathbf{D}) \pi_0(q, \rho, \tau_0)$$

$$\pi(q, \rho, \tau_0 | \mathbf{D}) \propto \tau_0^{\frac{n-1}{2}-1} \exp \left\{ - \frac{\tau_0}{2} \sum_{t=1}^{n-1} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2 \right\} \frac{q^{n-1}}{\prod_{t=1}^{n-1} [N_{t+1} + h_t N_t]}$$

As densidades condicionais a posteriori $\pi(\tau_0 | q, \rho, \mathbf{D})$ e $\pi(\rho | q, \tau_0, \mathbf{D})$ podem ser obtidas diretamente da densidade conjunta a posteriori e são dadas por:

$$\pi(\tau_0 | q, \rho, \mathbf{D}) \propto \tau_0^{\frac{n-1}{2}-1} \exp \left\{ - \frac{\tau_0}{2} \sum_{t=1}^{n-1} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2 \right\}$$

$\pi(\tau_0 | q, \rho, \mathbf{D})$ é uma densidade Gama com parâmetros $\frac{n-1}{2}$ e $\frac{1}{2} \sum_{t=1}^{n-1} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2$.

$$\pi(\rho | q, \tau_0, \mathbf{D}) \propto \exp \left\{ - \frac{\tau_0}{2} \sum_{t=1}^{n-1} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2 \right\}$$

Para identificar qual é a densidade $\pi(\rho | q, \tau_0, \mathbf{D})$, reescrevemos na forma matricial. Denotando $G_t = 1 - X_t^{(q)}$ e escrevendo o EMVP de ρ na forma matricial, temos:

$$\hat{\rho} = (G_t' G_t)^{-1} G_t' G_t$$

Então, temos que o somatório na expressão de $\pi(\rho|q, \tau_0, \mathbf{D})$ em forma matricial é escrito como:

$$\sum_{t=1}^{n-1} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2 = (\mathbf{Y}_h^{(q)} - \rho \mathbf{G}_t)' (\mathbf{Y}_h^{(q)} - \rho \mathbf{G}_t) = \epsilon' \epsilon + (\hat{\rho} - \rho)' \mathbf{G}_t' \mathbf{G}_t (\hat{\rho} - \rho)$$

Portanto $\pi(\rho|q, \tau_0, \mathbf{D})$ pode ser escrita como uma densidade gaussiana com média $\hat{\rho}$ e variância $(\mathbf{G}_t' \mathbf{G}_t)^{-1} \tau_0^{-1}$.

$$\pi(\rho|q, \tau_0, \mathbf{D}) \propto \exp \left\{ - \frac{\tau_0}{2} (\hat{\rho} - \rho)' \mathbf{G}_t' \mathbf{G}_t (\hat{\rho} - \rho) \right\}$$

Para gerar amostras de q , considerar novamente a reparametrização:

$$\phi = \log \left(\frac{q - a}{b - q} \right)$$

$$\text{com } \begin{cases} \phi = -\infty & \text{se } q = a \\ \phi = +\infty & \text{se } q = b \end{cases}$$

Considerando ϕ com uma densidade a priori gaussiana com média $(a + b)/2$ e uma variância ξ^2 , esta densidade é usada como núcleo de transição, no algoritmo Metropolis-Hastings, para gerar amostras de ϕ . Então, podemos escrever a densidade condicional de ϕ como:

$$\pi(\phi|\rho, \tau_0) \propto N((a + b)/2, \xi^2) \psi(\rho, \tau_0, q)$$

onde:

$$\psi(\rho, \tau_0, q) = \exp \left\{ - \frac{\tau_0}{2} \sum_{t=1}^{n-1} \left(y_{h,t}^{(q)} - \rho[1 - X_t^{(q)}] \right)^2 \right\}$$

A amostra do parâmetro q pode ser obtida com os valores gerados de ϕ , usando-se a inversa da transformação, dada por $q = (a + be^\phi)/(1 + e^\phi)$.