**Inferência Bayesiana em Modelos de Dinâmica de
Populações Biológicas com Termo de
Perturbação Assimétrico**

Carlos Patricio Montenegro-Silva

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
DOUTOR EM CIÊNCIAS

Programa: Estatística
Orientador: Prof. Dra. Márcia D'Elia Branco

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

São Paulo, Janeiro de 2016

Inferência Bayesiana em Modelos de Dinâmica de Populações Biológicas com Termo de Perturbação Assimétrico

Esta versão da dissertação/tese contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 20/01/2016. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof^a. Dr^a. Márcia D'Elia Branco - IME-USP
- Prof. Dr. Heleno Borfarine - IME-USP
- Prof. Dr. Helio dos Santos Migon UFRJ
- Prof. Dr. Manuel Galea Rojas - PUC-Chile
- Prof. Dr. Reinaldo Boris Arellano-Valle - PUC-Chile

Agradecimientos

Es larga la lista de personas e instituciones a quienes quiero agradecer en estas líneas. En primer lugar, a mis ancestros, mis abuelos, mi madre, mi padre, mis hermanos. Particularmente a Javier, que como hermano mayor y matemático sembró en mí la curiosidad por el álgebra. En mi camino por el Instituto de Fomento Pesquero, es extensa la lista de colegas y amigos con quienes crecimos, estudiamos, aprendimos y nos motivamos por ir siempre un poco más lejos. Jorge Sáteler, Miguel Donoso, Verónica Alegría y Hernán Miranda, son algunos de los colegas de quienes aprendí aspectos técnicos y principalmente humanos. Particularmente mi colega y amigo Cristian Canales, quien fue un gran maestro que me abrió las puertas al mundo del modelamiento estadístico computacional. Largas líneas de los más diversos lenguajes de programación escribimos por más de una década. A los colegas y profesores de la Universidad de Valparaíso, quienes fueron los primeros en permitirme sistematizar y profundizar en el conocimiento de la estadística matemática. Particularmente van mis sinceros agradecimientos a los profesores Ronny Vallejos, Soledad Torres, María Angélica Maulén, Raúl Fierro, Cristian Meza y Manuel Galea, entre otros. En la Universidad de São Paulo, también es extensa la lista de profesores y colegas sin los cuales no estaría escrita esta tesis. Particularmente a los profesores Heleno Bolfarine, Chang Chiann, Silvia Ferrari, Nikolai Koley, Silvia Elían, Sergio Wechler y Carlos Pereira. En especial, van mis agradecimientos a mi profesora orientadora de programa y de tesis, la profesora Márcia Branco. Su rigurosidad conceptual y su calidad humana la hicieron para mí una maestra a quien le estaré agradecido eternamente. Agradezco también a los miembros de la banca de defensa de tesis, quienes con su mirada aguda me mostraron el camino hacia nuevos desafíos de investigación. A los colegas del IME, con quienes pasamos largas horas resolviendo las listas de ejercicios y apoyándonos cuando el camino estuvo difícil. Diego Gallardo, Manuel González, Tiago Magalhães, Leandro Correia, entre otros. También, quiero agradecer a todos mis amigos brasileños, con quienes aprendí a sambar y ellos aprendieron que la cueca es mucho más que una prenda de vestir. A la familia Caggiano, particularmente a Christiane, quien fue mucho más que una gran parceira. Finalmente a mi hija Carla, quien fue, es y será la luz que ilumina mi camino. Te amo hija querida!

Resumo

Neste trabalho de tese, estudamos o modelo de crescimento logístico de populações biológicas utilizando a abordagem de espaço de estados. Os estados não observados são as biomassas anuais, a equação de observação é linear e a equação de estado é não linear. As distribuições de probabilidade utilizadas para os termos de erro de observação aditivos são: Normal, t-student, Skew-normal e Skew-t. As distribuições Log-normal, Log-t, Log-skew-normal e Log-skew-t são consideradas para os erros de observação multiplicativos. A inferência nos modelos é realizada considerando-se métodos Bayesianos e as distribuições *a posteriori* de interesse são aproximadas utilizando-se algoritmos MCMC e a aproximação de Laplace. Apresentamos duas aplicações, a primeira referente a pesca de camarão marinho na costa do Chile, na qual a variável observável é o rendimento médio anual de pesca (captura por unidade de esforço média). Na segunda é considerada a pesca de lagostim vermelho na costa de Chile, na qual além do rendimento médio anual da pesca, observa-se as estimativas anuais de biomassa vulnerável, obtidas através de estudos de área varrida. Para o primeiro conjunto de dados, os modelos com erros de observação multiplicativos têm melhor performance, particularmente os modelos Log-skew-normal e Log-skew-t. Considerando estes resultados, no segundo caso utilizamos somente erros multiplicativos e a distribuição *a posteriori* preditiva mostra que cada variável observável parece ter sua própria família de distribuição de probabilidades. Além disso, os resultados também revelam uma crescente complexidade do modelo ao incorporar a classe mais geral de distribuições assimétricas.

MONTENEGRO-SILVA, C.P. **Inferência Bayesiana em Modelos de Dinâmica de Populações Biológicas com Termos de Perturbação Assimétricos e de Caudas Pesadas**. 2016. 101 f. Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2016.

Palavras-chave: Modelos dinâmicos não-lineares, modelo de crescimento logístico, MCMC, camarão chileno.

Abstract

We study the logistic population growth model using a state-space approach. The non observable states are the annual biomass of the population with a linear observation equation and a non-linear state equation. The probability distribution used for the additives observation error terms are Normal, Student-t, Skew-normal and Skew-t, and Log-normal, Log-t, Log-skew-normal and Log-skew-t for multiplicative observation errors terms. The inference about the parameters of the models is performed using Bayesian methods, with MCMC algorithms and Laplace approximations. We present two applications to real data sets. The first in marine shrimp population off the coast of Chile, where observable variable is the average annual fishing yield. The second application is for the population of the red squat lobster off Chile, where in addition to the average annual fishing yield, a second observable variable was included. In the first case, the multiplicative observational errors models presented the best results. Particularly the Log-skew-normal and Log-skew-t models has the better performances. Considering these results, in the second application we use only multiplicative observation errors models.

MONTENEGRO-SILVA, C.P. **Bayesian inference in biological population dynamic models with skewed and heavy tailed perturbation terms**. 2016. 101 p. Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2016.

Keywords: Nonlinear dynamic models, logistic growth model, MCMC, Chilean shrimp.

Sumário

Lista de Símbolos	vi
Lista de Figuras	viii
Lista de Tabelas	xi
1 Introdução	1
1.1 Motivação	1
1.2 Modelos dinâmicos	2
1.3 Objetivos do trabalho de tese	3
1.3.1 Objetivo geral	3
1.3.2 Objetivos específicos	3
1.4 Organização da tese	3
2 Modelos dinâmicos	4
2.1 Introdução	4
2.2 Modelo espaço de estados linear gaussiano	5
2.2.1 Inferência sequencial	6
2.2.2 Dados faltantes e previsão	8
2.3 Modelo espaço de estados não linear gaussiano	9
2.3.1 Estimação de parâmetros através do Filtro de Kalman Estendido	10
2.4 Modelo espaço de estados não linear e não gaussiano	13
3 Modelo dinâmico logístico	14
3.1 Introdução	14
3.2 Modelo de crescimento exponencial	14
3.3 Modelo de crescimento logístico	16
3.4 Representação via espaço de estados do modelo de crescimento logístico	18
3.5 Algumas extensões	19
4 Distribuições assimétricas e de caudas pesadas	20
4.1 Introdução	20
4.2 Distribuições com suporte em $\mathbb{R}$	20
4.2.1 Distribuição Skew-Normal	20
4.2.2 Distribuição Skew-t	22
4.2.3 Distribuições de Mistura na Escala da Normal Assimétrica (SMSN)	23

4.3	Distribuições com suporte em $\mathbb{R}^+$	24
4.3.1	Distribuição Log-Skew-Normal Univariada	24
4.3.2	Distribuição Log-Skew-t Univariada	25
5	Inferência no modelo dinâmico logístico	27
5.1	Introdução	27
5.2	Inferência via filtro de Kalman estendido (EKF)	27
5.3	Inferência através da aproximação de Laplace	28
5.4	Inferência Bayesiana através de MCMC	30
5.4.1	Equação de estado	31
5.4.2	Modelos Observacionais	31
5.4.3	Aproximação das distribuições <i>a posteriori</i>	33
5.5	Crerios de seleção do modelo	34
6	Estudo de simulação	35
6.1	Introdução	35
6.2	Resultados	37
7	Aplicações a dados reais	47
7.1	Biomassa do camarão da zona central do Chile	47
7.2	Biomassa do lagostim vermelho no Chile	62
8	Discussão e Considerações Finais	69
	Referências Bibliográficas	72
A		80
A.1	Provas proposições	80
A.2	Algoritmos quase-Newton	82
A.3	Diferenciação Automática	84
A.4	Diferenciação implícita	86
B		87
B.1	Códigos computacionais	87
B.1.1	Códigos R: Modelo de simulação	87
B.1.2	Códigos OpenBugs: Aplicação dados reais caso 1	88
B.1.3	Código ADMB	90
B.2	Densidades condicionais completas	93
B.2.1	Modelo com erros de observação e proceso normais	93
B.2.2	Modelo com erros de observação e proceso lognormais	94
B.3	Trazas das cadeidas de Markov	97

Lista de Símbolos

$t = 1, \dots, n$	Conjunto de índices com $n \in \mathbb{N}$
Y_t	Variável aleatória observável a tempo t
$\mathbf{Y}_t$	Vetor aleatório observável a tempo t
X_t	Variável aleatória não observável a tempo t
$\mathbf{X}_t$	Vetor aleatório não observável a tempo t
$\mathbf{y}_{1:n}$	Amostra aleatória de Y_t de tamanho n
$\mathbf{x}_{1:n}$	Amostra aleatória de X_t de tamanho n
$(\mathbf{X}_t)_{t \geq 0}$	Cadeia de Markov não observável $\mathbb{R}^p$ -dimensional
$(\mathbf{Y}_t)_{t \geq 0}$	Série temporal observável $\mathbb{R}^m$ -dimensional
K	Capacidade de carga no modelo logístico
r	Taxa intrínseca de crescimento no modelo logístico
q	Coefficiente de capturabilidade
b	Taxa de nascimentos no modelo de crescimento exponencial
d	Taxa de morte no modelo de crescimento exponencial
H_t	captura ou <i>harvesting</i> no tempo t
τ^2	Variância do erro de processo
σ^2	Variância do erro de observação
Θ	Vetor de parâmetros do modelo logístico
θ	Vetor de parâmetros da distribuição de probabilidades dos erros de observação
(Ω, σ, P)	Espaço de probabilidade
$p(\cdot)$	Função de massa o densidade de probabilidade
u_t	Erros de processo
v_t	Erros de observação
D_t	Conjunto de informação disponível no tempo t
$\mathbf{G}_t$	Matriz de evolução
$\mathbf{F}_t$	Matriz de design
$\mathbf{f}_t$	Função de transição
$\mathbf{h}_t$	Função de observação
$\mathbf{W}_t$	Matriz de variâncias-covariâncias dos erros de processo
$\mathbf{V}_t$	Matriz de variâncias-covariâncias dos erros de observação
MSY	Máximo nível de exploração sustentável
H_{MSY}	Taxa de captura ou <i>harvesting</i> que maximiza os excedentes de produção
ϕ	Função de densidade de probabilidade de uma variável aleatória normal
Φ	Função de distribuição acumulativa de uma variável aleatória normal

t_ν	Função de densidade de probabilidade de uma variável aleatória t-student com ν graus de liberdade
T_ν	Função de distribuição acumulativa de uma variável aleatória t-student com ν graus de liberdade
λ	Parâmetro de forma das variáveis aleatórias skew-normal e skew-t
ν	Graus de liberdade das variáveis aleatórias t-student e skew-t
γ_1	Coefficiente de assimetria
γ_2	Coefficiente de kurtosis

Lista de Figuras

2.1	Representação do processo de atualização da informação no Filtro de Kalman.	8
2.2	Representação do processo de atualização da informação no Filtro de Kalman estendido.	11
3.1	Trajetórias de três populações com crescimento exponencial.	15
3.2	Efeitos denso-dependentes de nascimento e morte de uma população.	16
3.3	Taxa de mudança da população, no modelo de crescimento logístico, em função do tamanho da população	17
3.4	Trajetória da abundância de uma população com crescimento logístico	17
4.1	Funções de densidade da normal assimétrica com valores positivos e negativos de λ	22
4.2	Funções de densidade da skew-t. Esquerda, valores fixos de $\lambda = 5$ e variáveis de ν , direita, valores fixos de $\nu = 5$ e variáveis de λ	23
4.3	Funções de densidade de distribuições na família SMSN.	25
4.4	Densidades de probabilidade das distribuições log-skew-normal com $\mu = 0$, $\sigma = 1$ e diferentes valores de λ	26
6.1	Simulação de estados não observáveis com erro de processo normal. Série temporal de extensão 90 (acima) e simulação de 100 séries de observações com erro skew-t (abaixo).	36
6.2	Estimação de estados não observáveis. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das medias <i>a posteriori</i> e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo normal e erro de observação skew-t. Series temporais n=10, 20, 30, 50, 70 e 90.	39
6.3	Estimação de estados não observáveis, caso 2. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das medias <i>a posteriori</i> e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo normal e erro de observação skew-normal. Series temporais n=10, 20, 30, 50, 70 e 90.	40
6.4	Estimação de estados não observáveis, caso 2. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das medias <i>a posteriori</i> e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo normal e erro de observação t-student. Series temporais n=10, 20, 30, 50, 70 e 90.	41
6.5	Estimação de estados não observáveis, caso 2. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das medias <i>a posteriori</i> e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo e observação normais. Series temporais n=10, 20, 30, 50, 70 e 90.	42

6.6	Estimação de estados não observáveis. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das médias <i>a posteriori</i> e linha sólida valores verdadeiros. Procedimento de estimação via aproximação de Laplace, sob hipótese de erro de processo normal e erro de observação skew-t. Series temporais $n=10, 20, 30, 50, 70$ e 90	46
7.1	Rendimento de pesca (Y_t , toneladas/hora, linha preta sólida) e captura (H_t , milhares de toneladas, linha segmentada azul) anuais de camarão para a zona centro-sul do Chile entre os anos 1981 e 2012.	48
7.2	Diagramas de dispersão das discrepâncias observadas e preditas da autocorrelação dos erros de observação dos modelos aditivos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile.	50
7.3	Diagramas de dispersão das discrepâncias observadas e preditas da autocorrelação dos erros de observação dos modelos multiplicativos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile.	51
7.4	Distribuição <i>a posteriori</i> preditiva do rendimento da pesca para os modelos aditivos propostos (linhas) e rendimentos da pesca observados (pontos) de camarão no Chile.	52
7.5	Distribuição <i>a posteriori</i> preditiva do rendimento da pesca para os modelos multiplicativos propostos (linhas) e rendimentos da pesca observados (pontos) de camarão no Chile.	53
7.6	Estimativa de estados não observáveis (HPD de 95% da biomassa de camarão no Chile) com os modelos com erro de observação aditivos. A linha horizontal mostra a média <i>a posteriori</i> de B_{MSY}	57
7.7	Estimativa de estados não observáveis (HPD de 95% da biomassa de camarão no Chile) com os modelos com erro de observação multiplicativos. A linha horizontal mostra a média <i>a posteriori</i> de B_{MSY}	58
7.8	Estimativa das taxas de pesca (HPD de 95%), com modelos com erros de observação aditivos. A linha horizontal mostra a média <i>a posteriori</i> de H_{MSY} , pescaria do camarão no Chile.	60
7.9	Estimativa das taxas de pesca (HPD de 95%), com modelos com erros de observação multiplicativos. A linha horizontal mostra a média <i>a posteriori</i> de H_{MSY} , pescaria do camarão no Chile.	61
7.10	Série temporal de médio rendimento anual de pesca (acima, linha sólida, Y_{1t}) e biomassa vulnerável estimada em pesquisas de arrasto de fundo (acima, linha pontilhada, Y_{2t}) e capturas total anual do lagostim vermelho na zona centro-sul do Chile (abaixo, H_t).	63
7.11	Percentis da distribuição <i>a posteriori</i> preditiva do rendimento da pesca (Y_{1t}) de lagostim vermelho para os modelos propostos (linhas) e rendimentos da pesca observados (pontos).	64
7.12	Percentis da distribuição <i>a posteriori</i> preditiva do índice de biomassa vulnerável (Y_{2t}) de lagostim vermelho para os modelos propostos (linhas) e biomassa vulnerável observada (pontos).	65
7.13	Estimativa de estados não observáveis (HPD de 95% da biomassa de lagostim vermelho). A linha horizontal mostra a média <i>a posteriori</i> de B_{MSY}	66

B.1	Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dado de camarão da zona central do Chile, com erros de observação log-skew-normal (a).	97
B.2	Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dado de camarão da zona central do Chile, com erros de observação log-skew-normal (b).	98
B.3	Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dado de camarão da zona central do Chile, com erros de observação log-skew-t (a).	99
B.4	Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dado de camarão da zona central do Chile, com erros de observação log-skew-t (b).	100

Lista de Tabelas

6.1	Parâmetros do modelo de simulação para o modelo logístico. O modelo considera erro de processo normal e erro de observação skew-t.	36
6.2	Viés médio (MB) e a raiz do erro quadrático médio ($RMSE$) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação skew-t. . . .	43
6.3	Viés médio (MB) e a raiz do erro quadrático médio ($RMSE$) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação skew-normal. . . .	44
6.4	Viés médio (MB) e a raiz do erro quadrático médio ($RMSE$) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação t-student. . . .	44
6.5	Viés médio (MB) e a raiz do erro quadrático médio ($RMSE$) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação normais. . . .	45
6.6	Viés médio (MB) e a raiz do erro quadrático médio ($RMSE$) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação skew-t. Procedimento de estimação via aproximação de Laplace.	45
7.1	DIC (Deviance information criterion), número efetivo de parâmetros (p_D) e posterior predictive p-value da autocorrelação dos resíduos ($\hat{p}$) dos modelos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile.	49
7.2	Estatísticas da distribuição <i>a posteriori</i> dos parâmetros dos modelos aditivos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile. Se apresenta a medida de redução de escala potencial ($\hat{R}$). . . .	54
7.3	Estatísticas da distribuição <i>a posteriori</i> dos parâmetros dos modelos multiplicativos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile. Se apresenta a medida de redução de escala potencial ($\hat{R}$).	55
7.4	Estatísticas da distribuição <i>a posteriori</i> das quantidades de gestão derivadas do modelo logístico com erros de observação aditivos, população de camarão no Chile	56
7.5	Estatísticas da distribuição <i>a posteriori</i> das quantidades de gestão derivadas do modelo logístico com erros de observação multiplicativos, população de camarão no Chile	59
7.6	DIC (Deviance information criterion), número efetivo de parâmetros (p_D) e posterior predictive p-values da autocorrelação dos erros de observação das variáveis Y_{1t} ($\hat{p}_1$) e Y_{2t} ($\hat{p}_2$) dos modelos propostos para os dados do lagostim vermelho.	64
7.7	Estatísticas da distribuição <i>a posteriori</i> dos parâmetros do modelo de crescimento logístico para lagostim vermelho do Chile.	67

7.8	Estatísticas da distribuição <i>a posteriori</i> das quantidades de gestão derivadas do modelo de crescimento logístico para lagostim vermelho do Chile.	68
A.1	Valor e f em $x = \pi$	84
A.2	Valor da derivada de f em $x = \pi$	84
A.3	Valor e f em $x_1 = \pi, x_2 = 2$	85
A.4	Valor da derivada parcial $\frac{\partial f}{\partial x_1}$ em $x_1 = \pi, x_2 = 2$	85
A.5	Valor e derivadas de f	86

Capítulo 1

Introdução

1.1 Motivação

A pesquisa marinha moderna iniciou-se nos anos 30, quando foram feitos esforços para formalizar o estudo sistemático dos aspectos relacionados com a pesca e otimizar o uso dos recursos marinhos (Russell, 1931; Thompson e Bell, 1934). Apesar de ser uma área relativamente nova da investigação, nos últimos 80 anos tem surgido um grande conjunto de subáreas, entre as quais se destacam a biologia das espécies exploradas, o estudo do ambiente em que vivem, o estudo da dinâmica populacional, o estudo das pescarias multiespecíficas e, mais recentemente, a abordagem ecossistêmica integrada, entre outras (Garcia et al., 2003). Neste contexto, uma subárea que vem sendo desenvolvida há várias décadas pelos países que tem importante atividade pesqueira, é o estudo das populações marinhas com o propósito de encontrar uma forma sustentável de fazer sua exploração. Os esforços desenvolvidos nesta direção vêm permitindo o desenvolvimento de uma área de pesquisa aplicada conhecida como *Stock Assessment*, a qual consiste em modelagem (normalmente estatístico) das populações marinhas (Hilborn e Walters, 1992; Quinn e Deriso, 1999).

A importância social e econômica da pesca é notável em países com grandes Zonas Econômicas Exclusivas de pesca (ZEE). É especialmente importante para os países em desenvolvimento, porque às vezes a pesca tem um valor econômico que equivale à metade de todos os produtos vendidos nesses países (FAO, 2014). Enquanto isso, em países desenvolvidos a pesca gera significativas divisas com exportações de produtos derivados da pesca, como a farinha de peixe, enlatados e congelados. Assim, a pesca e a aquicultura são uma fonte não só de saúde, mas também de riqueza. Nos últimos anos o emprego no setor pesqueiro cresceu mais rapidamente que a população mundial e é a base dos meios de subsistência de centenas de milhões de pessoas no mundo (FAO, 2014).

De acordo com a Organização para a Alimentação e Agricultura das Nações Unidas (FAO), a pesca de captura e aquicultura forneceu cerca de 158 milhões de toneladas de peixe para consumo humano em 2012, proporcionando uma oferta per capita aparente de 19,2 kg (anuais), que está entre os mais altos registros históricos. A produção mundial da pesca de captura em 2012 foi de cerca de 92 milhões de toneladas, sendo China, Indonésia, Estados Unidos, Peru, Rússia, Japão, Índia e Chile os oito principais países produtores da pesca (FAO, 2014).

A FAO acredita que a pesca de captura provavelmente já atingiu o máximo potencial dos oceanos. 80% das unidades populacionais de peixes do mundo avaliadas estão totalmente exploradas ou sobre exploradas e, conseqüentemente, é urgente monitorar sua gestão eficiente. No ano de 2007, cerca de 28% dos estoques monitorados pela FAO ou foram sobre explorados ou esgotados, rendendo menos do que seu máximo potencial, devido ao excesso de pressão da pesca. Assim, considerando o nível de deterioração dos estoques de recursos marinhos, é necessário desenvolver estudos para estimar os níveis de abundância destas populações e seus níveis de exploração sustentável, e é neste ponto que a pesquisa marinha tem um papel fundamental a desempenhar. Um dos principais

objetivos dos estudos das populações marinhas é conhecer seus níveis populacionais, ou seja, sua abundância. Esta abundância pode ser medida em número de indivíduos ou em biomassa (peso total dos indivíduos da população) e, geralmente, é difícil de medir diretamente.

Para o estudo de populações de recursos marinhos biológicos, podem-se utilizar várias abordagens e a escolha do método de estudo dependerá, entre outros fatores, do tipo de população a ser analisada. No caso de espécies com baixa mobilidade (recursos que geralmente vivem perto da costa), podem-se usar métodos de quantificação diretos, através de procedimentos de amostragem para populações finitas em um contexto espacial, tomando-se amostras diretas da população para estimar os níveis relativos (densidades) ou totais de abundância (Seber, 1982; Gunderson, 1993 ; Gallucci et al., 1996).

Para as espécies migradoras, como o atum e o espadarte, entre outros, os métodos indiretos são geralmente utilizados, juntamente com técnicas complementares, para ajudar a atender seus padrões de migração, como a marcação e a telemetria (Hilborn e Walters, 1992; Quinn e Deriso, 1999; Anderson-Sprecher, 1994).

Os métodos indiretos têm a vantagem de ser de baixo custo monetário, dado que normalmente as agências pesqueiras contam com dados que se relacionam com a abundância ou biomassa. Um exemplo é o rendimento da pesca, uma variável que mede uma quantidade (em peso ou unidades) por unidade de esforço de pesca (por exemplo, uma viagem de pesca). Neste trabalho vamos usar um método indireto de estimativa de abundância de recursos marinhos, em condições de pouca informação, o qual será especificado no Capítulo 3.

1.2 Modelos dinâmicos

Métodos indiretos de estimativa da abundância marinha pertencem à classe de modelos dinâmicos, uma vez que o interesse central encontra-se na estimativa de variáveis não observáveis ao longo do tempo. Este tipo de modelos hoje tem um extenso desenvolvimento, particularmente usando a representação de espaço de estados (Kitagawa, 2010; Durbin e Koopman, 2012) .

O processo de inferência pode ser feito num paradigma clássico (freqüentista) ou dentro de um contexto Bayesiano. Cada uma destas abordagens tem seus pontos fortes e fracos e foge do objetivo desta tese a discussão deste aspecto. Aos leitores interessados em aprofundar essa discussão sugerimos a leitura de Bayarri e Berger, 2004; Efron, 2005; Aitkin, 2010, entre outros. Nesta tese utilizaremos uma abordagem Bayesiana de inferência.

Na representação de espaço de estados, a modelagem é feita através de dois componentes: (1) a equação de evolução do sistema (ou equação de estado ou processo) e (2) equação de observação, que relaciona as quantidades observáveis com os estados latentes. Ambas as equações têm componentes estocásticos (erros de processo e observação) e, portanto, as características probabilísticas deles devem ser definidas. Normalmente, para estes componentes aleatórios, os estudos de biologia das populações utilizam a distribuição Normal ou Log-normal, com a hipótese de erros aditivos para o primeiro caso e multiplicativos para o segundo (Millar e Meyer, 2000; de Valpine, 2002; Dennis et al., 2006; Maunder e Deriso, 2011). Neste trabalho, propomos a utilização de uma classe mais geral e flexível de distribuições de probabilidade que permitam acomodar simultaneamente assimetria e observações atípicas. Em particular, propomos o uso da distribuição skew-t, da classe de modelos de mistura na escala da normal assimétrica (scale mixture of skew-normal, SMSN) para os termos de perturbação dos modelos (Branco e Dey, 2001; Lachos et al., 2013).

1.3 Objetivos do trabalho de tese

1.3.1 Objetivo geral

O objetivo geral desta tese é utilizar distribuições de probabilidade flexíveis para modelar os componentes estocásticos do modelo de crescimento logístico de populações, através da abordagem de espaço de estados.

1.3.2 Objetivos específicos

1. Estudar e propor procedimentos de inferência para estimar os parâmetros do modelo de crescimento logístico e estimação de seus estados não observáveis, com termos de perturbação assimétricos.
2. Estudar as propriedades frequentistas dos procedimentos de inferência frente a diferentes extensões das séries temporais de dados.
3. Aplicar os procedimentos de inferência desenvolvidos para conjuntos de dados reais.

1.4 Organização da tese

Esta tese está estruturada na forma descrita a seguir. No Capítulo 2 fazemos uma apresentação detalhada dos modelos dinâmicos através da representação espaço de estados, apresentamos uma visão generalizada dos modelos espaço de estados para usar mais tarde essa representação para o caso de modelos dinâmicos não lineares e não gaussianos. Por fim, apresentamos um conjunto de opções para lidar com a inferência (aproximada) nesta classe de modelos, enfatizando duas técnicas que usamos nesta tese que são a simulação estocástica através dos algoritmos MCMC (Markov chain Monte Carlo) e a aproximação de Laplace.

No capítulo 3, apresentamos em detalhe o modelo de crescimento logístico de dinâmica populacional que é desenvolvido a partir do modelo de crescimento exponencial de populações. Este é o modelo central que é analisado nesta tese e corresponde a equação de evolução do modelo espaço de estado. No Capítulo 4, apresentamos a classe flexível de distribuições de probabilidade SMSN, para os erros aditivos, e classe log-SMSN para os erros multiplicativos.

No Capítulo 5, apresentamos os procedimentos inferenciais utilizados para estimar os parâmetros e estados não observáveis do modelo dinâmico proposto no Capítulo 3. Apresentamos os procedimentos de inferência com base no algoritmo MCMC e aproximação de Laplace. Apresentamos também os critérios de seleção dos modelos, especificamente o DIC (Deviance Information Criterion; Spiegelhalter et al., 2002) e os PPCs (*Posterior Predictives Checks*; Gelman et. al., 2014).

No Capítulo 6, apresentamos um estudo de simulação que esta focado a estudar o desempenho dos modelos propostos diante de diferentes extensões da série temporal dos estados não observáveis ($n=10, 20, 30, 50, 70$ e 90). Analisamos medidas de viés, precisão e exatidão do processo inferencial na estimativa dos parâmetros do modelo de crescimento logístico. No Capítulo 7, apresentamos a aplicação dos modelos propostos para dois estudos de caso de pesca de crustáceos da costa Chilena. O primeiro é a pesca marinha do camarão no centro-sul do Chile e o segundo é a partir da pesca do lagostim vermelho na mesma área. Para o primeiro temos uma equação de observação, sem dados faltantes e no segundo caso temos duas equações de observação (independentes) com dados faltantes.

No Capítulo 8, desenvolvemos uma discussão da abordagem de modelagem utilizada, apresentando as vantagens e desvantagens dos modelos e procedimentos inferenciais propostos e algumas possíveis extensões para os modelos analisados. Finalmente, apresentamos as conclusões e recomendações do trabalho e as perspectivas para trabalhos futuros decorrentes da tese desenvolvida.

Capítulo 2

Modelos dinâmicos

2.1 Introdução

Os modelos dinâmicos são uma ferramenta muito útil para modelar uma grande variedade de problemas em que o comportamento de um processo ou sistema é estudado no domínio do tempo. Pode-se dizer que um modelo dinâmico é uma expressão matemática que descreve as regras que governam a mudança de um sistema (Gurney e Nisbet, 1998). A maioria dos modelos dinâmicos expressam estas regras de atualização especificando a relação entre o estado atual do sistema e o estado futuro. Alternativamente, eles podem ser expressos através de uma equação diferencial especificando-se a taxa de mudança das variáveis de estado. Quando estes modelos incorporam o uso de um modelo probabilístico, temos um modelo dinâmico estocástico (Hartig et al., 2011).

Para o ajuste de modelos dinâmicos estocásticos a uma série de dados, várias abordagens podem ser usadas. Uma abordagem muito utilizada é a de modelos espaço de estados (Commandeur e Koopman, 2007; Petris et al., 2009; Durbin e Koopman, 2012). Um modelo espaço de estados considera uma Cadeia de Markov não observável $(\mathbf{X}_t)_{t \geq 0}$ $\mathbb{R}^p$ -dimensional e $\mathbf{Y}_t$, $\mathbb{R}^m$ -dimensional é uma medição imprecisa de $\mathbf{X}_t$. Assim, esta representação assume que o passo no tempo do sistema sob estudo é determinado por uma série não observável de vetores aleatórios $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ que estão associados a uma série de vetores aleatórios observáveis $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$. Os modelos espaço de estados nos quais os estados são variáveis aleatórias discretas (ou seja o espaço dos estados é discreto), são conhecidos como *Hidden Markov models* (Petris et al., 2009). A seguir apresentamos uma definição formal de um modelo espaço de estados.

Definição 2.1. *Um modelo espaço de estados consiste em uma série temporal $\mathbb{R}^p$ -dimensional $(\mathbf{X}_t : t = 0, 1, \dots)$ e uma série temporal $\mathbb{R}^m$ -dimensional $(\mathbf{Y}_t : t = 1, 2, \dots)$ que satisfaçam as seguintes condições:*

1. $(\mathbf{X}_t)_{t \geq 0}$ é uma cadeia de Markov;
2. $\mathbf{Y}_t$ são condicionalmente independentes dado $(\mathbf{X}_t)_{t \geq 0}$;
3. $\mathbf{Y}_t$ depende de $(\mathbf{X}_t)_{t \geq 0}$ apenas através de $\mathbf{X}_t$.

$(\mathbf{X}_t)_{t \geq 0}$ é um processo estocástico com espaço de estados S e conjunto índice (ou parâmetro) em T , definidos sobre um espaço de probabilidade (Ω, σ, P) . O conjunto índice T é considerado contável, definindo um processo estocástico tempo discreto e o espaço de estado é um subconjunto de $\mathbb{R}^p$, que representa o suporte do vector de parâmetros (Gamerman e Lopes, 2006). Podemos considerar a função $\mathbf{X}_t : T \times \Omega \rightarrow \mathbb{R}$ de forma que para todo $t \in T$ a função $\mathbf{X}(t, \cdot) : \Omega \rightarrow \mathbb{R}$ é uma variável aleatória.

$(\mathbf{X}_t)_{t \geq 0}$ é uma Cadeia de Markov se $\mathbf{X}_t$ depende dos valores passados, só através de $\mathbf{X}_{t-1}$. A lei de probabilidade da cadeia é especificada assumindo uma densidade inicial $p(\mathbf{x}_0)$ e as densidades de transição $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ para $\mathbf{X}_t$, com $t > 0$. Logo, condicionalmente a $(\mathbf{X}_t)_{t \geq 0}$, os $\mathbf{Y}_t$ são

independentes, e os $\mathbf{Y}_t$ dependem só de $\mathbf{X}_t$. A estrutura de dependência de um modelo espaço de estados pode ser representada por

$$\begin{array}{ccccccccc} \mathbf{X}_0 & \rightarrow & \mathbf{X}_1 & \rightarrow & \mathbf{X}_2 & \rightarrow & \dots & \rightarrow & \mathbf{X}_{t-1} & \rightarrow & \mathbf{X}_t & \rightarrow & \mathbf{X}_{t+1} & \rightarrow & \dots \\ & & \downarrow & & \downarrow & & & & \downarrow & & \downarrow & & \downarrow & & \\ & & \mathbf{Y}_1 & & \mathbf{Y}_2 & & & & \mathbf{Y}_{t-1} & & \mathbf{Y}_t & & \mathbf{Y}_{t+1} & & \end{array}$$

Como consequência da definição acima tem-se que um modelo espaço de estados está completamente especificado através da distribuição inicial dos estados $p(\mathbf{x}_0)$ e as densidades condicionais $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ e $p(\mathbf{y}_t | \mathbf{x}_t)$, $t \geq 1$. Ou seja para qualquer $t > 0$:

$$p(\mathbf{x}_{0:t}, \mathbf{y}_{1:t}) = p(\mathbf{x}_0) \prod_{j=1}^t p(\mathbf{x}_j | \mathbf{x}_{j-1}) p(\mathbf{y}_j | \mathbf{x}_j) \quad (2.1)$$

onde $\mathbf{x}_{0:t} = (\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_t)$ e $\mathbf{y}_{1:t} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t)$.

2.2 Modelo espaço de estados linear gaussiano

Definição 2.2. Um modelo espaço de estados linear gaussiano (multivariado) é caracterizado por

- Uma equação de estado: $\mathbf{X}_t = \mathbf{G}_t \mathbf{X}_{t-1} + \mathbf{u}_t$, $\mathbf{u}_t \sim N_p(\mathbf{0}, \mathbf{W}_t)$, $t = 0, 1, \dots$,
- Uma equação de observação: $\mathbf{Y}_t = \mathbf{F}_t \mathbf{X}_t + \mathbf{v}_t$, $\mathbf{v}_t \sim N_m(\mathbf{0}, \mathbf{V}_t)$, $t = 1, \dots, n$,
- Uma informação inicial: $(\mathbf{X}_0 | \mathbf{D}_0) \sim N_p(\mathbf{m}_0, \mathbf{C}_0)$

onde $\mathbf{X}_t$ é um vetor p -dimensional de estados, $\mathbf{Y}_t$ é um vetor m -dimensional de variáveis observáveis, $\mathbf{G}_t$ e $\mathbf{F}_t$ são matrizes conhecidas de ordem $p \times p$ e $m \times p$ respectivamente, $(\mathbf{u}_t)_{t \geq 1}$ e $(\mathbf{v}_t)_{t \geq 1}$ são duas seqüências de vetores aleatórios gaussianos independentes de média zero e variâncias conhecidas. Além disso, é assumido que $\mathbf{X}_0$ é independente de $\mathbf{u}_t$ e $\mathbf{v}_t$.

Neste ponto é importante destacar que o processo inicia-se baseado em um conjunto de informações disponível antes de se observar os dados, ou seja, no tempo $t = 0$. Este conjunto (subjetivos ou não) é denotado por $\mathbf{D}_0$. Toda a informação sobre o futuro ($t > 0$) será condicional a $\mathbf{D}_0$. Ao chegar no tempo t , a informação disponível está concentrada em $\mathbf{D}_t = \{\mathbf{y}_t, \mathbf{D}_{t-1}\}$. Assim, quando o tempo evolui de t para $t + 1$, $\mathbf{y}_{t+1}$ é inserido no conjunto de informações, portanto $\mathbf{D}_{t+1} = \{\mathbf{y}_{t+1}, \mathbf{D}_t\}$.

Quando definimos uma distribuição *a priori* para $\mathbf{X}_0$, esta especificação aplica-se, em particular, quando $\mathbf{Y}_1, \mathbf{Y}_2, \dots$ representa a continuação de uma série observada anteriormente. Em tais casos, a distribuição *a priori* inicial é vista como um resumo da informação do passado com $\mathbf{X}_0$ tendo a interpretação concreta do estado *final* do vetor para os dados históricos. Caso contrário, $\mathbf{X}_0$ não possui tal interpretação e o modelo pode ser equivalentemente inicializado especificando uma distribuição *a priori* $(\mathbf{X}_1 | \mathbf{D}_0)$ para $t = 1$ (West e Harrison, 1997).

No caso particular do modelo espaço de estados linear gaussiano univariado $\mathbf{Y}_t = Y_t$ é de ordem 1×1 , $\mathbf{X}_t$ é um vetor $p \times 1$ de parâmetros de estado, $\mathbf{G}_t$ é uma matriz de ordem $p \times p$ e $\mathbf{F}_t$ é um vetor de ordem $p \times 1$.

Um dos modelos espaço de estados mais simples e difundido é o modelo de nível local (*local level model*), ou também conhecido como modelo polinômico de primeira ordem (West e Harrison, 1997; Durbin e Koopman, 2012), ou *random walk plus noise* (Petrus et al., 2009). A seguir apresentamos a definição formal deste modelo.

Definição 2.3. Um modelo espaço de estados linear gaussiano univariado de nível local é caracterizado por

- Uma equação de estado: $X_t = X_{t-1} + u_t$, $u_t \sim N(0, \tau^2)$, $t = 0, 1, \dots$,
- Uma equação de observação: $Y_t = X_t + v_t$, $v_t \sim N(0, \sigma^2)$, $t = 1, \dots, n$,
- Uma informação inicial: $(X_0 | D_0) \sim N(m_0, C_0)$

onde Y_t é uma variável aleatória observável, X_t é uma variável não observável e as seqüências v_t e u_t são internamente independentes, mutuamente independentes, e independentes do $(X_0 | D_0)$ e $t = 1, \dots, n$.

Note-se que as variâncias podem variar no tempo, mas no geral assume-se que são constantes. Assim as equações de observação e de evolução do modelo de nível local podem ser expressas da forma:

$$\begin{aligned} Y_t | X_t, \sigma^2 &\sim N(X_t, \sigma^2) \\ X_t | X_{t-1}, \tau^2 &\sim N(X_{t-1}, \tau^2). \end{aligned} \quad (2.2)$$

A distribuição *a posteriori* condicional conjunta, dos estados não observáveis e os parâmetros do modelo de nível local (no caso σ^2 e τ^2 desconhecidos) é:

$$p(x_{0:n}, \theta | y_{1:n}) \propto p(x_0 | \tau^2) \prod_{t=1}^n p(x_t | x_{t-1}, \tau^2) p(y_t | x_t, \sigma^2) p(\theta) \quad (2.3)$$

onde $p(y_t | x_t, \sigma^2)$ é a densidade das observações, $p(x_t | x_{t-1}, \tau^2)$ é a densidade de transição, $\theta = (\sigma^2, \tau^2)$ e $p(\theta)$ é a distribuição *a priori* dos parâmetros do modelo.

Este resultado segue devido ao fato que, condicionalmente a X_t , os Y_t são independentes. Além disso, os Y_t dependem só de X_t e, X_t dependem dos valores passados só através de X_{t-1} . Para obter a distribuição *a posteriori* marginal dos estados não observáveis integramos em relação aos parâmetros:

$$p(x_{1:n} | y_{1:n}) \propto \int p(x_{1:n} | y_{1:n}, \theta) p(\theta | y_{1:n}) d\theta. \quad (2.4)$$

A estimação da distribuição *a posteriori* marginal dos parâmetros, obtém-se integrando sobre os estados não observáveis através de:

$$p(\theta | y_{1:n}) \propto \left(\int \dots \int p(y_{1:n} | x_{1:n}, \theta) p(x_{1:n} | \theta) dx_1 \dots dx_n \right) p(\theta). \quad (2.5)$$

2.2.1 Inferência sequencial

A inferência sequencial envolve três operações básicas (por simplicidade de notação, omitimos a dependência das densidades sob os parâmetros):

- Densidade preditiva para os estados um passo a frente (*a priori* para X_{t+1}):

$$p(x_{t+1} | y_{1:t}) = \int p(x_{t+1} | x_t) p(x_t | y_{1:t}) dx_t \quad (2.6)$$

- Densidade preditiva para as observações um passo a frente:

$$p(y_{t+1} | y_{1:t}) = \int p(y_{t+1} | x_{t+1}) p(x_{t+1} | y_{1:t}) dx_{t+1} \quad (2.7)$$

- Densidade de filtragem (atualização ou *a posteriori* para X_{t+1}):

$$p(x_{t+1} | y_{1:t+1}) = \frac{p(y_{t+1} | x_{t+1}) p(x_{t+1} | y_{1:t})}{p(y_{t+1} | y_{1:t})}. \quad (2.8)$$

Uma das características atraentes dos modelos espaço de estados é que a estimação e previsão podem ser feitas sequencialmente, a medida que novos dados estejam disponíveis. No entanto, na análise de séries temporais, muitas vezes tem-se observações $Y_t = y_t$ para um certo período, $t = 1, \dots, n$ e deseja-se reconstruir retrospectivamente o comportamento do sistema subjacente as observações. Neste caso, pode-se utilizar um algoritmo *backward* recursivo para calcular as distribuições condicionais de X_t dado $Y_1, \dots, Y_n$, para qualquer $t < n$, a partir da distribuição de filtragem $p(x_n | y_{1:n})$ e estimando para atrás toda a história dos estados. O resultado para os modelos de espaço de estados gerais está contido na seguinte proposição.

Proposição 2.1. *Para um modelo espaço de estados geral tem-se:*

- (i) *Condicionalmente a $Y_{1:n}$, a sequência de estados $(X_0, \dots, X_n)$ tem probabilidades de transição backward dadas por:*

$$p(x_t | x_{t+1}, y_{1:n}) = \frac{p(x_{t+1} | x_t) p(x_t | y_{1:t})}{p(x_{t+1} | y_{1:t})}. \quad (2.9)$$

- (ii) *A distribuição suavizada de X_t dados $Y_{1:n}$ pode ser calculada segundo a seguinte equação recursiva, partindo de $p(x_n | y_{1:n})$:*

$$p(x_t | y_{1:n}) = p(x_t | y_{1:t}) \int \frac{p(x_{t+1} | x_t)}{p(x_{t+1} | y_{1:t})} p(x_{t+1} | y_{1:n}) dx_{t+1} \quad (2.10)$$

A prova desta proposição se encontra no Apêndice A.1. A seguir apresentamos formalmente o Filtro de Kalman (KF).

Proposição 2.2. *Considere um modelo espaço de estados linear gaussiano como o definido anteriormente. Seja $X_{t-1} | Y_{1:t-1} \sim N(m_{t-1}, C_{t-1})$. Então as seguintes afirmações valem:*

- (i) *A distribuição preditiva um passo a frente de X_t dado $Y_{1:t-1}$ é gaussiana com parâmetros*

$$\begin{aligned} a_t &= E(X_t | Y_{1:t-1}) = G_t m_{t-1}, \\ R_t &= \text{Var}(X_t | Y_{1:t-1}) = G_t C_{t-1} G_t' + W_t. \end{aligned}$$

- (ii) *A distribuição preditiva um passo a frente de Y_t dado $Y_{1:t-1}$ é gaussiana com parâmetros*

$$\begin{aligned} f_t &= E(Y_t | Y_{1:t-1}) = F_t a_t, \\ Q_t &= \text{Var}(Y_t | Y_{1:t-1}) = F_t R_t F_t' + V_t. \end{aligned}$$

- (iii) *A distribuição de filtragem de X_t dado $Y_{1:t}$ é gaussiana com parâmetros*

$$\begin{aligned} m_t &= E(X_t | Y_{1:t}) = a_t + R_t F_t' Q_t^{-1} e_t, \\ C_t &= \text{Var}(X_t | Y_{1:t}) = R_t - R_t F_t' Q_t^{-1} F_t R_t. \end{aligned}$$

onde $e_t = Y_t - f_t$.

A prova desta proposição se encontra no Apêndice A.1.

Figura 2.1: Representação do processo de atualização da informação no Filtro de Kalman.

2.2.2 Dados faltantes e previsão

Não é raro em aplicações ter que se lidar com uma série de tempo contendo um ou mais dados faltantes. A estrutura dos modelos espaço de estados é tal que as observações faltantes podem ser facilmente acomodadas. Se a observação no tempo t está faltando, então Y_t não contém informação, de modo que:

$$p(x_{t+1} | y_{1:t+1}) = \frac{p(y_{t+1} | x_{t+1}) p(x_{t+1} | y_{1:t})}{p(y_{t+1} | y_{1:t})} = p(x_{t+1} | y_{1:t}). \quad (2.11)$$

Esquemáticamente:

$$\begin{array}{ccccccccccc} X_0 & \rightarrow & X_1 & \rightarrow & X_2 & \rightarrow & \dots & \rightarrow & X_{t-1} & \rightarrow & X_t & \rightarrow & X_{t+1} & \rightarrow & \dots \\ & & \downarrow & & & & & & \downarrow & & & & \downarrow & & \\ & & Y_1 & & & & & & Y_{t-1} & & & & Y_{t+1} & & \end{array}$$

Isto significa que, neste caso, a distribuição de filtragem no tempo t (ou densidade *a posteriori* dos estados) é equivalente a distribuição preditiva dos estados no tempo $t-1$. Em particular, para o caso do modelo espaço estados linear gaussiano, dado que $X_t | Y_{1:t-1} \sim N(a_t, R_t)$, tudo que se precisa fazer é definir $m_t = a_t$ e $C_t = R_t$. Note que, num modelo espaço estados linear gaussiano, se Y_t é faltante, é o mesmo que definir $F_t = 0$ ou $V_t = \infty$. No primeiro caso, não existe ligação entre Y_t e X_t , e na segunda, a observação é tão imprecisa que não contém informação sobre X_t (Petrís et al., 2009).

Para o caso de ser preciso fazer inferência a respeito das observações faltantes, precisa-se somente utilizar a equação de observação (que relaciona os estados não observáveis com as observações) e obter a densidade preditiva das observações. Desta forma conseguimos completar os dados faltantes. Esquemáticamente:

$$\begin{array}{ccccccccccc} X_0 & \rightarrow & X_1 & \rightarrow & X_2 & \rightarrow & \dots & \rightarrow & X_{t-1} & \rightarrow & X_t & \rightarrow & X_{t+1} & \rightarrow & \dots \\ & & \downarrow & & \downarrow & & & & \downarrow & & \downarrow & & \downarrow & & \\ & & Y_1 & & \hat{Y}_2 & & & & Y_{t-1} & & \hat{Y}_t & & Y_{t+1} & & \end{array}$$

$\hat{Y}_t$ é a média da densidade preditiva dos dados faltantes. A seguinte proposição apresenta formulas recursivas para obter a distribuição de previsão para os estados e as observações para um modelo espaço de estados geral.

Proposição 2.3. *Para um modelo espaço de estados geral tem-se para $k > 0$:*

- *A distribuição de previsão do estado k passos no frente é:*

$$p(x_{t+k} | y_{1:t}) = \int p(x_{t+k} | x_{t+k-1}) p(x_{t+k-1} | y_{1:t}) dx_{t+k-1}. \quad (2.12)$$

- *A distribuição de previsão da observação k passos no frente é:*

$$p(y_{t+k} | y_{1:t}) = \int p(y_{t+k} | x_{t+k}) p(x_{t+k} | y_{1:t}) dx_{t+k}. \quad (2.13)$$

A prova desta proposição é dada no Apêndice A.1. Assim para fazer previsão dos estados não observáveis k passos a frente ($k > 0$), usamos o mesmo procedimento descrito para o caso de dados faltantes. Considerando que o vetor de observações não contém informação sobre $X_{t+1} \dots X_{t+k}$, a densidade *a posteriori* dos estados é equivalente a densidade preditiva dos estados. Da mesma forma, para o caso de precisar fazer inferência das observações k passos a frente (ou seja previsão), utilizamos a densidade preditiva das observações. Esquematicamente:

$$\begin{array}{ccccccc} X_t & \rightarrow & X_{t+1} & \rightarrow & \dots & \rightarrow & X_{t+k} \\ \downarrow & & & & & & \downarrow \\ Y_{1:t} & & & & & & \hat{Y}_{t+k} \end{array}$$

2.3 Modelo espaço de estados não linear gaussiano

Uma classe de modelos mais flexíveis são os modelos dinâmicos não lineares gaussianos (Carlin et al., 1992; Gordon et al., 1993; Kitagawa, 1996; Geweke e Tanizaki, 1999; Shumway e Stoffer, 2011).

Definição 2.4. *Um modelo espaço de estados não linear gaussiano é caracterizado por*

- *Uma equação de estado: $\mathbf{X}_{t+1} = \mathbf{f}_t(\mathbf{X}_t, \mathbf{u}_t)$, $\mathbf{u}_t \sim N_p(\mathbf{0}, \mathbf{W}_t)$, $t = 0, 1, \dots$,*
- *Uma equação de observação: $\mathbf{Y}_t = \mathbf{h}_t(\mathbf{X}_t, \mathbf{v}_t)$, $\mathbf{v}_t \sim N_m(\mathbf{0}, \mathbf{V}_t)$, $t = 1, \dots, n$,*
- *Uma informação inicial: $(\mathbf{X}_0 | \mathbf{D}_0) \sim N_p(\mathbf{m}_0, \mathbf{C}_0)$*

$\mathbf{X}_t$ é um vetor de estado p -dimensional, $\mathbf{u}_t$ é o erro de processo l -dimensional, $\mathbf{f}_t : \mathbb{R}^p \times \mathbb{R}^l \rightarrow \mathbb{R}^p$ é a função de transição, $\mathbf{Y}_t$ é um vetor de observações m -dimensional, $\mathbf{v}_t$ é o erro de observação r -dimensional e $\mathbf{h}_t : \mathbb{R}^p \times \mathbb{R}^r \rightarrow \mathbb{R}^m$ é a função de observação. $\mathbf{v}_t$'s e $\mathbf{u}_t$'s são seqüências de vetores aleatórios gaussianos independentes de média zero e variâncias $\mathbf{V}_t$ e $\mathbf{W}_t$, respectivamente. As funções vetoriais $\mathbf{f}_t$ e $\mathbf{h}_t$ são conhecidas mas podem depender de parâmetros desconhecidos. Além disso, é assumido que $\mathbf{X}_0$ é independente de $\mathbf{u}_t$ e $\mathbf{v}_t$.

A inferência nos modelos não-lineares gaussianos podem ser feitas usando várias abordagens. Antes da década de 80, métodos de aproximação foram desenvolvidos para contornar a complexidade da análise de tais modelos. Por exemplo, técnicas de linearização e a soma de filtros gaussianos (Alspach e Sorenson, 1972; Anderson e Moore (1979). Várias técnicas de linearização tem sido desenvolvidas. A forma mais direta é aproximação de primeira ordem via Series de Taylor. No entanto isto requer que as funções $\mathbf{f}_t$ e $\mathbf{h}_t$ sejam ao menos uma vez diferenciáveis em seus argumentos vetoriais. Este procedimento de aproximação é conhecido como Filtro de Kalman Estendido (EKF), e será apresentado na próxima subseção.

2.3.1 Estimação de parâmetros através do Filtro de Kalman Estendido

Nesta seção apresentamos o Filtro de Kalman Estendido (EKF), o qual corresponde a um procedimento de estimação dos estados não observáveis e parâmetros num modelo espaço estados não linear através de aproximação via expansão em séries de Taylor, tanto na equação de observação quanto da equação de estado. Apresentamos separadamente os casos com variância de observação conhecida e desconhecida.

a) Variância de observação conhecida no modelo não linear gaussiano

Por simplicidade vamos apresentar aqui o caso univariado (Y_t). Seguindo Harvey (1990), o EKF pode ser utilizado na presença de um modelo espaço de estado não linear (NLSSM) como o apresentado na definição 2.4., da forma:

$$\begin{aligned} \mathbf{X}_t &= \mathbf{f}_t(\mathbf{X}_{t-1}) + \mathbf{u}_t, \quad \mathbf{u}_t \sim N(\mathbf{0}, \mathbf{W}_t), \quad t = 0, 1, \dots, \\ Y_t &= h_t(\mathbf{X}_t) + v_t, \quad v_t \sim N(0, V_t), \quad t = 1, \dots, n, \end{aligned} \quad (2.14)$$

$$\mathbf{X}_1 | D_0 \sim N(\mathbf{m}_0, \mathbf{C}_0).$$

$\mathbf{X}_t$ é um vetor de estado p -dimensional, $\mathbf{f}_t$ é uma função não-linear de evolução, $\mathbf{f}_t : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}^p$, $\mathbf{u}_t$ é um ruído branco gaussiano, com média zero e matriz de variância-covariância $\mathbf{W}_t$, Y_t é um vetor de observação unidimensional, h_t é uma função não-linear que mapeia o vetor $\mathbf{X}_t$ para a reta real e v_t é o erro de observação com média zero e variância V_t . Além disso, $(\mathbf{m}_0, \mathbf{C}_0)$ são assumidos conhecidos, e D_0 representa o conjunto de informação inicial disponível.

Diversas técnicas de linearização têm sido desenvolvidas para esta classe de modelos, quase todas baseadas no uso de aproximação linear local para as não-linearidades. A abordagem mais simples, baseia-se no uso de aproximação por séries de Taylor de primeira ordem para as equações não lineares. Isso requer que as funções não lineares de observação e evolução sejam pelo menos uma vez diferenciáveis em seus argumentos (West e Harrison, 1997). Vamos começar com a equação de evolução, fazendo a expansão em série de Taylor em torno da média condicional de $\mathbf{X}_{t-1} | D_{t-1}$ ($\mathbf{m}_{t-1}$),

$$\mathbf{f}_t(\mathbf{X}_{t-1}) \approx \mathbf{f}_t(\mathbf{m}_{t-1}) + \mathbf{F}_t[\mathbf{X}_{t-1} - \mathbf{m}_{t-1}], \quad (2.15)$$

onde $\mathbf{F}_t$ é a matriz de derivadas parciais da função de evolução avaliada em $\mathbf{m}_{t-1}$ da forma:

$$\mathbf{F}_t = \left[\frac{\partial \mathbf{f}_t(\mathbf{X}_{t-1})}{\partial \mathbf{X}_{t-1}'} \right]_{\mathbf{X}_{t-1} = \mathbf{m}_{t-1}}. \quad (2.16)$$

Assumindo que os outros termos são insignificantes, a equação de evolução resulta

$$\mathbf{X}_t \approx \mathbf{f}_t(\mathbf{m}_{t-1}) + \mathbf{F}_t[\mathbf{X}_{t-1} - \mathbf{m}_{t-1}] + \mathbf{u}_t = \mathbf{b}_t + \mathbf{F}_t \mathbf{X}_{t-1} + \mathbf{u}_t, \quad (2.17)$$

onde $\mathbf{b}_t = \mathbf{f}_t(\mathbf{m}_{t-1}) - \mathbf{F}_t \mathbf{m}_{t-1}$ é conhecida. A expressão acima é conhecida como a versão linearizada da equação de evolução (linearizado sob o valor esperado de $\mathbf{X}_{t-1}$).

Considerando a suposição de erros gaussianos, resulta $\mathbf{X}_t | D_{t-1} \sim N(\mathbf{a}_t, \mathbf{R}_t)$ onde:

$$\mathbf{a}_t = \mathbf{f}_t(\mathbf{m}_{t-1}), \quad \mathbf{R}_t = \mathbf{F}_t \mathbf{C}_{t-1} \mathbf{F}_t' + \mathbf{W}_t. \quad (2.18)$$

onde $\mathbf{C}_{t-1}$ é a variância condicional de $\mathbf{X}_{t-1} | D_{t-1}$.

Da mesma forma procedemos com a equação de observação. Agora, a função de observação não-linear é linearizada em torno do valor esperado de $\mathbf{X}_{t-1}|D_t(\mathbf{a}_t)$, resultando em:

$$h_t(\mathbf{X}_t) \approx h_t(\mathbf{a}_t) + \mathbf{H}_t'[\mathbf{X}_t - \mathbf{a}_t], \quad (2.19)$$

com

$$\mathbf{H}_t = \left[\frac{\partial h_t(\mathbf{X}_t)}{\partial \mathbf{X}_t} \right]_{\mathbf{X}_t=\mathbf{a}_t}. \quad (2.20)$$

$\mathbf{H}_t$ é o vetor de derivadas de h_t avaliado em $\mathbf{a}_t$. Assumindo que o termo linear domina a expansão, a equação de observação linearizada torna-se:

$$Y_t \approx h_t(\mathbf{a}_t) + \mathbf{H}_t'[\mathbf{X}_t - \mathbf{a}_t] + v_t = g_t + \mathbf{H}_t'\mathbf{X}_t + v_t, \quad (2.21)$$

onde $g_t = h_t(\mathbf{a}_t) - \mathbf{H}_t'\mathbf{a}_t$ é conhecido. A expressão acima é conhecida como a versão linearizada da equação de observação (linearizada em torno de $\mathbf{a}_t$). A partir dessas equações segue que a previsão um passo à frente para as observações é $Y_t|D_{t-1} \sim N(h_t(\mathbf{a}_t), Q_t)$ onde $Q_t = \mathbf{H}_t'\mathbf{R}_t\mathbf{H}_t + V_t$.

Considerando a suposição de erros gaussianos, a densidade atualizada (ou distribuição *a posteriori* dos estados) é $\mathbf{X}_t|D_t \sim N(\mathbf{m}_t, \mathbf{C}_t)$ com

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t, \quad \mathbf{C}_t = \mathbf{R}_t - \mathbf{R}_t\mathbf{H}_t'Q_t^{-1}\mathbf{H}_t\mathbf{R}_t, \quad (2.22)$$

e

$$\mathbf{A}_t = \mathbf{R}_t\mathbf{H}_tQ_t^{-1}, \quad e_t = Y_t - h_t(\mathbf{a}_t). \quad (2.23)$$

Este último resultado segue da aplicação da equação de densidade de filtragem apresentada anteriormente.

Figura 2.2: Representação do processo de atualização da informação no Filtro de Kalman estendido.

b) Variância de observação desconhecida no modelo não linear gaussiano

Normalmente a sequência de variância de observação V_t não será conhecida. No entanto, o analista geralmente tem algum conhecimento prévio sobre as características da sequência. Uma opção é considerar $V_t = V$, com V desconhecida. Neste caso, é possível realizar uma análise Bayesiana conjugada com todas as variâncias e covariâncias escaladas por V . O modelo é dado por

$$\mathbf{X}_t = \mathbf{f}_t(\mathbf{X}_{t-1}) + \mathbf{u}_t, \quad \mathbf{u}_t \sim N(\mathbf{0}, V\mathbf{W}_t^*), \quad t = 0, 1, \dots, \quad (2.24)$$

$$Y_t = h_t(\mathbf{X}_t) + v_t, \quad v_t \sim N(0, V), \quad t = 1, \dots, n, \quad (2.25)$$

$$\mathbf{X}_1 | D_0, V \sim N(\mathbf{m}_0, V \mathbf{C}_0^*), \quad (2.26)$$

$$\phi | D_0 \sim G(n_0/2, d_0/2), \quad (2.27)$$

para algum $\mathbf{m}_0, \mathbf{C}_0^*, \mathbf{W}_t^*, n_0$ e d_0 conhecidos. Note que, de 2.27 resulta que a média de $\phi = V^{-1}$, condicionada a informação prévia é $E(\phi | D_0) = n_0/d_0 = 1/S_0$, onde S_0 é geralmente uma estimativa pontual da variância observacional. Considerando as propriedades da distribuição Normal-Gamma (ver Apêndice A1), incondicionalmente em V , temos os seguintes resultados:

$$\mathbf{X}_t | D_{t-1} \sim t(\mathbf{a}_t, \mathbf{R}_t, n_{t-1}), \quad (2.28)$$

$$\phi | D_{t-1} \sim G(n_{t-1}/2, d_{t-1}/2), \quad (2.29)$$

onde

$$\mathbf{a}_t = \mathbf{f}_t(\mathbf{m}_{t-1}), \quad \mathbf{R}_t = \mathbf{F}_t \mathbf{C}_{t-1} \mathbf{F}_t' + \mathbf{W}_t. \quad (2.30)$$

e $\mathbf{F}_t$ é a matriz de derivadas da equação de evolução avaliada em $\mathbf{m}_{t-1}$ dada por

$$\mathbf{F}_t = \left[\frac{\partial \mathbf{f}_t(\mathbf{X}_{t-1})}{\partial \mathbf{X}_{t-1}'} \right]_{\mathbf{X}_{t-1} = \mathbf{m}_{t-1}}. \quad (2.31)$$

A previsão um passo à frente para as observações é

$$Y_t | D_{t-1} \sim t(h_t(\mathbf{a}_t), Q_t, n_{t-1}), \quad (2.32)$$

onde $h_t(\mathbf{a}_t)$ é o valor esperado de $\mathbf{X}_{t-1} | D_t$ avaliado na função h_t e $Q_t = \mathbf{H}_t' \mathbf{R}_t \mathbf{H}_t + S_{t-1}$, com

$$\mathbf{H}_t = \left[\frac{\partial h_t(\mathbf{X}_t)}{\partial \mathbf{X}_t} \right]_{\mathbf{X}_t = \mathbf{a}_t}. \quad (2.33)$$

As distribuições atualizadas (distribuições *a posteriori*) para os estados e a precisão são, respectivamente,

$$\mathbf{X}_t | D_t \sim t(\mathbf{m}_t, \mathbf{C}_t, n_t), \quad (2.34)$$

$$\phi | D_t \sim G(n_t/2, d_t/2), \quad (2.35)$$

onde $\mathbf{m}_t$ foi definido em (2.22), $n_t = n_{t-1} + 1$, $d_t = n_t S_t$ e

$$\mathbf{C}_t = \frac{S_t}{S_{t-1}} (\mathbf{R}_t - \mathbf{A}_t \mathbf{A}_t' Q_t), \quad S_t = S_{t-1} \frac{S_{t-1}}{n_t} \left(\frac{e^2}{Q_t} - 1 \right). \quad (2.36)$$

2.4 Modelo espaço de estados não linear e não gaussiano

Uma classe de modelos mais flexíveis são os modelos dinâmicos não lineares e não gaussianos (Carlin et al., 1992; Gordon et al., 1993; Kitagawa, 1996; Geweke e Tanizaki, 1999; Shumway e Stoffer, 2011).

Definição 2.5. *Um modelo espaço de estados não linear e não gaussiano é caracterizado por:*

- Uma equação de estado: $\mathbf{X}_{t+1} = \mathbf{f}_t(\mathbf{X}_t, \mathbf{u}_t)$, $\mathbf{u}_t \sim g_l(\mathbf{u}_t)$, $t = 0, 1, \dots$,
- uma equação de observação: $\mathbf{Y}_t = \mathbf{h}_t(\mathbf{X}_t, \mathbf{v}_t)$, $\mathbf{v}_t \sim q_r(\mathbf{v}_t)$, $t = 1, \dots, n$,
- uma informação inicial: $(\mathbf{X}_0 | \mathbf{D}_0) \sim g_l(\mathbf{u}_0)$

onde $\mathbf{X}_t$ é um vetor de estado p -dimensional, $\mathbf{u}_t$ é o erro de processo l -dimensional, $\mathbf{f}_t : \mathbb{R}^p \times \mathbb{R}^l \rightarrow \mathbb{R}^p$ é a função de transição, $\mathbf{Y}_t$ é um vetor de observações m -dimensional, $\mathbf{v}_t$ é o erro de observação r -dimensional e $\mathbf{h}_t : \mathbb{R}^p \times \mathbb{R}^r \rightarrow \mathbb{R}^m$ é a função de observação. $\mathbf{u}_t$'s e $\mathbf{v}_t$'s são seqüências de ruído branco com média zero e com funções de densidade de probabilidade conhecidas $g_l(\mathbf{u}_t)$ e $q_r(\mathbf{v}_t)$, respectivamente. As funções vetoriais $\mathbf{f}_t$ e $\mathbf{h}_t$ são conhecidas mas podem depender de parâmetros desconhecidos.

A inferência nos modelos não-lineares e não-gaussianos podem ser feitas usando várias abordagens. Kitagawa (1987) propôs um método numérico baseado em aproximações lineares por partes (*piecewise linear approximation*) para as funções de densidade de previsão, filtragem e suavização para os modelos de espaço de estado não-gaussianos e não estacionários. Pole e West (1988) utilizaram técnicas de quadratura gaussiana em uma análise Bayesiana de modelos dinâmicos não-lineares. West e Harrison (1997) fornecem uma explicação detalhada destes métodos. Além destas propostas, existem outras abordagens de inferência em modelos não-lineares e não-gaussianos, as quais podem ser classificadas em quatro grandes grupos: (1) integração de Monte Carlo e os métodos que utilizam o amostragem de importância (*Importance sampling*), (2) os métodos de Monte Carlo que baseados em cadeias de Markov (*Markov chain Monte Carlo*, MCMC), (3) os algoritmos sequenciais de Monte Carlo, também conhecidos como filtros de partículas e (4) os métodos que utilizam alguma variante da aproximação de Laplace.

No primeiro caso, são métodos de simulação baseados na amostragem de importância (Ripley, 1987) os quais podem-se utilizar nas abordagens clássica ou Bayesiana. Durbin e Koopman (2012) fornecem uma apresentação detalhada destes métodos. Na década de noventa houve um grande desenvolvimento do segundo grupo de técnicas (métodos MCMC). O método MCMC mais comum é o amostrador de Gibbs (*Gibbs sampler* ou *Gibbs sampling*). Trabalhos importantes nessa direção são Carlin et al. (1992), Carter e Kohn (1994), Carter e Kohn (1996), Shepard e Pitt (1997), Gamerman (1998), Geweke e Tanizaki (1999). O terceiro grupo, com muito desenvolvimento nos últimos anos, tem o atrativo de permitir realizar inferência *on-line*, ou seja, quando estiver disponível uma nova observação, não é necessário simular uma nova cadeia de Markov para obter as distribuições *a posteriori* das quantidades de interesse. Algumas referências relevantes desta abordagem são Gordon et al. (1993), Kitagawa (1996), Pitt e Shephard (1999), Doucet et al. (2000), Migon et al. (2005) e Prado e West (2010). No quarto grupo, existem várias aproximações, como o algoritmo INLA (Integrated Nested Laplace Approximation, Martino e Rue, 2008) e os algoritmos de diferenciação automática em conjunto com aproximação de Laplace para a verosimilhança marginal (Skaug e Fournier, 2006, Fournier et al., 2012).

Capítulo 3

Modelo dinâmico logístico

3.1 Introdução

Para desenvolver uma modelagem estatística em populações marinhas, deve-se propor modelos de dinâmica populacional. A escolha deste modelo vai depender de múltiplos fatores. No caso de se ter pouca informação sobre o recurso marinho em estudo, ou seja, ter um baixo nível de desenvolvimento na coleta dos dados, é comum utilizar modelos indiretos. Estes modelos tem a vantagem de ter baixo custo monetário. Um dos modelos desta classe mais utilizados, para estudo de populações marinhas exploradas pelo homem, é o modelo dinâmico logístico, também conhecido como modelo de produção ou modelo de excedentes de produção (Hilborn e Walters, 1992). Este modelo tem uma longa história na pesquisa marinha, pois permitiu desenvolver um conceito chave na tomada de decisões pesqueiras, que é o conceito de máximo rendimento sustentável, ou máximo nível de exploração sustentável (MSY, do inglês *Maximum Sustainable Yield*). Existem várias versões deste modelo, entre as mais conhecidas estão as de Schaefer (1954), Pella e Tomlinson (1969), Fletcher (Quinn e Deriso, 1999) e a versão de Gilbert (1992). No presente trabalho de tese, trabalharemos com a versão de Schaefer do modelo logístico pois o foco principal é o estudo dos termos estocásticos da equações de observação e do processo, mais do que a forma explícita da equação de evolução.

A seguir vamos apresentar o modelo de crescimento logístico de populações que, embora baseado no modelo de crescimento exponencial de populações, permite incorporar o efeito denso-dependente nas taxas de crescimento populacional (Haddon, 2011).

3.2 Modelo de crescimento exponencial

O crescimento populacional em sua forma abstrata mais simples, pode ser pensado como uma população que cresce em um ambiente ilimitado, seja de espaço ou alimento (Haddon, 2011). O modelo que descreve este comportamento é conhecido como modelo de crescimento populacional exponencial, no qual as taxas de nascimento e morte são constantes para todos os tamanhos da população. Este modelo é formalmente dado por

$$\frac{dX_t}{dt} = (b - d)X_t \equiv rX_t, \quad (3.1)$$

em que X_t é a abundância de indivíduos, t é o tempo, d é a taxa de morte, b é a taxa de nascimentos e r é a taxa intrínseca de crescimento da população. Integrando a equação diferencial temos

$$\begin{aligned} \frac{dX_t}{dt} &= rX_t \\ \Rightarrow \int_{X_0}^{X_t} \frac{dX_t}{X_t} &= r \int_0^t dt \end{aligned}$$

$$\Rightarrow \ln \frac{X_t}{X_0} = rt$$

$$X_t = X_0 e^{rt}. \quad (3.2)$$

Na Figura 3.1 apresentamos a trajetória teórica de três populações que crescem segundo o modelo de crescimento exponencial, com diferentes valores de r .

Figura 3.1: Trajetórias de três populações com crescimento exponencial.

Vamos generalizar o modelo de crescimento exponencial para o caso em que as taxas de morte d e de nascimento b são funções lineares da abundância. Assim

$$b_X = b_{max} - b_1 X \quad (3.3)$$

$$d_X = d_{min} + d_1 X \quad (3.4)$$

onde b_{max} é a taxa de nascimento quando $X = 0$ e d_{min} é a taxa de morte quando $X = 0$ (intercepto dos modelos lineares). Note que, a taxa de crescimento da população decresce conforme aumenta a abundância (equação 3.3) e que, a taxa de decrescimento da população aumenta, conforme aumenta a abundância (3.4). Isto é chamado efeito denso-dependente (Haddon, 2011). A Figura 3.2 apresenta um esquema gráfico destes dois processos.

Se substituirmos no modelo exponencial as taxas denso-dependentes obtemos:

$$\begin{aligned} \frac{dX_t}{dt} &= [(b_{max} - b_1 X_t) - (d_{min} + d_1 X_t)] X_t \\ &= (b_{max} - d_{min}) X_t - (b_1 + d_1) X_t^2. \end{aligned} \quad (3.5)$$

Reparametrizando o modelo, fazendo $r = b_{max} - d_{min}$ temos

$$\begin{aligned} \frac{dX_t}{dt} &= r X_t - (b_1 + d_1) X_t^2 \\ &= r X_t \left(1 - \frac{(b_1 + d_1) X_t}{r} \right). \end{aligned} \quad (3.6)$$

Figura 3.2: Efeitos denso-dependentes de nascimento e morte de uma população.

Fazendo a substituição $K = r/(b_1 + d_1)$ resulta

$$\frac{dX_t}{dt} = rX_t \left(1 - \frac{X_t}{K}\right). \quad (3.7)$$

Assim, obtemos o modelo de crescimento logístico de populações biológicas (Krebs, 1985).

3.3 Modelo de crescimento logístico

No modelo de crescimento logístico de populações, a abundância está limitada por um valor K e os níveis de taxa de crescimento populacional são uma função quadrática da abundância (Haddon, 2001). O modelo de crescimento logístico de populações biológicas pode ser expresso na forma:

$$\frac{dX_t}{dt} = rX_t \left(1 - \frac{X_t}{K}\right) \equiv rX_t - \frac{rX_t^2}{K}, \quad X_t \geq 0, \quad r > 0, \quad K > 0, \quad t = 0, 1, \dots, \quad (3.8)$$

onde X_t é a biomassa (abundância ou peso) no tempo t da população, r é a taxa intrínseca de crescimento da população e K é a capacidade de carga, ou seja, a abundância máxima que o ambiente pode manter (Haddon, 2001). Este modelo é denso-dependente, isto é, a taxa de mudança da abundância da população depende do seu tamanho. A expressão acima é uma equação diferencial quadrática, separável, autônoma e homogênea (Quinn and Deriso, 1999). A Figura 3.3 apresenta o gráfico da equação 3.8. A solução ao problema de valor inicial $X(t_0) = X_0$ pode ser obtida através de frações parciais (Braun, 1983), resultando na expressão

$$X(t) = \frac{K}{1 + \frac{K-X_0}{X_0}e^{-rt}}. \quad (3.9)$$

A Figura 3.4 apresenta o gráfico da equação (3.9). O modelo de crescimento logístico tem dois pontos de equilíbrio: $X = 0$ e $X = K$. O máximo obtém-se derivando a função $g(X_t) = \frac{dX_t}{dt}$ em relação a X_t e igualando a zero. Assim, o máximo da função é $K/2$.

Figura 3.3: Taxa de mudança da população, no modelo de crescimento logístico, em função do tamanho da população

Figura 3.4: Trajetória da abundância de uma população com crescimento logístico

A versão do modelo em equação de diferenças finitas de ordem um é dada por

$$X_{t+1} = X_t + rX_t \left(1 - \frac{X_t}{K}\right). \quad (3.10)$$

Se a população esta submetida à pesca, então podemos acrescentar outro termo à expressão anterior, tal que seja descontada a biomassa extraída pelo homem. Assim

$$X_{t+1} = X_t + rX_t \left(1 - \frac{X_t}{K}\right) - H_t, \quad (3.11)$$

onde H_t é a captura (ou *harvesting*) em peso (ou abundância) no tempo t , a qual é considerada conhecida.

Uma das características interessantes deste modelo é que pode-se calcular facilmente o máximo nível de exploração sustentável (MSY). Este máximo é atingido quando $\frac{dg(X_t)}{dX_t} = 0 \Rightarrow X_t = K/2$. Assim, de (3.11) resulta

$$MSY = r \frac{K}{2} \left(1 - \frac{K}{2} \frac{1}{K}\right) = \frac{rK}{4}. \quad (3.12)$$

Este indicador tem sido amplamente utilizado para a gestão da pesca em diferentes partes do mundo (Quinn and Deriso, 1999). Além disso, este modelo permite determinar a taxa de captura que maximiza os excedentes de produção (H_{MSY}), dada por

$$H_{MSY} = \frac{r}{2}. \quad (3.13)$$

3.4 Representação via espaço de estados do modelo de crescimento logístico

A versão estocástica deste modelo é dada pela equação de estado na forma:

$$X_{t+1} = X_t + rX_t \left(1 - \frac{X_t}{K}\right) - H_t + u_t, \quad X_t \geq 0, \quad r > 0, \quad K > 0, \quad H_t \geq 0, \quad t = 0, 1, \dots, \quad (3.14)$$

Neste trabalho, utilizamos para a equação de estado a hipóteses erros de processo aditivos, normalmente distribuídos e com variância constante, ou seja, $u_t \sim N(0, \tau^2)$. Adicionalmente, assume-se que a condição inicial de sistema é $X_1|D_0 \sim N(\mu_0, \tau_0^2)$, onde D_0 representa o conjunto de informação inicial disponível. Considerando a existência de uma variável observável, que pode se relacionar com os estados não observáveis, temos a equação de observação da forma:

$$Y_t = qX_t + v_t, \quad Y_t \geq 0, \quad q > 0, \quad t = 1, \dots, n. \quad (3.15)$$

Y_t é a variável observável (por exemplo, o rendimento da pesca), q é o coeficiente de proporcionalidade, denominado coeficiente de capturabilidade, X_t é a biomassa da espécie e v_t são variáveis aleatórias independentes e identicamente distribuídas. Note que, nas equações 3.14 e 3.15 os erros são aditivos. Na presente tese consideramos também erros de observação multiplicativos, ou seja:

$$Y_t = qX_te^{v_t}, \quad (3.16)$$

os que utilizando transformação log, ficam do tipo aditivo apresentado acima, da forma,

$$\log(Y_t) = \log(q) + \log(X_t) + v_t, \quad t = 1, \dots, n. \quad (3.17)$$

No Capítulo 5 detalhamos ambas especificações do modelo. Se consideramos um conjunto de variáveis aleatórias observáveis independentes, condicionais a X_t a equação de observação fica da forma:

$$Y_{it} = q_i X_t + v_{it}, \quad Y_t \geq 0, \quad q_i > 0, \quad i = 1, \dots, m, \quad (3.18)$$

onde v_{it} 's são variáveis aleatórias independentes. Desta forma, estas equações permitem a utilização da abordagem espaço de estados, incluindo as especificações para as distribuições dos termos de erro de processo e de observação, além das condições iniciais do sistema. Montenegro (2010) utilizou este modelo com distribuições lognormal para ambos termos de erro e um enfoque Bayesiano para a estimação dos parâmetros e estados não observáveis (ver também Meyer e Millar, 1999 e Millar e Meyer, 2000).

3.5 Algumas extensões

Uma versão generalizada do modelo dinâmico logístico tem a forma:

$$\frac{dX_t}{dt} = \frac{r}{p} X_t \left[1 - \left(\frac{X_t}{K} \right)^p \right] \quad (3.19)$$

onde p é o parâmetro de assimetria da curva de produção. Quando $p=1$ temos o modelo logístico de dois parâmetros na qual a curva de produção é simétrica em torno de $K/2$. Na versão generalizada, a taxa de mudança da população não é simétrica em torno do máximo (Haddon, 2001).

A versão em equação de diferenças finitas de ordem um, com remoções (ou harvesting) fica da forma:

$$X_{t+1} = X_t + \frac{r}{p} X_t \left[1 - \left(\frac{X_t}{K} \right)^p \right] - H_t \quad (3.20)$$

Esta equação é conhecida também como o modelo de biomassa dinâmica de Pella e Tomlinson (Pella e Tomlinson, 2008).

Outra extensão interessante é aquela que considera um modelo hierárquico, onde a captura (H_t) é um processo estocástico não observável e os dados de desembarques (D_t), são considerados uma série observável que se relaciona linearmente com as capturas. Desta maneira ao modelo especificado em 3.14, 3.15, acrescentamos duas equações: a equação de evolução das capturas através de um passeio aleatório da forma

$$H_{t+1} = H_t + \epsilon_t, \quad H_t \geq 0, \quad t = 1, \dots, n, \quad (3.21)$$

onde ϵ_t é um ruído branco. Por sua parte, a equação de observação que relaciona as variáveis aleatórias observáveis com as não observáveis é

$$D_t = \alpha H_t + \xi_t, \quad D_t \geq 0, \quad \alpha > 0, \quad t = 1, \dots, n, \quad (3.22)$$

onde ξ_t é um ruído branco.

Capítulo 4

Distribuições assimétricas e de caudas pesadas

4.1 Introdução

Em biologia de populações geralmente é considerada a distribuição normal para os erros aditivos e a distribuição log-normal para os erros multiplicativos (Millar e Meyer, 2000; de Valpine, 2002; Dennis et al, 2006; Maunder e Deriso, 2011), isto é motivado pelas boas propriedades da distribuição normal em modelos lineares. No entanto, a presença de assimetria nas variáveis ecológicas nos motivou a usar outras alternativas que podem acomodar melhor assimetria e a presença de observações atípicas. Classes mais flexível de distribuições de probabilidade têm sido propostas na literatura para lidar com estes dois pontos (ver, por exemplo, Genton, 2004). Uma breve revisão de alguns destes modelos é apresentada neste capítulo.

Considere a notação usual ϕ e Φ para a função de densidade de probabilidade (pdf) e a função de distribuição acumulativa (cdf) de uma variável aleatória normal, respectivamente, t_ν e T_ν indicam a pdf e cdf de uma distribuição t-student com ν graus de liberdade (df).

Em geral dizemos que uma distribuição de probabilidade é simétrica em torno de x_0 , se e somente se, existe um valor x_0 tal que $f(x_0 - \delta) = f(x_0 + \delta)$ para todo número real δ . f é a função de densidade de probabilidade, se a variável for contínua, ou função de massa de probabilidade, se for discreta. Caso contrário, dizemos que a função de distribuição é assimétrica. Uma classe geral de distribuições assimétricas flexíveis é a distribuição Skew-normal (SN). O primeiro tratamento sistemático desta classe de distribuições foi feito por Azzalini (1995,1996). A partir desta classe, posteriormente foram estendidas outras distribuições de probabilidade como a Skew-t e, mais geral, a classe de mistura na escala da normal assimétrica (SMSN, Branco e Dey, 2001). Neste capítulo, apresentamos algumas distribuições desta classe, como também as versões para erros multiplicativos com suporte em $\mathbb{R}^+$ da família log-SMSN (Marchenko e Genton, 2010).

4.2 Distribuições com suporte em $\mathbb{R}$

4.2.1 Distribuição Skew-Normal

Definição 4.1. Uma variável aleatória Y tem distribuição skew-normal com parâmetros de posição μ , de escala σ^2 e forma λ ($Y \sim SN(\mu, \sigma^2, \lambda)$) se sua função densidade de probabilidade é dada por:

$$p(y) = \frac{2}{\sigma} \phi\left(\frac{y - \mu}{\sigma}\right) \Phi\left(\lambda \frac{y - \mu}{\sigma}\right), \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \lambda \in \mathbb{R}, \sigma > 0. \quad (4.1)$$

O parâmetro de forma λ está associado a assimetria da distribuição. Valores negativos de λ indicam assimetria negativa e valores positivos, assimetria positiva. Se $\lambda = 0$ a densidade acima

coincide com a densidade de uma normal.

A função de distribuição associada a densidade acima é denotada por $F(y; \mu, \sigma^2, \lambda)$ e dada por:

$$F(y; \mu, \sigma^2, \lambda) = 2\Phi_2\left(\frac{y - \mu}{\sigma}, 0 \mid \Omega\right), \quad (4.2)$$

onde $\Phi_2(\cdot \mid \Omega)$ é a função de distribuição de uma normal bivariada, com média zero e matriz de covariâncias Ω .

A seguir apresentamos as representações estocásticas e construção por condicionamento da normal assimétrica padrão (ver Bayes, 2005) e algumas outras propriedades.

- Representação estocástica

Se Y_1 e Y_2 são v.as. independentes com distribuição normal padrão, então

$$Z = \delta|Y_1| + \sqrt{1 - \delta^2} Y_2 \text{ tem distribuição } SN(\lambda) \text{ com } \delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}. \quad (4.3)$$

- Construção por condicionamento

Se $Y = (Y_1, Y_2)^T \sim N_2(0, \Omega)$ e $|\delta| < 1$, então $Z = Y_2 \mid Y_1 > 0 \sim SN(\lambda)$.

- Representação hierárquica

$$\text{Se } Y|U \sim N(\mu + \sigma\delta U, \sigma^2(1 - \delta^2)) \text{ e} \quad (4.4)$$

$$U \sim N(0, 1)I_{(u>0)}, \quad (4.5)$$

então $Y \sim SN(\mu, \sigma^2, \lambda)$, onde $I(\cdot)$ é a função indicadora.

- A função geradora de momentos da v.a. $Y \sim SN(\mu, \sigma^2, \lambda)$ é dada por:

$$M_Y(t) = 2\exp(t\mu + \frac{t^2\sigma^2}{2})\Phi(\delta\sigma t). \quad (4.6)$$

- A esperança e variância de $Y \sim SN(\mu, \sigma^2, \lambda)$ são, respectivamente

$$E(Y) = \mu + \sigma\sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1 + \lambda^2}} \text{ e } V(Y) = \sigma^2\left(1 - \frac{2}{\pi}\frac{\lambda^2}{1 + \lambda^2}\right). \quad (4.7)$$

- Os coeficientes de assimetria e kurtosis são dados, respectivamente, por

$$\gamma_1 = \sqrt{\frac{2}{\pi}}\left(\frac{4}{\pi} - 1\right)\left(\frac{\lambda}{\sqrt{1 + \lambda^2}}\right)^3\left(1 - \frac{2}{\pi}\frac{\lambda^2}{1 + \lambda^2}\right)^{-3/2}, \quad (4.8)$$

$$\gamma_2 = \frac{8}{\pi^2}(\pi - 3)\left(\frac{\lambda}{\sqrt{1 + \lambda^2}}\right)^4\left(1 - \frac{2}{\pi}\frac{\lambda^2}{1 + \lambda^2}\right)^{-2}. \quad (4.9)$$

Figura 4.1: Funções de densidade da normal assimétrica com valores positivos e negativos de λ .

4.2.2 Distribuição Skew-t

Definição 4.2. Uma variável aleatória Y tem distribuição skew-t com parâmetros de posição μ , de escala σ^2 , parâmetro de forma λ e ν graus de liberdade ($Y \sim ST(\mu, \sigma^2, \lambda, \nu)$) se sua função densidade de probabilidade é dada por:

$$p(y) = \frac{2}{\sigma} t_{\nu} \left(\frac{y - \mu}{\sigma} \right) T_{\nu+1} \left(\frac{\lambda(y - \mu)}{\sigma} \sqrt{\frac{\nu + 1}{d^2(y) + \nu}} \right), \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \lambda \in \mathbb{R}, \nu > 0, \sigma > 0 \quad (4.10)$$

onde $d^2(y) = (y - \mu)^2 / \sigma^2$, t_{ν} e T_{ν} são as funções de densidade de probabilidade e de distribuição de uma v.a. t-student com ν graus de liberdade, respectivamente.

Se $X \sim SN(0, \sigma^2, \lambda)$ e $U \sim \text{Gamma}(\nu/2, \nu/2)$ variáveis aleatórias independentes, dizemos que a distribuição de $Y = \mu + U^{-1/2}X$ é t-student assimétrica (skew-t). Assim a skew-t é uma mistura na escala da normal assimétrica, com o fator de escala uma variável aleatória com distribuição Gamma. Note que $Y|U = u \sim SN(\mu, u^{-1}\sigma^2, \lambda)$.

A seguir apresentamos a representação hierárquica de uma variável aleatória skew-t.

Seja $Y \sim ST(\mu, \sigma^2, \lambda, \nu)$, então Y admite a seguinte representação hierárquica

$$\begin{aligned} Y|T = t, U = u &\sim N(\mu + \Delta t, u^{-1}\Gamma); \\ T|U = u &\sim HN(0, u^{-1}); \\ U &\sim \text{Gamma}(\nu/2, \nu/2), \end{aligned} \quad (4.11)$$

onde $\Delta = \sigma\delta$ e $\Gamma = \sigma^2(1 - \delta^2)$, com $\delta = \lambda/\sqrt{1 + \lambda^2}$. Denotamos $HN(0, \sigma^2)$ a uma variável aleatória com distribuição *half normal*, ou seja é uma $TN(\mu, \sigma^2; (a, b))$ (normal truncada) no intervalo $(0, \infty)$ e $\mu = 0$. A prova deste resultado pode ser encontrada, por exemplo, em Lachos et al., 2013.

Em relação aos momentos de $Y \sim ST(\mu, \sigma^2, \lambda, \nu)$, temos que a esperança e a variância são dados, respectivamente por

$$E[Y] = \mu + (\nu/\pi)^{1/2} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \sigma \delta, \quad \nu > 1, \quad (4.12)$$

$$Var[Y] = \frac{\nu}{\nu-2} \sigma^2 - (\nu/\pi) \left(\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \right)^2 \sigma^2 \delta^2, \quad \nu > 2. \quad (4.13)$$

Os coeficientes de assimetria e kurtosis são:

$$\gamma_1 = \eta \left[\frac{\nu(3-\delta^2)}{\nu-3} - \frac{3\nu}{\nu-2} + 2\eta^2 \right] \left[\frac{\nu}{\nu-2} - \eta^2 \right]^{-3/2} \quad \nu > 3, \quad (4.14)$$

$$\gamma_2 = \left[\frac{3\nu^2}{(\nu-2)(\nu-4)} - \frac{4\eta^2\nu(3-\delta^2)}{\nu-3} + \frac{6\eta^2\nu}{\nu-2} - 3\eta^4 \right] \left[\frac{\nu}{\nu-2} - \eta^2 \right]^{-2} - 3 \quad \nu > 4, \text{ onde} \quad (4.15)$$

$$\eta = \delta(\nu/\pi)^{1/2} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}. \quad (4.16)$$

Figura 4.2: Funções de densidade da skew-t. Esquerda, valores fixos de $\lambda = 5$ e variáveis de ν , direita, valores fixos de $\nu = 5$ e variáveis de λ .

4.2.3 Distribuições de Mistura na Escala da Normal Assimétrica (SMSN)

Definição 4.3. Uma variável aleatória Y tem distribuição de mistura na escala da normal ($Y \sim SMN(\mu, \sigma^2, \nu)$) se sua função densidade de probabilidade é dada por:

$$p(y|\mu, \sigma^2, \nu) = \int_0^\infty \phi(y|\mu, u^{-1}\sigma^2) h(u|\nu) du \quad (4.17)$$

onde U é o fator de escala e $h(\cdot|\nu)$ é a densidade misturadora.

A distribuição de $Y = \mu + U^{-1/2}X$, onde $\mu \in \mathbb{R}$, $X \sim N(0, \sigma^2)$ e U uma variável aleatória positiva, independente de X , com densidade $h(\cdot|\nu)$, é uma mistura de escala da normal, sendo U o fator de escala e $h(\cdot|\nu)$ a densidade misturadora, já que $Y|U = u \sim N(\mu, u^{-1}\sigma^2)$. Aplicações desta

classe de modelos podem ser encontradas em Lachos et al., (2011) e Garay et al., (2011).

Definição 4.4. *Uma variável aleatória Y tem distribuição de mistura na escala da skew-normal ($Y \sim SMSN(\mu, \sigma^2, \lambda, \nu)$) se sua função densidade de probabilidade é dada por:*

$$p(y|\mu, \sigma^2, \lambda, \nu) = 2 \int_0^\infty \phi(y|\mu, u^{-1}\sigma^2) \Phi(u^{1/2}\lambda\sigma^{-1}(y - \mu)) h(u|\nu) du \quad (4.18)$$

onde $h(\cdot|\nu)$ a densidade misturadora e U o fator de escala.

Assim, versões assimétricas das distribuições que estão na família SMSN podem ser construídas. Em particular, temos:

- Se $P(U = 1) = 1$, temos a distribuição skew-normal.
- Se $U \sim \text{Gamma}(\nu/2, \nu/2)$, com $\nu > 0$, temos a distribuição skew-t.
- Se $U \sim \text{beta}(\nu, 1)$, com $\nu > 0$, temos a distribuição skew-slash com densidade dada por:

$$p(y|\mu, \sigma^2, \lambda, \nu) = 2\nu \int_0^1 u^{\nu-1} \phi(y|\mu, u^{-1}\sigma^2) \Phi(u^{1/2}\lambda\sigma^{-1}(y - \mu)) du \quad (4.19)$$

- Se U é binária com $P(U = \tau) = \rho = 1 - P(U = 1)$, onde $0 < \tau < 1$, $0 < \rho \leq 1$, temos a distribuição skew-normal contaminada.

Se fizermos $\lambda = 0$ nos casos acima, teremos as versões simétricas destas distribuições. A seguir apresentamos a representação hierárquica de uma variável aleatória SMSN.

Seja $Y \sim SMSN(\mu, \sigma^2, \lambda, \nu)$. Então Y admite a seguinte representação hierárquica

$$\begin{aligned} Y|T = t, U = u &\sim N(\mu + \Delta t, u^{-1}\Gamma); \\ T|U = u &\sim TN(b, u^{-1}; (b, \infty)); \\ U &\sim h(\cdot|\nu), \end{aligned} \quad (4.20)$$

onde $\Delta = \sigma\delta$ e $\Gamma = \sigma^2(1 - \delta^2)$, com $\delta = \lambda/\sqrt{1 + \lambda^2}$. Denotamos $TN(\mu, \sigma^2; (a, b))$ a uma variável aleatória com distribuição normal truncada no intervalo (a, b) , com média e variância antes do truncamento μ e σ^2 . A prova deste resultado pode ser encontrada em Lachos et al., 2013.

4.3 Distribuições com suporte em $\mathbb{R}^+$

4.3.1 Distribuição Log-Skew-Normal Univariada

Definição 4.5. *Y tem distribuição log-skew-normal se $Z = \ln Y \sim SN(\mu, \sigma^2, \lambda)$, e sua fdp é dada por:*

$$p(y) = \frac{2}{\sigma y} \phi\left(\frac{\ln y - \mu}{\sigma}\right) \Phi\left(\lambda\left(\frac{\ln y - \mu}{\sigma}\right)\right), y > 0. \quad (4.21)$$

Denota-se $Y \sim LSN(\mu, \sigma^2, \lambda)$.

Como caso particular temos a distribuição lognormal, quando assumimos $\lambda = 0$, sua fdp é dada por:

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}}, y > 0. \quad (4.22)$$

Figura 4.3: Funções de densidade de distribuições na família SMSN.

Denota-se $Y \sim LN(\mu, \sigma^2)$.

A esperança e variância de $Y \sim LSN(\mu, \sigma^2, \lambda)$ são:

$$E(Y) = 2e^{(\mu + \frac{\sigma^2}{2})} \Phi\left(\frac{\lambda\sigma}{\sqrt{1 + \lambda^2}}\right), \quad (4.23)$$

$$Var(Y) = 2e^{(2\mu + 2\sigma^2)} \left[e^{\sigma^2} \Phi\left(\frac{2\lambda\sigma}{\sqrt{1 + \lambda^2}}\right) - 2 \left[\Phi\left(\frac{\lambda\sigma}{\sqrt{1 + \lambda^2}}\right) \right]^2 \right]. \quad (4.24)$$

4.3.2 Distribuição Log-Skew-t Univariada

Definição 4.6. Uma variável aleatória Y tem distribuição log-skew-t, denotada por $LST(\mu, \sigma, \lambda, \nu)$ se sua fdp é dada por

$$p(y) = \frac{2}{\sigma y} t_\nu\left(\frac{\ln y - \mu}{\sigma}\right) T_{\nu+1}\left(\frac{\lambda(\ln y - \mu)}{\sigma} \sqrt{\frac{\nu + 1}{d^2(y) + \nu}}\right), \quad y > 0, \mu \in \mathbb{R}, \lambda \in \mathbb{R}, \nu > 0, \sigma > 0, \quad (4.25)$$

onde $d^2(y) = (\ln y - \mu)^2 / \sigma^2$ e T é a função de distribuição da t -student.

É interessante observar que não existem os momentos das distribuições log-t e log-skew-t (Marchenko e Genton, 2010). No presente trabalho de tese utilizamos as distribuições Log-Normal, Log-t, Log-skew-normal e Log-skew-t univariadas, todas pertencentes a família log-SMSN.

Figura 4.4: Densidades de probabilidade das distribuições log-skew-normal com $\mu = 0$, $\sigma = 1$ e diferentes valores de λ .

Capítulo 5

Inferência no modelo dinâmico logístico

5.1 Introdução

A inferência nos modelos dinâmicos não lineares, tanto na estimação de parâmetros como no teste de hipóteses, não tem forma analítica fechada. Algumas alternativas de aproximação simples podem ser utilizadas em casos de normalidade dos erros de observação e de processo, quando as equações de observação ou de evolução são não lineares. Várias técnicas de linearização tem sido desenvolvidas para modelos não lineares gaussianos, por exemplo através de aproximação de primeira ordem de séries de Taylor para as funções não lineares. Para uma exposição detalhada dessas técnicas pode-se consultar West e Harrison (1997), Welch e Bishop (2001) e Durbin e Koopman (2012). No obstante, esta técnica tem grandes limitações quando no modelo além dos estados não observáveis, existem parâmetros desconhecidos. Nesse caso, é necessário técnicas de aproximação, tanto nos paradigmas clássico como Bayesiano. A seguir apresentamos os procedimentos de inferência considerados neste trabalho de tese. Os procedimentos são:

- Inferência via filtro de Kalman estendido (EKF)
- Inferência via aproximação de Laplace
- Inferência Bayesiana com o uso de MCMC

Antes de apresentar o detalhe dos procedimentos de inferência utilizados nesta tese, é importante apontar que nos modelos espaço de estados as quantidades de interesse desconhecidas podem ser divididas em duas classes: os estados não observáveis ou latentes (X_t) e os parâmetros das equações de observação e evolução (Θ).

5.2 Inferência via filtro de Kalman estendido (EKF)

A seguir apresentamos a inferência no modelo dinâmico logístico utilizando o EKF para o caso onde os erros de observação e processo são gaussianos e os parâmetros das equações de observação e evolução são conhecidos.

No caso do modelo de crescimento logístico de populações, a equação de observação é linear e a equação de evolução é uma função não linear. Lembre que a representação espaço de estados do modelo é:

$$\begin{aligned} X_t &= X_{t-1} + rX_{t-1} \left(1 - \frac{X_{t-1}}{K}\right) - H_{t-1} + u_t, \quad u_t \sim N(0, W_t), \quad t = 0, 1, \dots, \\ Y_t &= qX_t + v_t, \quad v_t \sim N(0, V_t), \quad t = 1, \dots, n, \end{aligned} \tag{5.1}$$

$$X_1|D_0 \sim N(m_0, C_0).$$

Fazendo a expansão em séries de Taylor de primeira ordem, em torno da média de X_{t-1} , m_{t-1} , na equação de evolução, temos

$$X_t \approx f_t(m_{t-1}) + F_t[X_{t-1} - m_{t-1}] + u_t = X_t^L, \quad (5.2)$$

onde

$$f_t(m_{t-1}) = m_{t-1} + rm_{t-1} \left(1 - \frac{m_{t-1}}{K}\right) - H_{t-1} \quad \text{e} \quad (5.3)$$

$$F_t = \left[\frac{\partial f_t(X_{t-1})}{\partial X_{t-1}} \right]_{X_{t-1}=m_{t-1}} = 1 - r - \frac{2r}{K}m_{t-1}. \quad (5.4)$$

Considerando a suposição de erros gaussianos, $X_t^L|D_{t-1} \sim N(a_t, R_t)$ onde:

$$a_t = E[f_t(m_{t-1}) + F_t[X_{t-1} - m_{t-1}] + u_t] = f_t(m_{t-1}), \quad (5.5)$$

$$R_t = V[f_t(m_{t-1}) + F_t[X_{t-1} - m_{t-1}] + u_t] = F_t^2 C_{t-1} + W_t. \quad (5.6)$$

onde C_{t-1} é a variância condicional de $X_{t-1}^L|D_{t-1}$.

A previsão um passo a frente para as observações é

$Y_t|D_{t-1} \sim N(f_t, Q_t)$, onde a partir da Proposição 2.2 (ii), se tem

$$\begin{aligned} f_t &= E(Y_t|Y_{1:t-1}) = qa_t, \\ Q_t &= Var(Y_t|Y_{1:t-1}) = q^2 R_t + V_t. \end{aligned}$$

A densidade atualizada dos estados é

$X_t^L|D_t \sim N(m_t, C_t)$ onde a partir da Proposição 2.2 (iii), se tem

$$m_t = E(X_t|Y_{1:t}) = m_{t-1} + rm_{t-1} \left(1 - \frac{m_{t-1}}{K}\right) - H_{t-1} + A_t e_t, \quad (5.7)$$

onde

$$A_t = \frac{qR_t}{Q_t}, \quad e_t = Y_t - qa_t, \quad (5.8)$$

$$C_t = Var(X_t|Y_{1:t}) = R_t - \frac{q^2 R_t^2}{Q_t}. \quad (5.9)$$

5.3 Inferência através da aproximação de Laplace

A inferência no modelo logístico utilizando aproximação de Laplace, tem sua motivação no uso do filtro de Kalman estendido (EKF). Nos casos de normalidade dos erros de observação e evolução e funções não lineares de observação e processo, pode-se obter forma fechada para as distribuições *a posteriori* de interesse, utilizando o EKF. Isto funciona sem problema quando os parâmetros destas funções são conhecidos. Se existem parâmetros desconhecidos, já não temos mais formas fechadas para as quantidades de interesse. Neste caso precisamos de métodos aproximados de estimação dos parâmetros e estados não observáveis.

Uma alternativa de grande utilidade para fazer inferência aproximada é a aproximação de Laplace. Esta consiste em utilizar a aproximação de Laplace para integrar a verossimilhança completa ao respeito dos estados latentes (Tierney e Kadane, 1986). Esta tem a forma geral

$$\int e^{nf(x)} dx \approx e^{nf(x^*)} \int \exp \left[-\frac{n(x-x^*)^2}{2\sigma^2} \right] dx = \sqrt{2\pi\sigma n^{-1/2}} e^{nf(x^*)} \quad (5.10)$$

onde x^* é o máximo da função $f(x)$ e $\sigma^2 = -1/f''(x^*)$. Para a nossa situação específica de interesse a função $f(x)$ é a log-verossimilhança completa.

Supondo independência condicional das observações, dado os estados não observáveis, a distribuição conjunta é

$$p(y_{1:n} | q, \tau^2, \theta, x_{1:n}) = \prod_{t=1}^n p(y_t | x_t, q, \theta). \quad (5.11)$$

onde $y_{1:n}$ é uma amostra da variável aleatória Y_t e θ corresponde ao vetor de parâmetros da distribuição de probabilidade dos erros de observação específica. A função de verossimilhança completa, ou seja, a verossimilhança conjunta das observações e dos estados latentes (não observáveis) é dada por

$$L(\Theta, x_{1:n}; y_{1:n}) = p(x_1 | \tau^2) \prod_{t=2}^n p(x_t | x_{t-1}, K, r, \tau^2) \prod_{t=1}^n p(y_t | x_t, q, \theta). \quad (5.12)$$

onde Θ é o vetor de parâmetros do modelo ($\Theta = K, r, q, \tau^2, \theta$). Para obter a verossimilhança marginal temos que integrar a verossimilhança completa em relação aos estados latentes, resultando em

$$L(\Theta; y_{1:n}) = \int L(\Theta, x_{1:n}; y_{1:n}) dx_{1:n} = \int p(y_{1:n} | x_{1:n}, \Theta) p(x_{1:n} | \Theta) dx_{1:n}, \quad (5.13)$$

O primeiro passo para obter a aproximação de Laplace para 5.13 é obter o máximo da verossimilhança completa $L(\Theta, x_{1:n}; y_{1:n})$ com respeito a $\mathbf{x} = x_{1:n}$, ou seja

$$\hat{\mathbf{x}}_{\Theta} = \arg \max_{x_{1:n}} \ln L(\Theta, \mathbf{x}; y_{1:n}). \quad (5.14)$$

Este valor pode ser obtido utilizando-se o algoritmo quase-Newton (Kuri, 2003, ver apêndice para mais detalhes do algoritmo). Notemos que $\hat{\mathbf{x}}_{\Theta}$ depende de Θ .

O segundo passo é a obtenção da matriz Hessiana $H(\hat{\mathbf{x}}_{\Theta}, \Theta)$ dada por:

$$H(\hat{\mathbf{x}}_{\Theta}, \Theta) = - \frac{\partial^2 \ln L(y_{1:n}, \Theta, \mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{\Theta}}. \quad (5.15)$$

Assim, a aproximação de Laplace da verossimilhança marginal é

$$L^*(\Theta; y_{1:n}) = |\det\{H(\hat{\mathbf{x}}_{\Theta}, \Theta)\}|^{-1/2} L(\Theta, \hat{\mathbf{x}}_{\Theta}; y_{1:n}). \quad (5.16)$$

Logo, a aproximação de Laplace a log-verossimilhança marginal é:

$$l^*(\Theta; y_{1:n}) = -0.5 \ln |\det\{H(\hat{\mathbf{x}}_{\Theta}, \Theta)\}| + g(y_{1:n}, \Theta, \hat{\mathbf{x}}_{\Theta}). \quad (5.17)$$

onde $g(y_{1:n}, \Theta, \mathbf{x}) = \ln\{L(\Theta, x_{1:n}; y_{1:n})\}$.

Encontrar o vetor de parâmetros Θ , que maximizam a aproximação de Laplace é um problema de otimização aninhada. Ou problema *interior* já foi descrito, e corresponde a aproximar a verossimilhança marginal via Laplace. O problema *exterior* é encontrar os valores que maximizam a

log-verossimilhança marginal, ou seja,

$$\hat{\Theta} = \arg \max_{\Theta} \ln L^*(\Theta; y_{1:n}). \quad (5.18)$$

Estes valores podem ser obtidos utilizando novamente o algoritmo quase-Newton, o qual precisa da derivada da log-verossimilhança marginal. Para obter a derivada de $l^*(\Theta; y_{1:n})$ em relação a Θ , precisamos a derivada de $\hat{\mathbf{x}}_{\Theta}$ em relação a Θ ($\frac{\partial \hat{\mathbf{x}}_{\Theta}}{\partial \Theta}$). Para isto precisamos fazer a derivada da função implícita (ver Apêndice A1), da qual se resulta

$$\frac{\partial \hat{\mathbf{x}}_{\Theta}}{\partial \Theta} = -\{H(\hat{\mathbf{x}}_{\Theta}, \Theta)\}^{-1} \frac{\partial^2 g(y_{1:n}, \Theta, \hat{\mathbf{x}}_{\Theta})}{\partial \mathbf{x} \partial \Theta} \equiv -\left\{ \frac{\partial^2 g(y_{1:n}, \Theta, \hat{\mathbf{x}}_{\Theta})}{\partial \mathbf{x} \partial \mathbf{x}'} \right\}^{-1} \frac{\partial^2 g(y_{1:n}, \Theta, \hat{\mathbf{x}}_{\Theta})}{\partial \mathbf{x} \partial \Theta} \quad (5.19)$$

A primeira derivada da log-verossimilhança marginal obtém-se a usando a regra da cadeia sob a função implícita, da qual se obtém

$$\frac{\partial l^*(y_{1:n}, \Theta, \hat{\mathbf{x}}_{\Theta})}{\partial \Theta} = \left\{ \frac{\partial \hat{\mathbf{x}}_{\Theta}}{\partial \Theta} \right\}^T \left\{ \frac{\partial l^*}{\partial \hat{\mathbf{x}}_{\Theta}} \right\} + \left\{ \frac{\partial l^*}{\partial \Theta} \right\} \quad (5.20)$$

Assim, a informação de Fisher observada se obtém a partir de $\frac{\partial}{\partial \Theta} \left\{ \frac{\partial l^*(y_{1:n}, \Theta, \hat{\mathbf{x}}_{\Theta})}{\partial \Theta} \right\}$, o que pode ser obtido através de diferenças finitas (Skaug and Fournier, 2006).

O procedimento de aproximação de Laplace pode ser resumido em:

1. inicializar o vetor Θ com valores arbitrários $\Theta^k = \Theta^0$,
2. com o valor $\Theta^k = \Theta^0$ otimizamos a verossimilhança com respeito aos estados não observáveis $\mathbf{x}$, para obter $\hat{\mathbf{x}}_{\Theta}$ e a Hessiana $H(\hat{\mathbf{x}}_{\Theta})$ (através do algoritmo quase-Newton),
3. usamos $\hat{\mathbf{x}}_{\Theta^k}$ e $H(\hat{\mathbf{x}}_{\Theta^k})$ para aproximar a log-verossimilhança marginal,
4. obtemos o valor da log-verossimilhança marginal e seu gradiente ($S(\Theta^{(k)})$, 5.20),
5. se o gradiente é maior que o valor de tolerância ϵ , o valor de Θ atualiza-se com o algoritmo quase-Newton da forma:

$$\Theta^{(k+1)} = \Theta^{(k)} + H(\Theta^{(k)})^{-1} S(\Theta^{(k)}).$$

Com este novo valor de $\Theta^{(k)}$ retorna-se ao passo 2. Caso contrario, o algoritmo convergiu e os valores de $\Theta^{(k)}$ são as estimativas de máxima verossimilhança finais ($\Theta^{(k)} = \hat{\Theta}$).

Usamos a plataforma de software ADMB (admb-project.org, Automatic Differentiation Model Builder, Skaug e Fournier, 2006), com seu modulo ADMB-RE (Random Effects) para estimar os parâmetros e estados não observáveis do modelo de crescimento logístico.

5.4 Inferência Bayesiana através de MCMC

A inferência Bayesiana no modelo de crescimento logístico depende da especificação da distribuição *a priori* conjunta dos parâmetros e a verossimilhança. Supondo independência condicional das observações, dado os estados não observáveis, a distribuição conjunta das observações está dada em 5.11. Por sua vez a verossimilhança completa, ou seja, a verossimilhança conjunta das observações e dos estados latentes (não observáveis) está dada em 5.12.

Assumindo independência entre os estados não observáveis e os parâmetros do modelo e independência *a priori* entre os parâmetros do modelo, a densidade *a priori* conjunta é:

$$p(K, r, q, \tau^2, \theta) = p(K)p(r)p(q)p(\tau^2)p(\theta), \quad (5.21)$$

onde (normal, skew-normal, t-student e skew-t, para os erros aditivos e log-normal, log-skew-normal, log-t e log-skew-t para os erros multiplicativos). Os outros parâmetros já foram definidos previamente.

Logo a distribuição *a posteriori* conjunta dos parâmetros e estados não observáveis é:

$$p(\Theta, x_{1:n} \mid y_{1:n}) \propto p(\Theta)p(x_1 \mid \tau^2) \prod_{t=2}^n p(x_t \mid x_{t-1}, K, r, \tau^2) \prod_{t=1}^n p(y_t \mid x_t, q, \theta), \quad (5.22)$$

5.4.1 Equação de estado

A equação de estado, também conhecida como equação de evolução do sistema, tem dois componentes: a componente funcional do modelo logístico e a componente aleatória (incerteza do processo). Este ruído de processo pode ter várias fontes e em populações biológicas provem principalmente da variabilidade demográfica e/ou variabilidade ambiental (Dennis et al., 2006; Mauner e Deriso, 2011).

Neste trabalho, para a equação de estado utilizamos a hipótese de erros de processo aditivos, normalmente distribuídos e com variância constante, ou seja, $u_t \sim N(0, \tau^2)$. Sendo assim, a probabilidade dos estados é:

$$p(x_t \mid x_{t-1}, K, r, \tau^2) = \frac{1}{\sqrt{2\pi\tau}} \exp \left(-\frac{(x_t - [x_{t-1} + rx_{t-1}(1 - \frac{x_{t-1}}{K}) - H_{t-1}])^2}{2\tau^2} \right). \quad (5.23)$$

Adicionalmente, assume-se que a condição inicial de sistema é $X_1 \mid D_0 \sim N(\mu_0, \tau_0)$, onde D_0 representa o conjunto de informação inicial disponível.

5.4.2 Modelos Observacionais

No presente trabalho, utilizamos duas formulações gerais para a equação de observação: erros aditivos, isto é,

$$Y_t = qX_t + v_t, \quad Y_t \geq 0, \quad q > 0, \quad t = 1, \dots, n, \quad (5.24)$$

e erros multiplicativos,

$$Y_t = qX_t e^{v_t}, \quad Y_t \geq 0, \quad q > 0, \quad t = 1, \dots, n, \quad (5.25)$$

No modelo de erros aditivos, utilizamos as distribuições normal, t-student, skew-normal e skew-t. No caso dos erros de observação multiplicativos, utilizamos as distribuições log-normal, log-t, log skew-normal e log skew-t.

No caso mais simples de erros aditivos, consideramos v_t variáveis aleatórias normais, independentes e identicamente distribuídas, $v_t \sim N(0, \sigma^2)$. Assim, a densidade condicional conjunta das observações é:

$$p(y_{1:n} \mid q, x_{1:n}, \sigma^2) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y_t - qx_t)^2}{2\sigma^2}}. \quad (5.26)$$

A equação de observação para o caso de erros aditivos t-student considera em 5.24 $v_t \sim t(0, \sigma^2, \nu)$. Como consequência disto:

$$p(y_{1:n} | q, x_{1:n}, \sigma^2, \nu) = \frac{c(\nu)^n}{(\pi\nu\sigma^2)^{n/2}} \prod_{t=1}^n \left(1 + \frac{(y_t - qx_t)^2}{\nu\sigma^2} \right)^{-\frac{\nu+1}{2}}, \quad (5.27)$$

onde $c(\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}$. Usando a representação hierárquica da distribuição t-student temos

$$\begin{aligned} Y_t | q, X_t, \sigma^2, \nu, U_t &\sim N(qX_t, U_t^{-1}\sigma^2) \text{ e} \\ U_t &\sim \text{Gamma}(\nu/2, \nu/2). \end{aligned} \quad (5.28)$$

Na equação de observação para o modelo de erros aditivos skew-normais, v_t são variáveis aleatórias independentes e identicamente distribuídas com $\mathbb{E}(v_t) = 0$, portanto $v_t \sim SN(\mu^*, \sigma^2, \lambda)$ onde $\mu^* = -\Delta\sqrt{\frac{2}{\pi}}$, $\Delta = \sigma\frac{\lambda}{\sqrt{1+\lambda^2}}$. Como consequência

$$p(y_{1:n} | q, x_{1:n}, \sigma, \lambda) = 2^n \sigma^{-n} \prod_{t=1}^n \phi\left(\frac{y_t - qx_t - \mu^*}{\sigma}\right) \Phi\left(\lambda \frac{y_t - qx_t - \mu^*}{\sigma}\right). \quad (5.29)$$

Usando a representação hierárquica da skew-normal temos

$$\begin{aligned} Y_t | q, X_t, \sigma, \lambda, U_t &\sim N(qX_t + \Delta U_t, \sigma^2(1 - \delta^2)) \text{ e} \\ U_t &\sim N(0, 1)I_{(u>0)}. \end{aligned} \quad (5.30)$$

Por sua vez, na equação de observação para os erros aditivos skew-t v_t são variáveis aleatórias independentes e identicamente distribuídas com $\mathbb{E}(v_t) = 0$, ou seja, $v_t \sim ST(\mu^* \kappa_1, \sigma^2, \lambda, \nu)$ onde $\kappa_1 = \sqrt{\frac{\nu}{2}} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}$. Consequentemente

$$p(y_{1:n} | q, x_{1:n}, \sigma, \lambda, \nu) = 2^n \sigma^{-n} \prod_{t=1}^n t_\nu\left(\frac{y_t - qx_t - \mu^* \kappa_1}{\sigma}\right) T_{\nu+1}\left(\frac{\lambda(y_t - qx_t - \mu^* \kappa_1)}{\sigma} \sqrt{\frac{\nu+1}{d^2(y_t) + \nu}}\right). \quad (5.31)$$

A representação hierárquica deste modelo é dada por

$$\begin{aligned} Y_t | U_t, T_t &\sim N(qX_t + \Delta T_t, U_t^{-1}\Gamma); \\ T_t | U_t &\sim N(0, u^{-1})I_{(u>0)}; \\ U_t &\sim \text{Gamma}(\nu/2, \nu/2), \end{aligned} \quad (5.32)$$

onde $\Gamma = \frac{\sigma^2}{1+\lambda^2}$.

Para a equação de observação com erros multiplicativos (5.25), podemos considerar sua versão transformada em escala log da forma:

$$\log(Y_t) = \log(q) + \log(X_t) + v_t \quad t = 1, \dots, n. \quad (5.33)$$

Para o propósito deste trabalho de tese, utilizaremos esta equação de observação para os modelos de erros multiplicativos log-normal, log-t, log-skew-normal e log-skew-t e para cada um destes, utilizamos os dados transformando em escala log e, portanto v_t na equação (5.33) são variáveis aleatórias normais, t-student, skew-normal e skew-t, respectivamente.

5.4.3 Aproximação das distribuições *a posteriori*

A distribuição *a posteriori* dos parâmetros se obtém integrando sob os estados não observáveis, da forma:

$$p(\Theta | y_{1:n}) \propto \int p(y_{1:n} | x_{1:n}, \Theta) p(x_{1:n} | \Theta) dx_{1:n} p(\Theta) \quad (5.34)$$

Por sua vez, a distribuição *a posteriori* dos estados não observáveis se obtém integrando *a posteriori* conjunta sobre os parâmetros do modelo.

$$p(x_{1:n} | y_{1:n}) = \int p(x_{1:n} | y_{1:n}, \Theta) p(\Theta | y_{1:n}) d\Theta. \quad (5.35)$$

Considerando que ditas distribuições não estão disponíveis em forma fechada, neste trabalho de tese, utilizamos métodos MCMC para aproxima-las. Para isso, utilizamos o software OpenBugs (Thomas et al., 2006) que foi executado desde o ambiente R (R Core Team, 2014), utilizando a livreria R2OpenBUGS (Sturtz et al., 2005). OpenBugs utiliza o algoritmo de Gibbs, amostrando das distribuições condicionais completas. A maioria das distribuições condicionais completas tem forma analítica fechada e podem ser encontradas em Lachos et al., 2013. Algumas delas não tem forma fechada e, portanto, precisamos um passo do algoritmo Metropolis-Hastings. No Apêndice B se apresentam as densidades condicionais completas para o caso de erros de observação normais e lognormais.

A ideia básica para fazer inferência via MCMC em modelos dinâmicos é a seguinte: obter a distribuição *a posteriori* dos estados não observáveis ($X_{1:n}$) e do vetor de parâmetros (Θ) via simulação, gerando amostras da densidade conjunta $p(\Theta, x_{1:n} | y_{1:n})$. A amostragem desta densidade conjunta é feita através da construção da Cadeia de Markov cuja distribuição estacionária é $p(\Theta, x_{1:n} | y_{1:n})$. O algoritmo consiste em:

1. Fixar valores iniciais para o vetor Θ digamos $\Theta = \Theta^0$,
2. gerar amostras de $X_{1:n}^i$ da densidade $p(x_{1:n} | y_{1:n}, \Theta^{i-1})$,
3. gerar amostras de Θ^i da densidade $p(\Theta | y_{1:n}, x_{1:n}^i)$

para $i = 1, \dots, M$. Depois do período *burn-in*, podemos tratar amostra gerada no passo (3) como sendo uma amostra da densidade $p(\Theta | y_{1:n})$. Uma das características atraentes deste esquema MCMC é que a amostragem das densidades condicionais é mais fácil do que a amostragem das marginais. As circunstâncias nas quais as amostras subsequentes das densidades marginais $p(x_{1:n} | y_{1:n}, \Theta^{i-1})$ e $p(\Theta | y_{1:n}, x_{1:n}^i)$ convergem para amostras de densidade conjunta $p(\Theta, x_{1:n} | y_{1:n})$, são apresentadas por exemplo em Gamerman e Lopes (2006).

Foram simuladas 4 cadeias de 40.000 iterações, descartando as primeiras 10.000. Como resultado, obteve-se 120.000 amostras da distribuição *a posteriori*, que foram usadas para obter as estatísticas de resumo das quantidades de interesse. Para avaliar a convergência da simulação MCMC usamos a medida de redução de escala potencial ($\hat{R}$) de Gelman e Rubin (Gilks et al., 1996).

No Apêndice B se apresentam os códigos computacionais nas linguagens OpenBugs, ADMB e R.

5.5 Critérios de seleção do modelo

Foram considerados dois critérios para seleção de modelos: o *Deviance Information Criterion* (DIC), descrito por Spiegelhalter et al. (2002); e a distribuição *a posteriori* preditiva da autocorrelação dos resíduos (ver Gelman et. al 1996, para uma apresentação completa do uso da distribuição *a posteriori* preditiva para a seleção de modelos).

O critério DIC baseia-se na medida *deviance*, que é dada por:

$$D(\Theta_m) = -2 \ln p(y_{1:n}|\Theta_m) \quad (5.36)$$

onde $p(y_{1:n}|\Theta_m)$ é a função de verossimilhança do modelo m com vetor de parâmetros Θ_m . A medida DIC para o modelo m é da forma:

$$DIC(m) = 2\overline{D(\Theta_m)} - D(\overline{\Theta}_m) = D(\overline{\Theta}_m) + 2\rho_m, \quad (5.37)$$

onde $\overline{D(\Theta_m)}$ é a média *a posteriori* e $D(\overline{\Theta}_m)$ é a deviance avaliada na média *a posteriori* dos parâmetros envolvidos do modelo. ρ_m pode ser interpretado como o número efetivo de parâmetros do modelo m e é dado por

$$\rho_m = \overline{D(\Theta_m)} - D(\overline{\Theta}_m) = -2 \int [\ln p(y_{1:n}|\Theta_m) - \ln p(y_{1:n}|\overline{\Theta}_m)] p(\Theta_m|y_{1:n}) d\Theta \quad (5.38)$$

Notemos que o DIC pode ser equivalentemente escrito na forma:

$$DIC(m) = 2\overline{D(\Theta_m)} - D(\overline{\Theta}_m) = -4 \mathbb{E}_{\Theta|Y} [\ln p(y_{1:n}|\Theta_m)] + 2 \ln p(y_{1:n}|\overline{\Theta}_m) \quad (5.39)$$

O DIC se assemelha com o critério AIC de Akaike (1973) e pode ser interpretado como uma medida da qualidade do ajuste do modelo, mais uma medida de sua complexidade. Valores menores do DIC indicam melhor adequação do modelo aos dados.

Os *posterior predictive checks* foram utilizados para avaliar se as discrepâncias observadas são concordantes com as discrepâncias preditas. Isto requer amostrar da distribuição *a posteriori* preditiva, o que foi feito a partir das M réplicas da distribuição *a posteriori* preditiva simulada através de MCMC. As discrepâncias foram calculadas utilizando os dados observados $(y_t, t = 1, \dots, n)$ e os dados gerados aleatoriamente $(\tilde{y}_t^{(j)}, j = 1, \dots, M)$, onde M é o tamanho da amostra Monte Carlo de cada réplica da distribuição *a posteriori*. Assim foram calculados dois tipos de resíduos:

- os obtidos com a distribuição *a posteriori* preditiva: $\tilde{v}_t^{(j)} = \tilde{y}_t^{(j)} - q^{(j)} X_t^{(j)}$, e
- e os obtidos com os dados reais: $\hat{v}_t^{(j)} = y_t - q^{(j)} X_t^{(j)}$.

Logo, dois tipos de autocorrelação com $lag = 1$ foram calculadas:

- a autocorrelação de resíduos preditos: $\tilde{\rho}_j(\tilde{v}_t^{(j)}, \tilde{v}_{t+1}^{(j)})$, e
- a autocorrelação dos resíduos realizados: $\hat{\rho}_j(\hat{v}_t^{(j)}, \hat{v}_{t+1}^{(j)})$.

Finalmente, foi calculado o valor-p *a posteriori* preditivo (Gelman et al., 2014) como a proporção de pontos para os quais a autocorrelação predita excede a autocorrelação observada da forma:

$$\hat{p} = \frac{\sum_{j=1}^M \mathbf{I}_j \left(\tilde{\rho}_j(\tilde{v}_t^{(j)}, \tilde{v}_{t+1}^{(j)}) \geq \hat{\rho}_j(\hat{v}_t^{(j)}, \hat{v}_{t+1}^{(j)}) \right)}{M}, \quad (5.40)$$

onde $\mathbf{I}(\cdot)$ é a função indicadora. Estas quantidades foram calculadas para todos os modelos de erros de observação propostos. Valores próximos a 0 ou 1 do valor-p *a posteriori* preditivo indicam que é pouco provável que os dados provenham da distribuição *a posteriori*.

Capítulo 6

Estudo de simulação

6.1 Introdução

Foi desenvolvido um estudo de simulação com o propósito de testar as propriedades frequentistas da estimação dos parâmetros frente a diferentes extensões das séries temporais. Este estudo considerou as seguintes etapas:

1. Geração dos estados não observáveis com erro de processo normal. Foi simulada uma série temporal de estados não observáveis de tamanho 90 ($n=90$).
2. Geração das observações com erro de observação aditivo skew-t. Foram geradas 100 séries de observações a partir da série de estados não observáveis simulados.
3. Estimação dos parâmetros e estados não observáveis. Estimamos os parâmetros e estados não observáveis para todas as séries de observações simuladas, considerando tamanhos crescentes das séries de dados: $n=10, 20, 30, 50, 70$ e 90 . Foram utilizados os procedimentos de estimação Bayesianos sob as hipóteses de erros de observação: (a) skew-t, (b) skew-normal, (c) t-student e (e) normais.
4. Foram calculados para cada parâmetro a média, o viés médio (MB) e a raiz do erro quadrático médio ($RMSE$) dados por:

$$MB(\hat{\theta}) = \frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}_i - \theta), \quad RMSE(\hat{\theta}) = \sqrt{\frac{1}{100} \sum_{i=1}^{100} (\hat{\theta}_i - \theta)^2}. \quad (6.1)$$

onde $\hat{\theta}_i$ corresponde ao valor estimado do parâmetro na i -ésima amostra (θ representa genericamente cada um dos parâmetros do modelo). Os parâmetros do modelo de simulação são apresentados na Tabela 6.1.

Tabela 6.1: *Parâmetros do modelo de simulação para o modelo logístico. O modelo considera erro de processo normal e erro de observação skew-t.*

Parâmetro		Valor
Equação de processo	K	40
	r	0.4
	τ	3
Equação de observação	q	10
	σ	12
	λ	3
	ν	7

Figura 6.1: *Simulação de estados não observáveis com erro de processo normal. Série temporal de extensão 90 (acima) e simulação de 100 séries de observações com erro skew-t (abaixo).*

6.2 Resultados

Nas Figuras 6.2, 6.3, 6.4 e 6.5, se apresentam os resultados das estimativas dos estados não observáveis obtidas a partir do estudo de simulação. Na Figura 6.2 se apresenta os resultados das estimativas obtidas utilizando o procedimento de inferência Bayesiano via MCMC, descrito na seção 5.4, sob o suposto de erros de observação skew-t, na Figura 6.3 sob o suposto de erros de observação de skew-normal, na Figura 6.4 sob suposto de erros t-student e na Figura 6.5 sob o suposto de erros de observações normais. Para cada um destes cenários, se apresentam as estimativas dos estados não observáveis considerando 6 casos de comprimento das séries temporais: $n=10, 20, 30, 50, 70$ e 90 . Nestas figuras se resumem as estimativas dos estados não observáveis a partir das 100 séries temporais de observações simuladas, apresentando-se os percentis de 2,5%, 50,0% e 97,5% das médias *a posteriori*. Com o objetivo de apresentar a qualidade dos procedimentos de inferências propostos, se acrescenta nas figuras a série verdadeira de estados não observáveis.

Nas Tabelas 6.2, 6.3, 6.4 e 6.5, se apresentam as respectivas estimativas dos parâmetros do modelo logístico estimados com o procedimento de inferência Bayesiano via MCMC, já descrito. Apresenta-se para cada caso de comprimento da série temporal e parâmetro do modelo, a média das estimativas, viés médio (MB) e a raiz do erro quadrático médio (RMSE). Para o cenário de erros skew-t, se apresentam estas estatísticas para todos os parâmetros, devido ao que os dados simulados foram gerados a partir desta distribuição de probabilidade, no entanto para os outros cenários somente se apresentam as estatísticas já citadas para os parâmetros do modelo logístico (K, r, τ, q).

Um dos resultados esperados mais evidente é a inferência feita com o suposto de erros skew-t é a de melhor performance, o que se confirma na Figura 6.2, onde para todos os casos de comprimento da série, os estados não observáveis verdadeiros estão contidos nos intervalos de confiança dos estados estimados. Além disso, conforme aumenta o comprimento da série temporal, a qualidade das estimativas dos estados vai melhorando.

Para os outros cenários de modelos observacionais considerados, a qualidade das estimativas é pior, o que é esperado, considerando que os dados foram simulados a partir de erros skew-t e a estimação é feita considerando erros normais, t-student e skew-normal (Figuras 6.3, 6.4 e 6.5). Um ponto interessante de destacar, é que para estes cenários a qualidade das estimativas é evidentemente pior, na medida em que a série temporal é menor.

Em relação aos parâmetros do modelo, no cenário de erros skew-t, a precisão e a exatidão das estimativas é relativamente boa para os parâmetros K, r, τ, q . No entanto que a qualidade das estimativas dos parâmetros de escala, forma e graus de liberdade (σ, λ, ν) é pior. Um resultado relevante de destacar é que a qualidade das estimativas dos parâmetros não melhora conforme aumenta o tamanho das séries temporais. Apesar de parecer este um resultado não esperado, existem 2 argumentos que permitem explicar este resultado:

- O esquema de simulação considerou a geração de **uma** série temporal de estados não observáveis, com erro de processo e 100 séries de variáveis observáveis. Considerando este esquema de simulação, parece difícil recuperar corretamente os valores dos parâmetros do modelo logístico.
- Deve-se lembrar que na análise de séries temporais, o comprimento da série não é equivalente ao tamanho da amostra, o qual de fato é de tamanho 1. Portanto, não é correto esperar consistência nas estimativas dos parâmetros.

A respeito das estimativas dos parâmetros do modelo logístico com os outros cenários de erros de observação, os resultados apresentam um comportamento similar ao descrito anteriormente, com bons níveis de precisão e exatidão para os parâmetros K, r e q . Um resultado interessante de destacar é a qualidade crescente na estimativa do desvio padrão do erro de processo (τ). Em particular para os casos de tamanho de comprimento da série maior a 50. Este resultado é esperado, pois quanto

o maior comprimento da série, maior informação se tem da variabilidade do processo não observável.

Um estudo de simulação complementar for feito considerando o procedimento de inferência aproximado através de Laplace. Porém este procedimento foi somente utilizado para o caso de maior interesse deste trabalho de tese, que o cenário de erros de observação skew-t. Os resultados obtidos a respeito das estimativas dos estados não observáveis são resumidos na Figura 6.6 e na Tabela 6.6. Nota-se um comportamento similar ao obtido via MCMC, com a diferença relevante de apresentar intervalos de confiança mais estreitos para todos os casos de comprimento da série temporal.

Figura 6.2: *Estimação de estados não observáveis. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das medias a posteriori e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo normal e erro de observação skew-t. Series temporais $n=10, 20, 30, 50, 70$ e 90 .*

Figura 6.3: *Estimação de estados não observáveis, caso 2. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das médias a posteriori e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo normal e erro de observação skew-normal. Series temporais $n=10, 20, 30, 50, 70$ e 90 .*

Figura 6.4: *Estimação de estados não observáveis, caso 2. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das medias a posteriori e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo normal e erro de observação t-student. Series temporais $n=10, 20, 30, 50, 70$ e 90 .*

Figura 6.5: *Estimação de estados não observáveis, caso 2. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das médias a posteriori e linha sólida valores verdadeiros. Procedimento de estimação via MCMC, sob hipótese de erro de processo e observação normais. Series temporais $n=10, 20, 30, 50, 70$ e 90 .*

Tabela 6.2: *Viés médio (MB) e a raiz do erro quadrático médio (RMSE) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação skew- t .*

n		K	r	τ	σ	λ	ν	q
	Valores verdadeiros	40	0.4	3	12	3	7	10
10	Media	40.87	0.39	2.16	17.34	-0.01	9.32	9.85
	MB	0.87	-0.01	-0.84	5.34	-3.01	2.32	-0.15
	RMSE	1.32	0.02	0.85	5.73	3.02	2.34	0.28
20	Media	41.62	0.41	2.17	21.74	0.06	9.20	9.95
	MB	1.62	0.01	-0.83	9.74	-2.94	2.20	-0.05
	RMSE	2.01	0.01	0.84	9.87	2.95	2.30	0.29
30	Media	41.96	0.42	2.01	21.30	0.17	9.55	9.78
	MB	1.96	0.02	-0.99	9.30	-2.83	2.55	-0.22
	RMSE	2.43	0.03	0.99	9.43	2.83	2.67	0.39
50	Media	41.19	0.38	1.93	19.71	0.19	9.62	9.70
	MB	1.19	-0.02	-1.07	7.71	-2.81	2.62	-0.30
	RMSE	2.25	0.02	1.07	7.87	2.81	2.83	0.54
70	Media	40.08	0.32	2.38	18.78	0.16	10.56	9.84
	MB	0.08	-0.08	-0.62	6.78	-2.84	3.56	-0.16
	RMSE	1.81	0.08	0.63	7.02	2.85	3.67	0.49
90	Media	40.58	0.32	2.25	17.55	0.34	10.12	9.80
	MB	0.58	-0.08	-0.75	5.55	-2.66	3.12	-0.20
	RMSE	2.18	0.08	0.76	5.79	2.67	3.43	0.57

Tabela 6.3: *Viés médio (MB) e a raiz do erro quadrático médio (RMSE) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação skew-normal.*

n		K	r	τ	q
	Valores verdadeiros	40	0.4	3	10
10	Média	39.22	0.38	0.66	9.80
	MB	-0.78	-0.02	-2.34	-0.20
	RMSE	2.32	0.02	2.36	0.55
20	Média	39.28	0.40	0.46	9.74
	MB	-0.72	0.00	-2.54	-0.26
	RMSE	2.41	0.01	2.55	0.63
30	Média	39.76	0.39	0.37	9.55
	MB	-0.24	-0.01	-2.63	-0.45
	RMSE	1.80	0.02	2.64	0.64
50	Média	39.18	0.35	1.21	10.26
	MB	-0.82	-0.05	-1.79	0.26
	RMSE	2.12	0.06	1.82	0.59
70	Média	38.5	0.31	2.23	9.63
	MB	-1.5	-0.09	-0.77	-0.37
	RMSE	2.24	0.09	0.83	0.55
90	Média	39.02	0.31	2.09	9.62
	MB	-0.98	-0.09	-0.91	-0.38
	RMSE	1.92	0.1	0.95	0.57

Tabela 6.4: *Viés médio (MB) e a raiz do erro quadrático médio (RMSE) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação t-student.*

n		K	r	τ	q
	Valores verdadeiros	40	0.4	3	10
10	Média	39.87	0.39	0.57	10.12
	MB	-0.13	-0.01	-2.43	0.12
	RMSE	2.19	0.02	2.45	0.58
20	Média	40.50	0.40	0.57	10.22
	MB	0.50	0.00	-2.43	0.22
	RMSE	2.42	0.01	2.45	0.65
30	Média	40.86	0.40	0.49	10.04
	MB	0.86	0.00	-2.51	0.04
	RMSE	2.21	0.01	2.52	0.50
50	Média	39.64	0.35	1.10	10.10
	MB	-0.36	-0.05	-1.90	0.10
	RMSE	2.24	0.06	1.93	0.61
70	Média	39.53	0.34	2.5	9.94
	MB	-0.47	-0.06	-0.5	-0.06
	RMSE	2	0.06	0.6	0.52
90	Média	40.24	0.33	2.28	9.84
	MB	0.24	-0.07	-0.72	-0.16
	RMSE	1.63	0.07	0.76	0.45

Tabela 6.5: Viés médio (MB) e a raiz do erro quadrático médio (RMSE) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação normais.

n		K	r	τ	q
	Valores verdadeiros	40	0.4	3	10
10	Média	40.18	0.38	0.56	10.02
	MB	0.18	-0.02	-2.44	0.02
	RMSE	2.00	0.02	2.46	0.53
20	Média	41.43	0.4	0.58	10.03
	MB	1.43	0.01	-2.42	0.03
	RMSE	2.69	0.01	2.43	0.56
30	Média	41.04	0.40	0.49	10.06
	MB	1.04	0.01	-2.51	0.06
	RMSE	2.50	0.01	2.52	0.55
50	Média	39.18	0.35	1.21	10.26
	MB	-0.82	-0.05	-1.79	0.26
	RMSE	2.12	0.06	1.82	0.59
70	Média	38.99	0.33	2.41	10.11
	MB	-1.01	-0.07	-0.59	0.11
	RMSE	2.12	0.07	0.65	0.53
90	Média	39.44	0.33	2.27	10.08
	MB	-0.56	-0.07	-0.73	0.08
	RMSE	2.03	0.07	0.79	0.53

Tabela 6.6: Viés médio (MB) e a raiz do erro quadrático médio (RMSE) das estimativas dos parâmetros do modelo com erro de processo normal e erro de observação skew- t . Procedimento de estimação via aproximação de Laplace.

n		K	r	τ	σ	λ	ν	q
	Valores verdadeiros	40	0.4	3	12	3	7	10
10	Média	40.96	0.42	2.25	17.43	0.08	9.37	9.90
	MB	0.88	0.01	-0.77	5.42	-2.95	2.42	-0.08
	RMSE	1.38	0.03	0.92	5.77	3.06	2.58	0.30
20	Média	41.68	0.42	2.18	21.82	0.09	9.28	10.01
	MB	1.67	0.07	-0.78	9.83	-2.90	2.26	-0.04
	RMSE	2.08	0.08	0.89	9.91	3.02	2.31	0.32
30	Média	42.05	0.52	2.09	21.35	0.23	9.59	9.82
	MB	1.98	0.11	-0.92	9.33	-2.78	2.63	-0.14
	RMSE	2.47	0.15	1.04	9.47	2.87	2.68	0.39
50	Média	41.24	0.38	2.01	19.77	0.25	9.63	9.79
	MB	1.24	0.05	-1.00	7.74	-2.72	2.66	-0.27
	RMSE	2.27	0.07	1.14	7.96	2.89	2.88	0.62
70	Média	40.08	0.42	2.45	18.83	0.26	10.58	9.92
	MB	0.14	0.01	-0.57	6.82	-2.83	3.65	-0.13
	RMSE	1.83	0.14	0.70	7.03	2.88	3.76	0.52
90	Média	40.61	0.33	2.29	17.58	0.38	10.13	9.86
	MB	0.63	-0.03	-0.68	5.61	-2.59	3.18	-0.12
	RMSE	2.20	0.08	0.82	5.89	2.69	3.45	0.60

Figura 6.6: *Estimação de estados não observáveis. Em linha pontilhada mediana e percentis de 2.5% e 97.5% das médias a posteriori e linha sólida valores verdadeiros. Procedimento de estimação via aproximação de Laplace, sob hipótese de erro de processo normal e erro de observação skew-t. Series temporais $n=10, 20, 30, 50, 70$ e 90 .*

Capítulo 7

Aplicações a dados reais

Neste capítulo apresentados dois conjuntos de dados de populações de recursos marinhos do Chile nas quais utilizamos os modelos apresentados nas seções anteriores. Os dois recursos pertencem à pesca de arrasto de crustáceos da zona central do Chile. O primeiro conjunto de dados é da pesca do camarão e o segundo da pesca do lagostim vermelho. Nos dois casos os objetivos são: estimar a biomassa anual dos recursos, estimar os parâmetros do modelo logístico que determinam sua produtividade e as quantidades de interesse para administração das pescarias.

7.1 Biomassa do camarão da zona central do Chile

O camarão (*Heterocarpus reedi*) é um crustáceo que representa um importante recurso da pesca de arrasto na zona central do Chile. Os dados usados nesta aplicação correspondem a 2 séries temporais (Figura 7.1):

- a série de capturas (toneladas) anuais de camarão na região centro-sul do Chile, entre os anos 1981 a 2013 ($n=32$);
- a série temporal dos rendimentos da pesca (toneladas/hora de arrastre), no mesmo período e zona.

A série de capturas (H_t) é considerada um vetor de constantes conhecidas; enquanto que, a série de rendimentos da pesca indexada no tempo em escala anual (Y_t), é considerada a variável aleatória observável que pode ser relacionada com os estados não observáveis (as biomassas anuais). Para este conjunto de dados utilizamos o modelo dinâmico logístico apresentado na seção 3.4:

$$X_{t+1} = X_t + rX_t \left(1 - \frac{X_t}{K}\right) - H_t + u_t, \quad X_t \geq 0, \quad r > 0, \quad K > 0, \quad H_t \geq 0, \quad t = 0, 1, \dots, \quad (7.1)$$

onde X_t é a biomassa da população no tempo t , r é a taxa intrínseca do crescimento da população, K é a capacidade de carga, que representa a abundância máxima hipotética que o ambiente pode sustentar (Krebs, 1985), H_t é a captura (ou *harvesting*) do ano t e u_t são variáveis aleatórias independentes e identicamente distribuídas. Nesta aplicação, consideramos $u_t \sim N(0, \tau^2)$. A equação de observação para os erros aditivos (Normal, t-student, skew-normal e skew-t) é:

$$Y_t = qX_t + v_t, \quad Y_t \geq 0, \quad q > 0, \quad t = 1, \dots, n. \quad (7.2)$$

para os erros de observação multiplicativos (log-normal, log-t, log-skew-normal e log-skew-t) é:

$$Y_t = qX_t e^{v_t} \quad (7.3)$$

onde Y_t é a variável observável (rendimento da pesca), q é um coeficiente de proporcionalidade e v_t são variáveis aleatórias independentes e identicamente distribuídas.

Figura 7.1: Rendimento de pesca (Y_t , toneladas/hora, linha preta sólida) e captura (H_t , milhares de toneladas, linha segmentada azul) anuais de camarão para a zona centro-sul do Chile entre os anos 1981 e 2012.

Foram escolhidas distribuições *a priori* para os parâmetros a partir das informações existentes sobre a espécie e de acordo com as recomendações de McAllister et al. (1994), Punt e Hilborn (1997) e Meyer e Millar (1999). Para os parâmetros da função logística propomos distribuições *a priori* informativas e para os outros parâmetros usamos distribuições próprias não informativas ou fracamente informativas. Dado que a capacidade de carga (K) e a taxa intrínseca de crescimento (r) são parâmetros estritamente positivos, distribuições *a priori* log-normais foram escolhidas para ambos os parâmetros. A distribuição *a priori* considerada para (K) foi uma log-normal com média igual ao valor obtido a partir de Montenegro et al. (2009), ou seja, $\mathbb{E}(K) = 33$ (milhares de toneladas), e um $CV(K) = 0,25$. Então $K \sim \text{Log-normal}(3.43, 0.25)$. A distribuição *a priori* para a taxa de crescimento (r) também foi escolhida log-normal com média igual ao valor obtido a partir de Montenegro et al. (2009), ou seja, $\mathbb{E}(r) = 0,33$ e $CV(r) = 0,25$. Portanto $r \sim \text{Log-normal}(-1.17, 0.25)$. As distribuições *a priori* para a variância do ruído de processo (τ^2), o coeficiente de capturabilidade (q) e o parâmetro de escala (σ^2) foram escolhidas como sendo não informativas, com distribuição *a priori* Gamma inversa com parâmetros (0.01, 0.01). Para os graus de liberdade da t-student e log-skew-t, foi considerada uma distribuição *a priori* hierárquica de tal forma que $\nu \mid \xi \sim \exp(\xi)$ e $\xi \sim U(0.02, 0.5)$. A distribuição *a priori* utilizada para o parâmetro de assimetria dos modelos skew-normal e log-skew-normal é $\lambda \sim t(0, \frac{\pi^2}{4}, 0, 5)$. Esta distribuição é uma aproximação para a *priori* de Jeffrey e foi sugerida em Bayes e Branco (2007).

A Tabela 7.1 mostra que dos quatro modelos aditivos propostos o que apresenta o menor DIC é o modelo com o erro de observação skew-t. O segundo melhor modelo é o modelo de erro de observação skew-normal. Ao usar erros de observação multiplicativos novamente os melhores modelos são aqueles com o parâmetro de assimetria: log-skew-t e log-skew-normal. Além disso, tendo em conta a distribuição *a posteriori* preditiva da autocorrelação dos resíduos, no caso dos modelos aditivos o skew-t e skew-normal são também os melhores para o conjunto de dados analisados. No caso de erros multiplicativos, os quatro modelos tem p-valor preditivo perto de 0.5, indicando a boa capacidade preditiva de todos. Esta evidência pode também ser observada nos gráficos de dispersão de autocorrelação dos resíduos observada versus a autocorrelação preditiva (Figuras 7.2 e 7.3). Uma comparação entre os dados observados e a distribuição preditiva *a posteriori* dos dados foi feita através da sobreposição da mediana e o HPD 95 % da distribuição preditiva *a posteriori*

Tabela 7.1: *DIC (Deviance information criterion), número efetivo de parâmetros (p_D) e posterior predictive p-value da autocorrelação dos resíduos ($\hat{\rho}$) dos modelos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile.*

Aditivos	DIC	p_D	$\hat{\rho}$
Normal	319.8	10.6	0.198
t-student	322.9	13.0	0.192
Skew-normal	297.2	29.6	0.416
Skew-t	266.3	4.6	0.510
Multiplicativos	DIC	p_D	$\hat{\rho}$
Log-normal	-40.13	26.5	0.487
Log-t	-46.62	25.8	0.542
Log-skew-normal	-54.46	16.5	0.482
Log-skew-t	-54.71	13.8	0.512

dos modelos propostos e os dados observados. Os modelos com erros normais e t-student tem o pior desempenho com os níveis mais elevados de incerteza e os modelos skew-normal e skew-t mostram baixos níveis de incerteza com intervalos de credibilidade mais estreitos (Figura 7.4). No caso de erros de observação multiplicativos, os quatro modelos reproduzem quase perfeitamente os dados observados, mas os modelos log-skew-normal e log-skew-t tiveram melhor desempenho preditivo do que os modelos de erros log-normal e log-t, com intervalos de credibilidade mais estreitos (Figura 7.5).

As Tabelas 7.2 e 7.3 mostram as medidas resumos das distribuições *a posteriori* dos parâmetros dos oito modelos propostos. Nota-se que, as diferenças nas estimativas dos parâmetros da função logística entre os modelos não são muito grandes. As médias *a posteriori* para a capacidade de suporte (K) situam-se entre 42,6 e 46,9 milhares de toneladas para os modelos aditivos e entre 43,8 e 44,8 milhares de toneladas para os modelos multiplicativos. As médias *a posteriori* para a taxa intrínseca de crescimento populacional (r) estão entre 0,42 e 0,46 para os modelos aditivo e entre 0,42 e 0,43 para os modelos multiplicativos. O desvio padrão do ruído do processo (τ) foi estimado quase identicamente entre os oito modelos propostos, com médias *a posteriori* entre 2,57 e 3,06. Os modelos normal, t-student, skew-normal e skew-t foram considerados homocedásticos e os modelos multiplicativos são intrinsecamente heterocedásticos (a variância depende dos parâmetros de localização e da escala). No caso do modelo com erro de observação t-student, os graus de liberdade foram estimados com um intervalo de credibilidade amplo, não mostrando evidência explícita de caudas pesadas nos dados. Sob a hipótese de distribuição skew-normal, a estimativa do parâmetro de assimetria mostra evidências de assimetria positiva nos dados.

Figura 7.2: Diagramas de dispersão das discrepâncias observadas e previstas da autocorrelação dos error de observação dos modelos aditivos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile.

Figura 7.3: Diagramas de dispersão das discrepâncias observadas e previstas da autocorrelação dos erros de observação dos modelos multiplicativos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile.

Figura 7.4: Distribuição a posteriori preditiva do rendimento da pesca para os modelos aditivos propostos (linhas) e rendimentos da pesca observados (pontos) de camarão no Chile.

Figura 7.5: Distribuição a posteriori preditiva do rendimento da pesca para os modelos multiplicativos propostos (linhas) e rendimentos da pesca observados (pontos) de camarão no Chile.

Tabela 7.2: Estatísticas da distribuição a posteriori dos parâmetros dos modelos aditivos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile. Se apresenta a medida de redução de escala potencial ($\hat{R}$).

Parâmetro	Média	Desv. Pad.	Mediana	Perc. 2.5%	Perc. 97.5%	$\hat{R}$
Normal						
K	46.58	8.19	45.79	32.61	65.14	1.004
r	0.42	0.07	0.42	0.28	0.58	1.002
$q \times 10^6$	9.11	1.77	9.00	5.93	13.04	1.005
τ	2.57	0.93	2.40	1.27	4.85	1.001
σ	37.57	5.63	36.97	28.33	50.35	1.001
t-student						
K	46.93	8.31	46.04	33.43	65.40	1.002
r	0.42	0.07	0.41	0.28	0.57	1.001
$q \times 10^6$	8.99	1.72	8.89	5.98	12.53	1.003
τ	2.62	0.93	2.45	1.30	4.90	1.001
σ	37.11	5.63	36.49	27.86	49.91	1.001
ν	20.14	22.68	12.66	3.20	83.47	1.001
Skew-normal						
K	43.53	7.15	42.74	31.67	60.03	1.052
r	0.46	0.08	0.46	0.30	0.63	1.020
$q \times 10^6$	9.12	1.56	9.08	6.14	12.70	1.078
τ	2.78	0.87	2.66	1.47	4.87	1.020
σ	32.17	7.01	31.50	20.47	47.75	1.001
λ	3.98	1.29	3.95	1.56	6.60	1.002
Skew-t						
K	42.58	7.72	41.61	30.30	60.30	1.166
r	0.44	0.08	0.43	0.28	0.61	1.072
$q \times 10^6$	10.22	1.98	10.09	6.41	14.08	1.266
τ	3.06	0.95	2.91	1.67	5.40	1.116
σ	11.53	4.75	10.53	4.01	22.99	1.005
λ	0.68	3.81	0.44	-3.36	5.73	1.102
ν	7.48	9.75	4.33	1.43	34.70	1.015

As estimativas dos parâmetros do modelo de crescimento logístico têm uma relevância imediata para gestão pesqueira. O máximo rendimento sustentável ($MSY = rK/4$), o nível de biomassa do máximo rendimento sustentável ($B_{MSY} = K/2$) e a taxa de pesca que maximiza o excedente de produção ($H_{MSY} = r/2$) são funções destes parâmetros. Nas Tabelas 7.4 e 7.5 a estimativa destas quantidades são apresentadas. Os níveis mais baixos de B_{MSY} são os obtidos com a distribuição skew-t e log-skew-t com médias *a posteriori* variando entre 21,3 (15,5-29,9 95% HPD) e 21,9 (16-30 95% HPD) milhares de toneladas. As estimativas de rendimento máximo sustentável (MSY) são semelhantes entre todos os modelos, variando entre 4,6 e 4,9 milhares de toneladas. Estes valores são menores do que os estimados de 5,5 milhares de toneladas por Montenegro et al. (2009). Nossos resultados mostraram quase nenhuma diferença na distribuição *a posteriori* da taxa de pesca do rendimento máximo sustentável (H_{MSY}) entre os oito modelos propostos, com média *a posteriori* entre 0,21 e 0,22. Considerando o procedimento sequencial de inferência, é simples estimar a densidade um passo à frente dos estados (X_{n+1}). A previsão de biomassa de camarão para o ano 2014 teve uma média *a posteriori* entre 34,8 milhares de toneladas (modelo skew-t) e 37,9 milhares de toneladas (modelo de t-student), com HPD de 95% entre 24,5 e 50,2 e entre 26,4 e 55,7 milhares de toneladas, respectivamente.

Tabela 7.3: Estatísticas da distribuição a posteriori dos parâmetros dos modelos multiplicativos utilizados para estimar os parâmetros do modelo de crescimento logístico da população de camarão no Chile. Se apresenta a medida de redução de escala potencial ($\hat{R}$).

Parâmetro	Média	Desv. Pad.	Mediana	Perc. 2.5%	Perc. 97.5%	$\hat{R}$
Log-normal						
K	44.73	7.95	43.69	32.16	63.33	1.015
r	0.42	0.07	0.42	0.28	0.58	1.007
$q \times 10^6$	9.58	1.81	9.53	6.22	12.89	1.025
τ	2.93	0.92	2.76	1.63	5.13	1.010
σ	0.09	0.03	0.08	0.03	0.15	1.001
Log-t						
K	44.47	8.07	43.35	31.55	62.14	1.049
r	0.43	0.08	0.43	0.28	0.60	1.025
$q \times 10^6$	9.70	1.91	9.67	6.42	13.15	1.066
τ	2.78	0.89	2.62	1.52	4.89	1.026
σ	0.07	0.02	0.07	0.02	0.13	1.001
ν	9.34	14.15	4.99	1.34	47.40	1.001
Log-skew-normal						
K	44.33	7.59	43.45	32.01	60.95	1.008
r	0.42	0.07	0.42	0.28	0.59	1.003
$q \times 10^6$	9.63	1.77	9.65	6.47	13.50	1.013
τ	3.03	0.93	2.89	1.64	5.16	1.003
σ	0.66	0.59	0.48	0.05	2.19	1.015
λ	-0.11	10.39	0.11	-20.20	18.63	1.001
Log-skew-t						
K	43.77	7.19	43.01	32.22	59.60	1.084
r	0.43	0.07	0.43	0.29	0.59	1.044
$q \times 10^6$	9.80	1.68	9.76	6.56	13.23	1.132
τ	2.83	0.89	2.68	1.54	4.96	1.066
σ	0.52	0.48	0.37	0.04	1.78	1.012
λ	0.27	10.22	0.18	-20.39	20.27	1.001
ν	8.23	11.92	4.48	1.23	39.75	1.001

Tabela 7.4: Estatísticas da distribuição a posteriori das quantidades de gestão derivadas do modelo logístico com erros de observação aditivos, população de camarão no Chile

Parâmetro	Média	Desv. Pad.	Mediana	Perc. 2.5%	Perc. 97.5%
Normal					
B_{MSY}	23.29	4.09	22.90	16.45	32.63
B_{2014}	37.90	7.46	37.06	25.59	55.38
H_{MSY}	0.21	0.03	0.21	0.14	0.29
MSY	4.84	0.68	4.84	3.47	6.22
t-student					
B_{MSY}	23.27	4.11	22.78	16.78	32.24
B_{2014}	37.90	7.55	36.78	26.40	55.70
H_{MSY}	0.21	0.03	0.21	0.14	0.28
MSY	4.83	0.69	4.83	3.43	6.24
Skew-normal					
B_{MSY}	22.75	4.06	22.34	16.35	31.56
B_{2014}	37.80	7.42	36.95	26.31	53.10
H_{MSY}	0.22	0.04	0.22	0.14	0.31
MSY	4.95	0.75	4.96	3.40	6.39
Skew-t					
B_{MSY}	21.29	3.86	20.89	15.49	29.85
B_{2014}	34.80	7.05	33.80	24.54	50.26
H_{MSY}	0.22	0.04	0.21	0.14	0.30
MSY	4.60	0.71	4.59	3.17	6.04

A distribuição *a posteriori* dos estados não observáveis (biomassa) para todos os modelos propostos mostraram uma tendência semelhante, com níveis mais baixos de biomassa entre os anos 1999 e 2001 e os valores máximos entre 2008 e 2009. Os intervalos de credibilidade mais estreitos foram obtidos com a distribuição skew-normal e skew-t no caso aditivo e com o log-skew-normal e log-skew-t no cenário de erros multiplicativos. Os modelos aditivos normais e t-student mostram uma tendência mais suavizada na biomassa estimada que os outros modelos, com os intervalos de credibilidades mais amplas. A comparação da biomassa no rendimento máximo sustentável B_{MSY} com as estimativa dos níveis de biomassa com todos os modelos propostos revelou três períodos: a primeira década em que os níveis de biomassa foram menores do que os B_{MSY} , uma segunda onde os níveis de biomassa mostraram uma tendência crescente chegando a níveis próximos aos da B_{MSY} , e um terceiro e último período, onde os níveis de biomassa são mais elevados do que os da B_{MSY} (Figuras 7.6 e 7.7).

Figura 7.6: Estimativa de estados não observáveis (HPD de 95% da biomassa de camarão no Chile) com os modelos com erro de observação aditivos. A linha horizontal mostra a média a posteriori de B_{MSY}

Figura 7.7: Estimativa de estados não observáveis (HPD de 95% da biomassa de camarão no Chile) com os modelos com erro de observação multiplicativos. A linha horizontal mostra a média a posteriori de B_{MSY} .

Tabela 7.5: Estatísticas da distribuição a posteriori das quantidades de gestão derivadas do modelo logístico com erros de observação multiplicativos, população de camarão no Chile

Parâmetro	Média	Desv. Pad.	Mediana	Perc. 2.5%	Perc. 97.5%
Log-normal					
B_{MSY}	22.37	3.97	21.88	16.26	31.60
B_{2014}	36.46	7.32	35.39	25.37	54.69
H_{MSY}	0.21	0.03	0.21	0.14	0.29
MSY	4.70	0.69	4.69	3.35	6.13
Log-t					
B_{MSY}	22.24	4.03	21.71	15.84	31.14
B_{2014}	36.20	7.31	35.19	25.25	51.71
H_{MSY}	0.21	0.04	0.21	0.14	0.30
MSY	4.75	0.67	4.74	3.43	6.11
Log-skew-normal					
B_{MSY}	22.17	3.79	21.73	16.04	30.63
B_{2014}	36.18	6.81	35.15	25.36	51.47
H_{MSY}	0.21	0.03	0.21	0.14	0.29
MSY	4.66	0.71	4.66	3.25	6.10
Log-skew-t					
B_{MSY}	21.89	3.59	21.47	16.04	29.99
B_{2014}	35.62	6.35	34.72	25.48	50.83
H_{MSY}	0.21	0.03	0.21	0.14	0.30
MSY	4.71	0.67	4.71	3.34	6.05

As taxas de pesca (a biomassa capturadas durante o tempo t dividida pela biomassa disponível em tempo t , (H_t/X_t)), mostra a pressão da pesca sobre a biomassa disponível de um recurso pesqueiro e pode ser comparada com a taxa de pesca no rendimento máximo sustentável (H_{MSY}) para realizar uma análise retrospectiva dos níveis de exploração. No caso do camarão chileno, todos os modelos analisados apresentaram níveis de taxa de pesca superior à taxa de pesca no rendimento máximo sustentável durante o período 1990-1999 e uma clara diminuição das taxas de pesca no ano de 2000 abaixo do H_{MSY} e se mantém constante até ao fim da série de tempo analisada (Figuras 7.8 e B.4). Os modelos com erro de observação aditivo mostram maior grau de incerteza na estimativa das taxas de colheita do que os modelos com termos de erro multiplicativos.

Os resultados obtidos para os dados do camarão da zona central do Chile, mostram que os modelos com erros de observação aditivos skew-normal e skew-t são mais adequados que os modelos normal e t-student. Por sua vez, os modelos com erros de observação multiplicativos log-skew-normal e log-skew-t apresentam melhor performance que as alternativas log-normal e log-t. Além disso, as estimativas de biomassa do camarão apresentam menores níveis de incerteza com os dois modelos assimétricos utilizados, sob ambas estruturas de erros de observação. No entanto, as diferenças nas estimativas dos parâmetros do modelo logístico têm relativamente baixas diferenças entre todos os modelos utilizados.

Figura 7.8: Estimativa das taxas de pesca (HPD de 95%), com modelos com erros de observação aditivos. A linha horizontal mostra a média a posteriori de H_{MSY} , pescaria do camarão no Chile.

Figura 7.9: Estimativa das taxas de pesca (HPD de 95%), com modelos com erros de observação multiplicativos. A linha horizontal mostra a média a posteriori de H_{MSY} , pescaria do camarão no Chile.

7.2 Biomassa do lagostim vermelho no Chile

O lagostim vermelho (*Pleuroncodes monodom*) igual que o camarão é um crustáceo que representa um importante recurso da pescaria de arrasto da zona central do Chile. Os dados usados nesta aplicação correspondem a 3 séries temporais (Figura 7.10):

- a série de capturas (toneladas) anuais de lagostim vermelho na região centro-sul do Chile, entre os anos 1968 a 2014 (47 anos),
- a série temporal dos rendimentos da pesca (toneladas/hora de arrastre), no mesmo período e zona (Y_{1t}),
- a série temporal das estimativas de biomassa vulnerável obtidas em pesquisas de arrasto de fundo (toneladas), no mesmo período e zona (Y_{2t}).

A série de capturas (H_t) é considerada um vetor de constantes conhecidas, enquanto que a série de rendimentos da pesca indexada no tempo em escala anual (Y_{1t}), é considerada uma variável aleatória observável que se relaciona com os estados não observáveis (as biomassas anuais) sendo que a série de estimativas de biomassa vulnerável obtidas em pesquisas de arrasto é considerada uma segunda variável observável (Y_{2t}). Ver Seção 3.4, equação 3.18, com $m = 2$.

Para os parâmetros da função logística propomos distribuições *a priori* informativas e para os outros parâmetros usamos distribuições próprias não informativas ou fracamente informativas. Dado que a capacidade de carga (K) e taxa intrínseca de crescimento (r) são parâmetros estritamente positivos, distribuições *a priori* log-normais foram escolhidas para ambos os parâmetros. A distribuição *a priori* para (K) foi escolhida para ser log-normal com média de 100 milhares de toneladas, e um $CV(K) = 0,25$. Este valor está baseado nos valores máximos de biomassa estimados por Bucarey et al (2014) para o período 1980 até 2014. Portanto $K \sim \text{Log-normal}(4.54, 0.25)$. A distribuição *a priori* para (r) também foi escolhida para ser log-normal com média igual ao valor obtido a partir de Montenegro (2010) para o lagostim amarelo, espécie da mesma família que o lagostim vermelho (Galatheidæ), ou seja, $E(r) = 0,33$ e $CV(r) = 0,25$. Portanto $r \sim \text{Log-normal}(-1.17, 0.25)$. As distribuições *a priori* para a variância do ruído de processo (τ^2), o coeficiente de capturabilidade (q) e o parâmetro de escala (σ^2) foram escolhidas como uma distribuição *a priori* não informativas, com uma distribuição *a priori* Gamma inversa com parâmetros (0.01, 0.01). Para os graus de liberdade da t-student e log-skew-t, foi considerada uma distribuição *a priori* hierárquica de tal forma que $\nu \mid \xi \sim \exp(\xi)$ e $\xi \sim U(0.02, 0.5)$. A distribuição *a priori* utilizada para o parâmetro de assimetria dos modelos skew-normal e log-skew-normal é $\lambda \sim t(0, \frac{\pi^2}{4}, 0, 5)$. Esta distribuição é uma aproximação para a *priori* de Jeffrey e foi sugerida em Bayes e Branco (2007).

A Tabela 7.6 apresenta a estatística DIC, o número efetivo de parâmetros e o p-valor *a posteriori* predictivo da autocorrelação dos erros de observação, de ambos índices de abundância (Y_{1t} , Y_{2t}). Utilizando o critério DIC os melhores modelos são o log-normal e log-t, no entanto ao utilizar o p-valor *a posteriori* predictivo mostra o modelo log-skew-normal como o mais adequado para o primeiro índice de abundância relativa, no entanto para que o segundo índice de abundância relativa o melhor modelo é o log-skew-t. Desta forma, uma estrutura de modelagem, considerando distribuição log-skew-normal para o primeiro índice e log-skew-t para o segundo, parece uma alternativa interessante de ser explorada. Na Figura 7.11 apresentamos os percentis da distribuição *a posteriori* predictiva dos rendimentos de pesca de lagostim colorado (Y_{1t}). Pode se perceber que nos modelos log-normal e log-t os intervalos de credibilidade são mais amplos que nos modelos log-skew-normal e log-skew-t mostrando maior precisão destes últimos.

Figura 7.10: *Série temporal de médio rendimento anual de pesca (acima, linha sólida, Y_{1t}) e biomassa vulnerável estimada em pesquisas de arrasto de fundo (acima, linha pontilhada, Y_{2t}) e capturas total anual do lagostim vermelho na zona centro-sul do Chile (abaixo, H_t).*

Tabela 7.6: DIC (*Deviance information criterion*), número efetivo de parâmetros (p_D) e posterior predictive p -values da autocorrelação dos erros de observação das variáveis Y_{1t} ($\hat{p}_1$) e Y_{2t} ($\hat{p}_2$) dos modelos propostos para os dados do lagostim vermelho.

	DIC	p_D	$\hat{p}_1$	$\hat{p}_2$
Log-normal	18.66	35.42	0.361	0.063
Log-t	11.72	31.99	0.355	0.057
Log-skew-normal	43.72	18.87	0.505	0.149
Log-skew-t	30.28	12.01	0.654	0.254

Como era de se esperar, os intervalos de credibilidade para os anos com dados faltantes, são mais amplos que nos anos em que se conta com informação. É importante sinalizar que, para os quatro modelos, os dados observados se encontram dentro dos intervalos de credibilidade (95%) da distribuição preditiva.

Figura 7.11: Percentis da distribuição a posteriori preditiva do rendimento da pesca (Y_{1t}) de lagostim vermelho para os modelos propostos (linhas) e rendimentos da pesca observados (pontos).

Na Figura 7.12 apresentamos os percentis da distribuição a posteriori preditiva do índice de biomassa vulnerável de lagostim vermelho para os modelos propostos, junto aos dados observados. Se apresenta um padrão similar ao obtido para o índice 1 (rendimentos de pesca) com todas as observações dentro dos intervalos de credibilidade com amplos intervalos de confiança para as seções da série temporal com dados faltantes e com os modelos log-skew-normal e log-skew-t como os mais

apropriados para o conjunto de dados analisado com intervalos de credibilidade de alta precisão.

Figura 7.12: Percentis da distribuição *a posteriori* preditiva do índice de biomassa vulnerável (Y_{2t}) de lagostim vermelho para os modelos propostos (linhas) e biomassa vulnerável observada (pontos).

A Tabela 7.7 apresenta as estatísticas da distribuição *a posteriori* dos parâmetros do modelo de crescimento logístico para lagostim vermelho para os quatro modelos de erro de observação multiplicativos considerados. Os quatro modelos estimam os parâmetros da função logística de evolução (K e r) com um valor médio muito similar e também com intervalos de credibilidade similares. Quanto do desvio padrão do erro de processo (τ) os quatro modelos apresentam estimações similares quanto à estimativa pontual e os intervalos de credibilidade. Quanto aos parâmetros associados aos erros de observação dos modelos, no caso da distribuição log-normal se apresenta uma maior variabilidade do índice 2 do que no índice 1 e quanto aos erros log-t ambos índices apresentam moderadas evidências de caudas pesadas. Quanto ao parâmetro de assimetria estimado para ambos índices com a distribuição log-skew-normal, não existe evidência suficiente nos dados para afirmar que exista assimetria positiva ou negativa. Este mesmo padrão se observa nos resultados obtidos com a distribuição log-skew-t, onde não há evidência de assimetria nos dados, mas sim existe uma moderada evidência de caudas pesadas no processo aleatório subjacente.

A Tabela 7.8 apresenta as estatísticas da distribuição *a posteriori* das quantidades de gestão derivadas do modelo de crescimento logístico para o lagostim vermelho. O MSY foi estimado em torno das 14 mil toneladas com os quatro modelos propostos, no entanto a projeção de biomassa para o ano de 2015 foi estimada uma média de no máximo de 95 mil toneladas para o caso da distribuição log-skew-t e 92 mil toneladas para o caso da log-skew-normal. Os quatro modelos estimam a taxa intrínseca de crescimento (r) em torno à 0.18. A B_{MSY} foi estimada cerca de 80 mil toneladas para o caso da distribuição log-skew-t e da log-t, no entanto que a log-normal e a log-skew-normal apresentam estimativas próximas das 77 milhares de toneladas.

Na Figura 7.13 se apresentam as estimativas dos estados não observáveis (HPD de 95% da biomassa de lagostim vermelho) mostrando-se tendências similares nos quatro modelos propostos e com altos níveis de incerteza para os anos em que as séries temporais apresentam dados faltantes.

Figura 7.13: Estimativa de estados não observáveis (HPD de 95% da biomassa de lagostim vermelho). A linha horizontal mostra a média *a posteriori* de B_{MSY} .

Tabela 7.7: Estatísticas da distribuição a posteriori dos parâmetros do modelo de crescimento logístico para lagostim vermelho do Chile.

Parâmetro	Média	Desvio padrão	Mediana	Perc. 2.5%	Perc. 97.5%
Log-normal					
K	155.60	28.22	153.60	107.60	220.80
r	0.35	0.16	0.33	0.12	0.71
q_1	7.91	1.77	7.72	4.98	11.92
q_2	0.37	0.09	0.36	0.22	0.57
τ	37.74	9.75	36.11	23.37	62.16
σ_1	0.13	0.07	0.12	0.03	0.30
σ_2	0.41	0.09	0.40	0.25	0.60
Log-t					
K	158.20	28.59	156.30	108.70	219.7
r	0.35	0.15	0.33	0.12	0.71
q_1	7.66	1.67	7.49	4.84	11.54
q_2	0.37	0.09	0.36	0.22	0.56
τ	38.31	9.82	36.8	23.59	62.23
σ_1	0.12	0.06	0.11	0.03	0.27
σ_2	0.34	0.10	0.34	0.16	0.55
ν_1	9.70	14.27	5.23	1.31	46.46
ν_2	12.53	17.00	7.15	1.60	57.49
Log-skew-normal					
K	154.00	28.13	150.70	108.30	218.20
r	0.35	0.15	0.33	0.12	0.70
q_1	7.98	1.75	7.88	4.85	11.49
q_2	0.38	0.10	0.38	0.22	0.59
τ	38.23	10.06	36.69	24.01	62.14
σ_1	0.92	1.03	0.55	0.05	3.75
σ_2	3.16	2.16	2.80	0.41	8.18
λ_1	-1.51	10.23	-2.53	-20.27	19.66
λ_2	-2.89	9.85	-3.17	-21.24	16.38
Log-skew-t					
K	158.10	29.76	155.70	106.60	222.60
r	0.34	0.15	0.32	0.12	0.69
q_1	7.67	1.86	7.47	4.92	11.65
q_2	0.38	0.10	0.37	0.22	0.60
τ	39.48	10.87	37.69	23.78	63.26
σ_1	0.80	0.93	0.47	0.04	3.34
σ_2	2.71	1.89	2.34	0.35	7.18
λ_1	1.39	10.46	1.90	-19.47	21.18
λ_2	2.20	10.21	2.42	-16.72	21.35
ν_1	9.34	13.99	4.95	1.20	45.95
ν_2	12.97	16.23	7.54	1.71	57.43

Tabela 7.8: Estatísticas da distribuição a posteriori das quantidades de gestão derivadas do modelo de crescimento logístico para lagostim vermelho do Chile.

Parâmetro	Média	Desv. Pad.	Mediana	Perc. 2.5%	Perc. 97.5%
Log-normal					
B_{MSY}	77.78	14.11	76.62	53.24	108.90
B_{2015}	93.52	23.42	90.33	55.86	146.10
H_{MSY}	0.18	0.08	0.17	0.06	0.36
MSY	13.73	6.27	12.93	4.04	28.22
Log-t					
B_{MSY}	79.09	14.29	77.85	54.66	109.90
B_{2015}	95.47	23.99	92.36	55.45	156.20
H_{MSY}	0.17	0.08	0.16	0.06	0.35
MSY	13.79	6.37	12.99	3.99	28.47
Log-skew-normal					
B_{MSY}	76.98	14.06	75.51	53.22	108.60
B_{2015}	91.57	23.32	87.75	56.24	146.20
H_{MSY}	0.18	0.08	0.17	0.06	0.35
MSY	13.56	6.18	12.74	4.03	27.61
Log-skew-t					
B_{MSY}	79.04	14.88	77.77	52.75	109.70
B_{2015}	95.04	24.31	92.53	55.8	149.10
H_{MSY}	0.17	0.08	0.16	0.06	0.34
MSY	13.56	6.23	12.74	3.93	27.91

Capítulo 8

Discussão e Considerações Finais

Existem várias questões que precisam ser abordadas para estudar a adequação de um modelo, para um problema específico. No caso do modelo dinâmico estudado neste trabalho, devem ser especificados: (1) a equação de evolução; (2) a equação de observação; (3) a distribuição de probabilidade do ruído do processo; (4) a distribuição de probabilidade do erro de observação; as distribuições *a priori* dos parâmetros do modelo e (5) os critérios de seleção dos modelos. Em modelagem de populações marinhas, a primeira questão depende do conhecimento disponível da população e dos dados disponíveis para fazer inferências. Os modelos disponíveis vão desde o modelo de crescimento logístico utilizado nesta tese, até os modelos mais complexos que apresentam uma estrutura explícita de idades ou tamanhos (Haddon, 2011).

Neste trabalho, fizemos a análise considerando uma situação de pouca informação disponível, especificamente na situação que existem séries de dados das capturas e de algum índice auxiliar que possa se relacionar com as abundâncias (estados não observáveis). A formulação do modelo logístico utilizado neste trabalho é a mais utilizada na literatura, mas existem várias outras formulações do modelo de crescimento logístico, como as de Gompertz, Ricker e Beverton-Holt (Birch, 199; De Lara e Doyen, 2008). Outra opção é o modelo generalizado de Pella e Tomlinson (1969), que incorpora um parâmetro que permite que a curva de produção seja assimétrica. Existem pelo menos duas versões deste modelo: a formulação Fletcher (Quinn e Deriso, 1999) e a versão de Gilbert (Gilbert, 1992). Um problema reconhecido com estas versões generalizadas é que, dependendo dos dados disponíveis, pode ou não ser possível estimar o parâmetro de forma (Prager, 2002). Dada a forma relativamente simples desta classe de modelos, poderia ser razoável utilizar mais de uma formulação da função de produção e testar a performance do processo inferencial para cada situação de aplicação de dados, em particular, através das técnicas de seleção de modelos (Burnham e Anderson, 1998). Considerando que o objetivo central desta tese, foi aplicar distribuições de probabilidade flexíveis para os erros de observação, usamos a versão padrão do modelo de crescimento logístico de Scheafer (1954).

Outro aspecto importante que precisa ser abordado, é a forma da equação de observação que relaciona as variáveis aleatórias observáveis com os estados latentes. A escolha da equação depende dos dados disponíveis e da relação teórica proposta entre estes dois tipos de variáveis. No caso de dinâmica de populações marinhas é comum assumir que um índice de abundância relativa, como as taxas de capturas são diretamente proporcionais à abundância. Na prática esta suposição pode não ser correta, pois pode existir incrementos temporais na capturabilidade (q), entre outros fatores. Harley et al., (2001) apresenta um exemplo desta formulação mais geral, na qual a relação entre as variáveis observáveis e não observáveis é não linear, ou seja, $Y_t = qX_t^\beta$ onde se $\beta = 1$ o modelo se reduz à relação linear e se $\beta \neq 1$, então a capturabilidade muda com a abundância. Quando $\beta > 1$, Y_t declina mais rápido do que X_t , em uma situação conhecida como *hyperdepletion*. Inversamente, se $\beta < 1$, Y_t declina mais rápido do que X_t , resultando a situação conhecida como *hyperstability*. Inserir esta forma de modelagem em nossos modelos propostos é relativamente fácil, apenas um

parâmetro deve ser adicionado ao modelo, mas ao custo de transformação da equação de observação linear em uma não-linear. Neste trabalho, foi utilizada uma relação linear entre as variáveis observáveis e não observáveis e considerando o coeficiente de capturabilidade tempo invariante por simplicidade. No entanto, se suspeita-se que existam mudanças temporais na capturabilidade, pode usar algumas das múltiplas alternativas de modelagem apresentadas em Wilberg et al.,(2010).

Uma das características mais interessantes dos modelos espaço de estados é a possibilidade de incluir um termo de erro de processo na equação de estado. Este erro pode ter várias fontes, principalmente da variabilidade demográfica e/ou variabilidade ambiental (Dennis et al., 2006; Maunder e Deriso, 2011). Em algumas séries de tempo a amplitude das variações não muda com o nível da série. Nestes casos, um ruído processo aditivo é adequado. Mas em outras séries tempo, a amplitude das variações aumenta à medida que o nível da tendência aumenta. Nesta situação, um ruído de processo multiplicativo poderia ser mais apropriado. Nas aplicações ecológicas o ruído de processo representa o efeito conjunto de um grande número de eventos aleatórios multiplicativos que se combinam para formar um erro processo multiplicativo log-normal. Isso corresponde à versão multiplicativa do teorema do limite central, o que afirma que o produto de número grande de variáveis aleatórias positivas, independentes e identicamente distribuídas tem aproximadamente uma distribuição log-normal (Limpert et al., 2001). Conforme informação de Maunder et al. (2011), os modelos de espaço de estados tradicionalmente refletem a variabilidade demográfica. Apesar disso, geralmente a variabilidade ambiental é maior que a variabilidade demográfica.

Em modelagem de populações marinhas, a hipótese mais comum sobre os erros de observação é que eles são multiplicativos Log-normais (Millar e Meyer, 2000; de Valpine, 2002; Dennis et al 2006; Maunder e Deriso, 2011). Outros pressupostos comuns são que os erros são aditivos e normalmente distribuídos, com coeficiente de variação constante (Pollacheck et al., 1993). A presença de erros de observação assimétricos nas variáveis ecológicas motivou o uso de outras distribuições de probabilidade que pode acomodar observações atípicas e assimétricas com erros multiplicativos e aditivos.

As distribuições de probabilidade dos erros de observação no modelo dinâmico logístico apresentado neste trabalho são adequadas para este tipo de estudos. Apresentamos duas famílias de distribuições com potencial para ser usado não apenas para os erros do modelo de observação do modelo espaço de estados, mas também para ser estendido para a equação de evolução. No caso dos erros aditivos sugerimos o uso da família SMSN descrita na seção 4.2, e no caso de erros multiplicativos o uso da família log-SMSN apresentada na seção 4.3. Em modelagem de populações biológicas os erros multiplicativos parecem ser mais adequados. Dentro desta classe, a distribuição log-skew-t surge como uma alternativa interessante à distribuição log-normal. O nível adicional de complexidade das famílias de distribuição aqui apresentada não é muito alto, especialmente quando se utilizam suas representações estocásticas, tornando mais fácil para o analista compreender a extensão das distribuições mais comuns como a normal e log-normal a estas famílias mais gerais e flexíveis. Além disso, se a abordagem de inferência é Bayesiana o analista tem uma maneira simples de implementar este tipo de modelos usando o software WinBUGS ou OpenBUGS (Thomas et al., 2006).

Neste trabalho foi utilizado o *DIC* como um dos critérios para comparar modelos. Porém *DIC* tem uma série de desvantagens quando é utilizado no contexto de modelos de efeitos aleatórios e modelos espaço de estado. Isso é por causa da falta de forma fechada da função de verossimilhança dos dados observados ou no contexto de modelos de espaço de estados, para a densidade de previsão um passo à frente para as observações. Neste caso o cálculo do *DIC* desde a amostra da distribuição posterior obtida via MCMC pode ser muito demorado ou não viável já que a $L(\Theta; y_{1:n})$ não tem forma fechada. Além disso, o *DIC* é definido com base na técnica de aumento de dados, tratando os estados não observáveis como parâmetros adicionais. Como foi apontado por Li et al. (2012), a partir de um ponto de vista teórico, o *DIC* obtido a partir da verossimilhança aumentada tem

dois problemas graves. Um dos problemas é devido ao aumento de dados, o número de variáveis latentes aumenta com o tamanho amostral, resultando numa verossimilhança não regular e, portanto invalidando a justificação assintótica do DIC e o outro é devido aos dados aumentados, a dimensão do espaço de parâmetros torna-se maior e por isso, é esperado que o DIC seja sensível às transformações das variáveis latentes (Li et al., 2012). Considerando a dificuldade em calcular a *DIC* em modelos dinâmicos existem várias alternativas para solucionar este problema. Uma solução possível é a utilização de algoritmos, tais como: filtro de partículas ou sequencial de Monte Carlo (Arulampalam et al., 2002).

Outra possibilidade é usar a aproximação de Laplace para a verossimilhança marginal, ou suas variantes como os algoritmos INLA (Martino e Rue, 2008). A implementação de Laplace para a aproximação da verossimilhança marginal pode ser utilizada facilmente através da utilização da plataforma de programação ADMB, com seu módulo de efeitos aleatórios (Skaug e Fournier, 2006). No caso dos algoritmos INLA existem algumas implementações no pacote R-INLA (Rue et al., 2015). Neste trabalho, utilizamos os *posterior predictive checks*, estimando-se a autocorrelação dos resíduos como um método para a avaliação da qualidade do modelo. Os resultados mostraram que estes tipos de diagnóstico são adequados para comparar a classe de modelos espaço de estados discutidos aqui. Estes resultados foram consistentes com aqueles obtidos utilizando o *DIC*, apontando para as distribuições skew-t e log-skew-t como as mais adequadas para os erros aditivos e multiplicativos, respectivamente.

A aplicação de distribuições de caudas pesadas e assimétricas tem sido feita em diversas áreas (finanças, economia, meteorologia, etc.), mas o uso dessas famílias de distribuição na investigação pesqueira não é comum. Em particular, aplicações recentes de sucesso com este tipo de distribuições, neste campo de pesquisa, podem ser encontradas nos artigos de Contreras-Reyes e Arellano-Valle (2013) e Contreras-Reyes et al. (2014) onde a classe SMSN foi usada para modelar os erros de observação no modelo de crescimento somático de von Bertalanffy (von Bertalanffy, 1938).

As extensões deste trabalho de teses e as possibilidades de pesquisas futuras são múltiplas e entre estas, identificamos as seguintes:

- Aplicar as distribuições de probabilidade utilizadas nos erros de observação para os erros da equação de evolução.
- Considerar modelos de dinâmica populacional mais complexos que o apresentado neste trabalho de teses, como os estruturados em grupos de idade ou grupos de tamanho, como no caso de Haddon (2011).
- Estudar outras distribuições de probabilidade da família Skew-Elíptica (Marchenko e Genton, 2010).
- Utilizar outros procedimentos de inferência aproximada como os filtros de partículas e os algoritmos INLA (Arulampalam et al., 2002; Rue et al., 2015)

Referências Bibliográficas

- [1] Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V.H., Enriquez, I., 2010. Robust Bayesian Analysis of Heavy-tailed Stochastic Volatility Models using Scale Mixtures of Normal Distributions. *Computational Statistics and Data Analysis*, 54, 2883-2898.
- [2] Aitchison, J., Brown, J.A.C., 1957. *The log-normal Distribution*, Cambridge University Press.
- [3] Aitkin, M., 2010. *Statistical Inference: an Integrated Bayesian/Likelihood Approach*. Boca Raton, Chapman and Hall/CRC Press.
- [4] Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle, in Petrov, B.N.; Csáki, F., 2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akadémiai Kiadó, pp. 267-281.
- [5] Alspach, D.L., Sorensen, H.W., 1972. Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Trans. Automat. Contr.*, AC-17, 439-447.
- [6] Anderson, B.D.O., Moore, J.B., 1979. *Optimal Filtering*. Englewood Cliffs: Prentice-Hall.
- [7] Anderson-Sprecher, R., 1994. Robust estimates of wildlife location using telemetry data. *Biometrics* 5:406-416.
- [8] Arulampalam, M. S., Maskell, S., Gordon, N.J., Clapp, T., 2002. A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing* 50, 174-88.
- [9] Azzalini, A., 1985. A class of distributions which includes the normal ones. *Scand. J. Statist.* 12, 171-178.
- [10] Azzalini, A., 1986. Further results on a class of distributions which includes the normal ones. *Statistica* 46, 199-208.
- [11] Azzalini, A., Dal Capello, T., Kotz, S., 2002. Log-skewnormal and log-skew-t distributions as models for family income data. *Journal of Income Distribution*. 11 (3-4), 12-20.
- [12] Azzalini, A., Capitanio, A., 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution. *J. Roy. Statist. Soc. Ser. B* 65, 367-389.
- [13] Azzalini, A., Dal Cappello, T., Kotz, S., 2003. Log-skew-normal and log-skew-t distributions as models for family income data. *J. Income Distribution*, 11, no. 3-4, 12-20.
- [14] Bayarri, M.J., Berger, J.O., 2004. The Interplay of Bayesian and Frequentist Analysis. *Statistical Science*, Vol. 19, No. 1, 58-80.
- [15] Bayes, C.L., 2005. *Inferência Bayesiana No Modelo Normal Assimétrico*. Dissertação de mestrado. IME-USP
- [16] Bayes, C.L., Branco, M. D., 2007. Bayesian inference for the skewness parameter of the scalar skewnormal distribution. *Braz. J. Probab. Stat.* 21, 141-163.

- [17] Birch, C.P., 1999. A new generalized logistic sigmoid growth equation compared with the Richards growth equation. *Annals of Botany* 83:713-723.
- [18] Branco, M. D., Dey, D.K., 2001. A general class of multivariate skew-elliptical distributions. *J. Multivariate Anal.* 79, 99-113.
- [19] Branco, M. D, Genton, M., Eliseo, B., 2012. Objective Bayesian Analysis of Skew-t Distributions. *Scand. J. Statist.* 40, 63-85.
- [20] Braun, M., 1983. *Differential equations and their applications*, 3rd ed. Springer-Verlag. New York-
- [21] Brodziak, J., Ishimura, G., 2011. Development of Bayesian production models for assessing the North Pacific swordfish population. *Fish Sci.*, 77: 23-34.
- [22] Brooks, S. P. 2002. Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. *Bayesian Measures of Model Complexity and Fit. Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, Vol. 64, No.4, 616-620.
- [23] Bucarey, D., Canales, C., Montenegro, C., Zilleruelo, M., Párraga, D., 2014. Informe Consolidado. Proyecto 2.13: Investigación del estatus y posibilidades de explotación biológicamente sustentables en langostino colorado, año 2014. Insituto de Fomento Pesquero, Chile, 82 pp. + Anexos.
- [24] Burnham, K.P., Anderson, D.R., 2002. *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach*, Second Edition. Springer.
- [25] Carlin, B.P., Polson, N.G., Stoffer, D.S., 1992. A Monte Carlo approach to nonnormal and nonlinear state-space modelling. *J. American Statistical Association*, 87, 493-500.
- [26] Carter, C.K., Kohn, R., 1994. On Gibbs sampling for state space models. *Biometrika*, 81:541-553.
- [27] Carter, C.K., Kohn, R., 1996. Markov chain Monte Carlo in conditionally Gaussian state space models. *Biometrika*, 83:589-601.
- [28] Casella, G., Berger, R. L., 2002. *Statistical Inference*. 2nd edition, Wadsworth Group.
- [29] Chai, H.S., Baily, K.R., 2008. Use of log-skew-normal distribution in analysis of continuous data with a discrete component at zero. *Statistics in Medicine* 27: 3643-3655.
- [30] Celexus, G., Forbesy, F., Robertz, C.P., Titterington, D.M., 2006. Deviance Information Criteria for Missing Data Models. *Bayesian analysis*, 4:651-674.
- [31] Commandeur, J., Koopman S.J., 2007. *An Introduction to State Space Time Series Analysis*. Oxford University Press.
- [32] Contreras-Reyes, J.E., Arellano-Valle, R.B., 2013. Growth estimates of cardinalfish (*Epigonus crassicaudus*) based on scale mixtures of skew-normal distributions. *Fish. Res.* 147, 137-144.
- [33] Contreras-Reyes, J.E., Arellano-Valle, R.B., Canales, T.M., 2014. Comparing growth curves with asymmetric heavy-tailed errors: Application to the southern blue whiting (*Micromesistius australis*). *Fish. Res.* 159, 88-94.
- [34] De Lara, M., Doyen, L., 2008. *Sustainable Management of Natural Resources. Mathematical Models and Methods*. Springer-Verlag Berlin Heidelberg.
- [35] de Valpine, P., 2002. Review of methods for fitting time-series models with process and observation error, and likelihood calculations for nonlinear, non-Gaussian state-space models. *Bulletin of Marine Science* 70: 455-471.

- [36] de Valpine, P., Hastings, A., 2002. Fitting population models with process noise and observation error. *Ecological Monographs* 72:57-76.
- [37] Dennis, B., Ponciano, J.M., Lele, S.R., Taper, M.L., Staples, D.F., 2006. Estimating density dependence, process noise, and observation error. *Ecological Monographs* 76:323-341
- [38] Doubleday, W.G., 1976. A least squares approach to analyzing catch at age data. *Int. Comm. Northwest Atl. Fish. Res. Bull.* 12: 69-81.
- [39] Doucet, A., Godsill, S., Andrieu, C., 2000. On sequential Monte Carlo sampling methods for Bayesian Filtering, *Statistics and Computing*, 10, 197-208.
- [40] Durbin, J., Koopman, S.J., 2001. Time series analysis by state space methods. Oxford Statistical Science Series.
- [41] Durbin, J., Koopman, S.J., 2012. Time series analysis by state space methods. Second Edition. Oxford Statistical Science Series.
- [42] Efron, B., 2005. Bayesians, Frequentists, and Scientists. *Journal of the American Statistical Association*, Vol. 100, No. 469, pp. 1-5.
- [43] Ellner, S.P., Guckenheimer, J., 2006. *Dynamic Models in Biology*. Princeton University Press.
- [44] FAO, 2014. Informe sobre el estado mundial de la pesca y la acuicultura, 2014. <http://www.fao.org/3/a-i3720s.pdf>, versión: [17.08.2015]
- [45] Farkas, M., 2001. *Dynamical Models in Biology*. Academic Press.
- [46] Fournier, D.A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., Sibert, J., 2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models, *Optimization Methods and Software*, 27:2, 233-249.
- [47] Gallucci, V. F., Saila, S. B., Gustafson, D. J., Rothschild, B. J., 1996. *Stock Assessment: Quantitative Methods and Applications for Small-Scale Fisheries*. CRC Press, Boca Raton, FL.
- [48] Gamerman D., Lopes, H. F., 2006. *Markov chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. 2nd Edition. Chapman and Hall.
- [49] Garay, A. M., Lachos, V. H., Abanto-Valle, C. A., 2011. Nonlinear regression models based on scale mixture of skew-normal distributions. *Journal of the Korean Statistical Society*, 40, 115-124.
- [50] García, S., Zerbi, A., Aliaume, C., Do-Chi, T., Lasserre, G., 2003. The ecosystem approach to fisheries. Issues, terminology, principles, institutional foundations, implementation and outlook. *FAO Fisheries Technical Paper No. 443*.
- [51] Gelman, A., Meng, X., Stern, H., 1996. Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statist.Sin.*, 6, 733-807.
- [52] Gelman, Carlin, J., Stern, H., Dunson, D., Vehtari, A. Rubin, D., 2014. *Bayesian Data Analysis*. Third Edition. Chapman and Hall CRC Texts in Statistical Science.
- [53] Geman, S., Geman D., 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intelligence* 6, 721:741.
- [54] Genton, M.G., 2004. *Skew-elliptical distributions and their applications: A Journey Beyond Normality*. Chapman and Hall/CRC.

- [55] Geweke, J., Tanizaki, H., 1999. On Markov Chain Monte Carlo Methods for Nonlinear and Non-Gaussian State-Space Models. *Commun. Statist.-Simulation*, 28(4), 867-894.
- [56] Geweke, J., Tanizaki, H., 2001. Bayesian estimation of state-space models using the Metropolis-Hastings algorithm within Gibbs sampling. *Computational Statistics and Data Analysis*, 37, 151-170.
- [57] Gilbert, D.J., 1992. A stock production modeling technique for fitting catch histories to stock index data. *New Zealand Fisheries Assessment Res. Doc.* 92/15.
- [58] Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. *Markov chain Monte Carlo in Practice*. Chapman and Hall, London, U.K.
- [59] Gordon, N.J., Salmond, D.J., Smith A.F.M., 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, Vol 140,2.
- [60] Gunderson, D.R., 1993. *Surveys of Fisheries Resources*. Wiley, New York.
- [61] Gurney, W.S.C., Nisbet, R.M., 1998. *Ecological dynamics*. Oxford University Press.
- [62] Haddon, M., 2001. *Modelling and quantitative methods in fisheries*. Chapman and Hall/CRC, Florida. 406 pp.
- [63] Haddon, M., 2011. *Modelling and quantitative methods in fisheries*. 2nd. Ed., Chapman and Hall/CRC, Florida. 449 pp.
- [64] Harley, S.J., Myers, R.A., Dunn, A., 2001. Is catch-per-unit-effort proportional to abundance?. *Can. J. Fish. Aquat. Sci.*, 58, 1760-1772.
- [65] Hartig, F., Calabrese, J.M., Reineking, B., Wiegand, T., Huth, A., 2011. Statistical inference for stochastic simulation models - theory and application. *Ecology Letters*, 14: 816-827.
- [66] Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57, 97:109.
- [67] Henze, N., 1986. A probabilistic representation of the skew-normal distribution. *Scand.J. Statist.*, 13, 271-275.
- [68] Hilborn, R., Walters, C.J., 1992. *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. London: Chapman and Hall.
- [69] Hilborn, R., Pikitch, E., McAllister, M., 1994. A Bayesian estimation decision analysis for an age-structured model using biomass survey data. *Fish.Res.* 19, 17-30.
- [70] Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T., 1999. Bayesian Model Averaging: A Tutorial. *Statistical Science*, Vol. 14, No. 4, 382-417.
- [71] Jungbacker, B., Koopman, S.J., 2007. Monte Carlo estimation for nonlinear non-Gaussian state space models. *Biometrika*, 94:827-839.
- [72] Kitagawa, G., 1987. Non-Gaussian State-Space Modeling of Nonstationary Time Series (with discussion). *J. Amer. Statist. Assoc.*, 82, 1032-1063.
- [73] Kitagawa, G., 1996. Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models. *Journal of Computational and Graphical Statistics*, Vol. 5, No. 1, pp. 1-25.
- [74] Kitagawa, G., 2010. *Introduction to time series modeling*. CRC Press, Taylor and Francis Group. A Chapman and Hall Book.
- [75] Knight, K., 2000. *Mathematical Statistics*. Chapman and Hall. CRC Press.

- [76] Kotz, S., Nadarajah, S., 2004. Multivariate t-distributions and their applications. Cambridge University Press.
- [77] Krebs, C.J., 1985. Ecology: The experimental Analysis of Distribution and Abundance. 3rd Edition, Harper and Row, New York.
- [78] Kuri, A.I., 2003. Advanced calculus with applications in Statistics. Second Edition. Wiley Series in Probability and Statistics.
- [79] Lachos, V.H., Bandyopadhyay, D., Garay, A. M., 2011. Heteroscedastic nonlinear regression models based on scale mixture of skew-normal distributions. *Statistics and Probability Letters*, 81, 1208-1217.
- [80] Lachos, V.H., Barbosa, C.R., Medina, A., 2013. Modelos não Lineares Assimétricos. 13a Escola de Modelos de Regressão. Maresias-SP, Brazil.
- [81] Li, Y., Zeng, T., Yu, J., 2012. Robust Deviance Information Criterion for Latent Variable Models. Research Collection School Of Economics. Singapore Management University, 8-2012.
- [82] Liseo, B., Loperfido, N., 2006. A note on reference priors for the scalar skew-normal distribution. *Journal of Statistical Planning and Inference*, 136 (373-379).
- [83] Limpert E., Sathel W., Abbt, M., 2001. Log-normal distributions across the Sciences: Keys and Clues. *BioScience*. Vol 51, No. 5.
- [84] Lynch, S., Western B., 2004. Bayesian Posterior Predictive Checks for Complex Models. *Sociological Methods and Research*, Vol. 32, No. 3, 301-335.
- [85] Marchenko, Y.V., Genton, M.G., 2010. Multivariate log-skew-elliptical distributions with applications to precipitation data. *Environmetrics* 21, 318-340.
- [86] Martino, S., Rue, H., 2008. Implementing approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations: A manual for the inla-program. Technical Report no 2, Department of mathematical sciences, Norwegian University of Science and Technology.
- [87] Maunder, M. N., Watters, G.M., 2003. A-SCALA: an age structured statistical catch-at-length analysis for assessing tuna stocks in the eastern Pacific Ocean. *Inter-American Tropical Tuna Commissions Bulletin*, 22: 433-582.
- [88] Maunder, M.N., Deriso, R.B., 2011. A statespace multistage life cycle model to evaluate population impacts in the presence of density dependence: illustrated with application to delta smelt (*Hypomesus transpacificus*) *Can. J. Fish. Aquat. Sci.* 68: 1285-1306.
- [89] McAllister, M.K., Pikitch, E.K., Punt, A.E., Hilborn, R., 1994. A bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. *Can. J. Fish. Aquat. Sci.*, 51, 2673-2687.
- [90] McAllister, M., Ianelli, J., 1997. Bayesian stock assessment using catch-age data the sampling-importance resampling algorithm. *Can.J.Fish.Aquat.Sci.* 54, 284-300.
- [91] McAllister, M.K., Kirkwood, G.P., 1998. Bayesian stock assessment: a review and example application using the logistic model. *ICES. J. Mar. Sci.* 55, 1031-1060.
- [92] Meinhold, R. J., Singpurwalla, N. D., 1983. Understanding the Kalman Filter. *The American Statistician*, Vol. 37, No. 2, pp. 123-127
- [93] Methot, R.D., 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. *International North Pacific Fishery Commission Bull.* 50:259-277.

- [94] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machine. *J. Chem. Phys.* 21, 1087:1091.
- [95] Meyer, R., Millar, R.B., 1999. BUGS in Bayesian Stock Assessments. *Can. J. Fish. Aquat. Sci.*, 56, 1078-1087.
- [96] Migon, H., Gamerman, D., Lopes, H.F., Ferreira, M.R., 2005. Dynamic Models. In *Handbook of Statistics*, 25, 557-592.
- [97] Millar, R.B., Meyer, R., 1999. Bayesian statespace modeling of age-structured data: fitting a model is just the beginning. *Can. J. Fish. Aquat. Sci.* 57, 43:50.
- [98] Millar, R.B., Meyer, R., 2000. Non-Linear State-Space Modeling of Fisheries Biomass Dynamics by Using MetropolisHastings within-Gibbs sampling. *Applied Statistics*, 49, 327-342.
- [99] Montenegro, C., Maunder M., Zilleruelo, M., 2009. Improving management advice through spatially explicit models and sharing information. *Fisheries Research*, 100, 191-199.
- [100] Montenegro, C., 2010. Modelamiento Bayesiano de la dinámica poblacional del recurso langostino amarillo, utilizando un modelo de crecimiento logístico en representación espacio de estados. Trabajo de tesis presentado para optar al grado Académico de Magíster en Estadística. Universidad de Valparaíso, Chile.
- [101] Ntzoufras, I., 2009. *Bayesian Modeling Using WinBUGS*. Wiley Series in Computational Statistics.
- [102] Papoulis, A., 1984. *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill.
- [103] Pella, J.J., Tomlinson, P.K., 1969. A generalized stock production model. *Inter-Am. Trop. Tuna Comm. Bull.* 13, 421-496.
- [104] Petris, G., Petroni, S., Campagnoli, P., 2009. *Dynamic linear models with R*. Springer.
- [105] Pitt, M., Shephard, N., 1999. Filtering via simulation: auxiliary particle filters. *Journal of the American Statistical Association*, 94, 590-599.
- [106] Polachek, T., Hilborn, R., Punt, A.E., 1993. Fitting surplus production models: comparing methods and measuring uncertainty. *Can. J. Fish. Aquat. Sci.* 50: 2597-2607.
- [107] Prado, R., West, M., 2010. *Time Series: Modelling, Computation and Inference*. Chapman Hall/CRC.
- [108] Prager, M.H., 2002. Comparison of logistic and generalized surplus-production models applied to swordfish, *Xiphias gladius*, in the north Atlantic Ocean. *Fish. Res.* 58, 41-57.
- [109] Punt, A., Kennedy, R., 1997. Population modelling of Tasmanian rock lobster, *Jasus edwardsii*, resources. *Mar.Freshwater Res.* 48, 967-980
- [110] Punt, A., 2003. Extending production models to include process error in the population dynamics. *Can. J. Fish. Aquat. Sci.*: 60: 1217-1228.
- [111] Punt, A., Hilborn, R., 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. *Rev. Fish Biol. Fish.* 7: 35-63.
- [112] Quinn, T.J., Deriso, R., 1999. *Quantitative Fish Dynamics*. Oxford University Press. New York.
- [113] R Core Team, 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

- [114] Ripley, B.D., 1987. Stochastic Simulation. New York: Willey.
- [115] Rue, H., Martino, S., Lindgren, F., Simpson, D., Riebler, A., Krainski, E.T., 2015. Functions which allow to perform full Bayesian analysis of latent Gaussian models using Integrated Nested Laplace Approximation. www.r-inla.org.
- [116] Russel, E.S., 1931. Some theoretical considerations on the overfishing problem. *Journal du Conseil International pour l'Exploration de la Mer*, 6:3-20.
- [117] Schaefer, M.B., 1954. Some aspects of the dynamics of populations important to the management of commercial marine fisheries. *IATTC Bull.* 1, 25-56.
- [118] Seber, G.A.F., 1982. The estimation of animal abundance and related parameters. 2nd. Ed., Macmillan, NY, 654 pp.
- [119] Shumway R., Stoffer, D., 2011. Time series analysis and its applications (with R examples). Third edition. Springer Texts in Statistics
- [120] Skaug, H., Fournier, D., 2006. Automatic approximation of the marginal likelihood in non-Gaussian hierarchical models. *Computational Statistics and Data Analysis*, 56:699-709.
- [121] Smith M.T., Addison, J.T., 2003. Methods for stock assessment of crustacean fisheries. *Fisheries Research*, 65, 231:256.
- [122] Spiegelhalter, D.j., Best, N.G., Carlin, B.P., Linde, A., 2002. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, Vol. 64, No.4, pp. 583-639.
- [123] Sturtz, S., Ligges U., Gelman, A., 2005. R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software*, 12(3), 1-16.
- [124] Sullivan, P. J., Lai, H.-L., Gallucci, V. F., 1990. A catch-at-Length Analysis that Incorporates a Stochastic Model of Growth. *Can.J.Fish.Aquat.Sci.* 47, 184-198.
- [125] Sullivan, P. J., 1992. A Kalman filter approach to catch-at-length analysis. *Biometrics* 48:237-257.
- [126] Tanizaki, H., 2001. Nonlinear and non-Gaussian state space modeling using sampling techniques. *Ann. Inst. Statist. Math.* Vol 53,1,63-81.
- [127] Tanizaki, H., Mariano, R., 1998. Nonlinear and non-Gaussian state space modeling with Monte Carlo simulations. *Journal of Econometrics* Vol 83,263-290.
- [128] Thomas, A., O'Hara, B., Ligges, U., Sturtz, S., 2006. Making BUGS Open. *R News* 6: 12-17.
- [129] Thompson, W.F., Bell, F.H., 1934. Biological statistics of the Pacific halibut fishery. (2) Effect of changes in intensity upon total yield and yield per unit of gear. *Int. Fisheries Comm. Report* 8.
- [130] Tierney, L., 1994. Markov chains for exploring posterior distributions (with discussion). *Ann. Stat.*, 22, 1701-1728.
- [131] Tierney, L., Kadane, J.B., 1986. Accurate approximations for posterior moments and marginal distributions. *J. Amer. Statist. Assoc.* 81, 82-86.
- [132] Von Bertalanffy, L., 1938. A quantitative theory of organic growth (inquiries on growth laws. II). *Hum. Biol.* 10, 181-213.
- [133] Watanabe, T. Omori, Y., 2004. A multi-move sampler for estimate non-Gaussian time series model: Comments on Shepard and Pitt (1997). *Biometrika*, 91, 246-248.

- [134] Welch, G., Bishop, G., 2001. An Introduction to the Kalman Filter. University of North Carolina at Chapel Hill, Department of Computer Science Chapel Hill, NC 27599-3175.
- [135] West, M., Harrison, P.J., 1997. Bayesian Forecasting and Dynamic Models (2nd ed.). Springer Verlag.
- [136] Wilberg M.J, Thorson J.T, Linton B.C, Berkson J., 2010. Incorporating time-varying catchability into population dynamic stock assessment models. *Rev. Fish. Sci.* 18(1): 7-24.

Apêndice A

A.1 Provas proposições

a) Prova Proposição 2.1.

Para mostrar (i) note que X_t e $Y_{t+1:n}$ são condicionalmente independentes dado X_{t+1} . Além disso X_{t+1} e $Y_{1:n}$ são condicionalmente independentes dado X_t .

Usando formula de Bayes, temos

$$\begin{aligned} p(x_t|x_{t+1}, y_{1:n}) &= p(x_t|x_{t+1}, y_{1:t}) \\ &= \frac{p(x_t|y_{1:t})p(x_{t+1}|x_t, y_{1:t})}{p(x_{t+1}|y_{1:t})} \\ &= \frac{p(x_t|y_{1:t})p(x_{t+1}|x_t)}{p(x_{t+1}|y_{1:t})} \end{aligned}$$

Para mostrar (ii), marginalize $p(x_t, x_{t+1}|y_{1:n})$ com respeito a X_{t+1} :

$$\begin{aligned} p(x_t|y_{1:n}) &= \int p(x_t, x_{t+1}|y_{1:n})dx_{t+1} \\ &= \int p(x_{t+1}|y_{1:n})p(x_t|x_{t+1}, y_{1:n})dx_{t+1} \\ &= \int p(x_{t+1}|y_{1:n})\frac{p(x_{t+1}|x_t)p(x_t, y_{1:t})}{p(x_{t+1}|y_{1:t})}dx_{t+1} \\ &= p(x_t|y_{1:t}) \int p(x_{t+1}|x_t)\frac{p(x_{t+1}|y_{1:n})}{p(x_{t+1}|y_{1:t})}dx_{t+1} \end{aligned}$$

b) Prova Proposição 2.2.

O vetor aleatório $(X_0, X_1, \dots, X_t, Y_1, \dots, Y_t)$ tem distribuição conjunta dada por:

$$p(x_{0:t}, y_{1:t}) = p(x_0) \prod_{j=1}^t p(x_j | x_{j-1})p(y_j | x_j) \quad (\text{A.1})$$

onde as distribuições marginais e condicionais são gaussianas. De resultados estandar da distribuição Normal Multivariada, a densidade conjunta de $(X_0, X_1, \dots, X_t, Y_1, \dots, Y_t)$ é gaussiana para todo $t \geq 1$. Consequentemente a distribuição de qualquer subvetor é também gaussiana, como também a distribuição condicional de alguns componentes dado alguns outros. Portanto a distribuição preditiva e a distribuição de filtragem são gaussianas e logo é suficiente encontrar suas médias e variâncias.

Para provar (i), seja $X_t|Y_{1:t-1} \sim N(a_t, R_t)$. a_t e R_t podem ser obtidos como:

$$\begin{aligned}
a_t &= E(X_t|Y_{1:t-1}) = E(E(X_t|X_{t-1}, Y_{1:t-1})|Y_{1:t-1}) \\
&= E(G_t X_{t-1}|Y_{1:t-1}) = G_t m_{t-1} \\
R_t &= \text{Var}(X_t|Y_{1:t-1}) \\
&= E(\text{Var}(X_t|X_{t-1}, Y_{1:t-1})|Y_{1:t-1}) + \text{Var}(E(X_t|X_{t-1}, Y_{1:t-1})|Y_{1:t-1}) \\
&= W_t + G_t C_{t-1} G_t'
\end{aligned}$$

Para provar (ii), seja $Y_t|Y_{1:t-1} \sim N(f_t, Q_t)$. f_t e Q_t podem ser obtidos como:

$$f_t = E(Y_t|Y_{1:t-1}) = E(E(Y_t|X_t, Y_{1:t-1})|Y_{1:t-1}) = E(F_t X_t|Y_{1:t-1}) = F_t a_t$$

$$\begin{aligned}
Q_t &= \text{Var}(Y_t|Y_{1:t-1}) \\
&= E(\text{Var}(Y_t|X_t, Y_{1:t-1})|Y_{1:t-1}) + \text{Var}(E(Y_t|X_t, Y_{1:t-1})|Y_{1:t-1}) \\
&= V_t + F_t R_t F_t'.
\end{aligned}$$

Para provar (iii), notemos que para obter a distribuição de filtragem utilizamos a fórmula de Bayes para combinar a distribuição *a priori* $p(x_t|y_{1:t-1})$ e a verosimilhança $p(x_t|x_t)$, ou seja:

$$p(x_t | y_{1:t}) = \frac{p(y_t | x_t) p(x_t | y_{1:t-1})}{p(y_t | y_{1:t-1})} \quad (\text{A.2})$$

Assim o problema é o mesmo como que no caso do modelo de regressão linear Bayesiano com:

$$Y_t = F_t X_t + v_t, \quad v_t \sim N(0, V_t)$$

com o parâmetro de regressão X_t com uma distribuição *a priori* conjugada gaussiana $N(a_t, R_t)$ (aqui V_t é conhecido). Assim temos que:

$$X_t|Y_{1:t} \sim N(m_t, C_t),$$

onde a média e a variância estão dadas por:

$$m_t = a_t + R_t F_t' Q_t^{-1} (Y_t - F_t a_t),$$

e

$$C_t = R_t - R_t F_t' Q_t^{-1} F_t R_t.$$

c) Prova Proposição 2.3.

Usando as propriedades de independencia condicional do modelo, temos

$$\begin{aligned}
p(x_{t+k}|y_{1:t}) &= \int p(x_{t+k}, x_{t+k-1}|y_{1:t}) dx_{t+k-1} \\
&= \int p(x_{t+k}|x_{t+k-1}, y_{1:t}) p(x_{t+k-1}|y_{1:t}) dx_{t+k-1} \\
&= \int p(x_{t+k}|x_{t+k-1}) p(x_{t+k-1}|y_{1:t}) dx_{t+k-1}
\end{aligned}$$

do qual segue (i). para provar (ii) nos baseamos novamente nas propriedades de independencia

condicional do modelo. Temos que

$$\begin{aligned} p(y_{t+k}|y_{1:t}) &= \int p(y_{t+k}, x_{t+k}|y_{1:t}) dx_{t+k} \\ &= \int p(y_{t+k}|x_{t+k}, y_{1:t}) p(x_{t+k}|y_{1:t}) dx_{t+k} \\ &= \int p(y_{t+k}|x_{t+k}) p(x_{t+k}|y_{1:t}) dx_{t+k} \end{aligned}$$

do qual segue (ii).

d) Distribuição Normal/Gamma univariada.

Sea $\phi \sim G[n/2, d/2]$ para qualquer $n > 0$ e $d > 0$, e suponha que a distribuição condicional de uma variável aleatória X dada ϕ é normal ($X|\phi \sim N[m, C\phi^{-1}]$, para algum $m > 0$ e C). Note que $E[X] \equiv E[X|\phi] = m$ não depende de ϕ . Mas $V[X|\phi] = C\phi^{-1}$. A distribuição conjunta de X e ϕ é chamada de Normal/Gamma (univariada). Note que

$$p(X, \phi) = \left(\frac{\phi}{2\pi C}\right)^{1/2} \exp\left[-\frac{\phi(X-m)^2}{2C}\right] \quad (\text{A.3})$$

$$\times \frac{d^{n/2}}{2^{n/2}\Gamma(\frac{n}{2})} \phi^{\frac{n}{2}-1} \exp\left[-\frac{\phi d}{2}\right] \quad (\text{A.4})$$

$$\propto \phi^{(\frac{n+1}{2})-1} \exp\left[-\frac{\phi}{2}\left\{\frac{(X-m)^2}{C} + d\right\}\right] \quad (\text{A.5})$$

como função de X e ϕ .

$$p(\phi|X) \propto \phi^{(\frac{n+1}{2})-1} \exp\left[-\frac{\phi}{2}\left\{\frac{(X-m)^2}{C} + d\right\}\right] \quad (\text{A.6})$$

assim

$$(\phi|X) \propto G\left[\frac{n^*}{2}, \frac{d^*}{2}\right]$$

onde $n^* = n + 1$ e $d^* = d + \frac{(X-m)^2}{C}$.

$$p(X) = p(X, \phi)/p(\phi|X) \propto [n + (X-m)^2/R]^{-\frac{n+1}{2}}, \quad (-\infty < X < \infty) \quad (\text{A.7})$$

onde $R = C(d/n) = C/E[\phi]$. Isto é proporcional à densidade da distribuição t -student com n graus de liberdade, moda m e escala R . Portanto $X \sim T_n[m, R]$, ou $(X-m)/R^{1/2} \sim T_n[0, 1]$, a distribuição t -student standard.

A.2 Algoritmos quase-Newton

A classe de algoritmos quase-Newton são métodos utilizados para encontrar zeros ou máximos locais e mínimos de funções, como uma alternativa ao método de Newton. Eles podem ser usados se a Jacobiana ou Hessiana não está disponível ou é demasiado cara para calcular em cada iteração. A seguir, apresentamos brevemente o algoritmo quase-Newton.

Dado um ponto inicial $x_0 \in (f)$, $H_0 > 0$ para $k = 1, 2, \dots$ até que um critério de parada é satisfeito,

1. computar direção quase-Newton $\Delta x = -H_{k-1}^{-1} \nabla f(x^{k-1})$, onde H é uma aproximação da Hessiana,

2. determinar tamanho do passo t
3. computar $x^k = x^{k-1} + t\Delta x$
4. computar H_k

Diferentes métodos usam regras diferentes para atualizar H no passo 4. A seguir apresentamos alguns dos métodos mais utilizados.

a) Broyden-Fletcher-Goldfarb-Shannon (BFGS)

Neste caso, H é atualizada da forma

$$H_k = H_{k-1} + \frac{yy^T}{y^T s} - \frac{H_{k-1}ss^T H_{k-1}}{s^T H_{k-1}s}$$

onde $s = x^k - x^{k-1}$ e $y = \nabla f(x^k) - \nabla f(x^{k-1})$ e,

$$H_k^{-1} = \left(I - \frac{sy^T}{y^T s}\right)H_{k-1}^{-1}\left(I - \frac{ys^T}{y^T s}\right) + \frac{ss^T}{y^T s}$$

Note que $y^T s > 0$ para f estritamente convexa.

b) Davidon-Fletcher-Powell (DFP)

Aqui H_{k-1} e X são objetivos em

$$\begin{array}{ll} \text{minimize} & \text{tr}(H_{k-1}X^{-1}) - \log \det(H_{k-1}X^{-1}) - n \\ \text{sujeito a} & Xs = y \end{array}$$

- minimizar entropia relativa entre $\mathcal{N}(0, H_{k-1})$ e $\mathcal{N}(0, X)$
- O problema é convexo em X^{-1} (com restrição $s = X^{-1}y$)
- solução é dual da formula BFGS

$$H_k = \left(I - \frac{ys^T}{s^T y}\right)H_{k-1}\left(I - \frac{sy^T}{s^T y}\right) + \frac{yy^T}{s^T y}$$

c) Metodo quase-Newton de memória limitada

A principal desvantagem do método quase-Newton é armazenar H_k ou H_k^{-1} por memória limitada.

- Por exemplo, em vez disso armazenamos os dados mais recentes de

$$s_j = x^j - x^{j-1}, \quad y_j = \nabla f(x^j) - \nabla f(x^{j-1})$$

- Avaliamos $\Delta x = H_k^{-1}\nabla f(x^k)$ recursivamente, usando

$$H_j^{-1} = \left(I - \frac{s_j y_j^T}{y_j^T s_j}\right)H_{j-1}^{-1}\left(I - \frac{y_j s_j^T}{y_j^T s_j}\right) + \frac{s_j s_j^T}{y_j^T s_j}$$

para $j = k, k-1, \dots, k-m+1$, assumindo que $H_{k-m}^{-1} = I$.

- custo por iteração é $O(nm)$ e armazenamento é $O(nm)$.

A.3 Diferenciação Automática

A diferenciação automática, também conhecida como diferenciação algorítmica, é um conjunto de técnicas e ferramentas com as quais avaliamos numericamente a derivada de uma função definida através de um código fonte numa linguagem de programação.

A AD está baseada no fato que para avaliar uma função numa linguagem de programação dado é executado uma sequência de operações aritméticas (adição, subtração, multiplicação e divisão) e chamados as funções elementares (exp, log, sin, cos, etc). Aplicando a regra da cadeia, simultaneamente com estas operações é possível calcular a derivada de qualquer ordem tão exatas como a aritmética da máquina permita.

A base da AD é a decomposição de diferenciais que providencia a regra da cadeia. Para uma composição de funções $f(x) = g(h(x))$ obtém-se, a partir da regra da cadeia,

$$\frac{df}{dx} = \frac{df}{dh} \frac{dh}{dx}$$

Existem duas formas de aplicar a AD, a forma para frente (forward) e a forma para atrás (reverse). A forma forward obtém-se ao aplicar a regra da cadeia de direita para esquerda, isto é, primeiro calcula-se $\frac{dh}{dx}$ e depois $\frac{df}{dh}$, em quanto que a forma reverse obtém-se quando se aplica a regra da cadeia de esquerda para direita.

Os seguintes exemplos ilustram o funcionamento da forma forward:

Exemplo 1 Deseja-se calcular o valor e a derivada de

$$f(x) = x^2 \cos(x)$$

no ponto $x = \pi$.

A seguinte tabela contém as operações necessárias para avaliar a função.

Tabela A.1: Valor e f em $x = \pi$

w_1	=	x	=	π
w_2	=	w_1^2	=	π^2
w_3	=	$\cos(w_1)$	=	-1
w_4	=	$w_2 w_3$	=	$-\pi^2$
y	=	w_4	=	$-\pi^2$

Para aplicar a forma para frente, armazena-se numa nova variável a derivada de cada operação com respeito a variável x . Esta derivada pode-se calcular ao mesmo tempo que se realiza a operação. A seguinte tabela ilustra este procedimento.

Tabela A.2: Valor da derivada de f em $x = \pi$

$\dot{w}_1$	=	$\dot{x}$	=	1
$\dot{w}_2$	=	$2w_1 \dot{w}_1$	=	2π
$\dot{w}_3$	=	$-\sin(w_1) \dot{w}_1$	=	0
$\dot{w}_4$	=	$w_2 \dot{w}_3 + \dot{w}_2 w_3$	=	-2π
$\dot{y}$	=	$\dot{w}_4$	=	-2π

Na tabela anterior, as variáveis $\dot{\omega}_i$ armazenam a derivada com respeito a x da operação ω_i . Como é desejado calcular a derivada com respeito à variável x , então a nova variável $\dot{\omega}_1 = \frac{d\omega_1}{dx}$ se inicia com valor 1. Ao finalizar o cálculo, na variável $\dot{\omega}_4$ tem-se o valor da derivada de f com respeito a x .

O seguinte exemplo mostra como pode se aplicar a forma para frente para calcular as derivadas parciais de funções com mais de uma variável.

Exemplo 2 Deseja-se calcular o valor e o gradiente de

$$f(x_1, x_2) = x_1 x_2 + \sin(x_1)$$

no ponto $x_1 = \pi$ e $x_2 = 2$.

A função f deste exemplo pode se expressar através das seguintes operações

Tabela A.3: Valor e f em $x_1 = \pi, x_2 = 2$

w_1	=	x_1	=	π
w_2	=	x_2	=	2
w_3	=	$w_1 w_2$	=	2π
w_4	=	$\sin(w_1)$	=	0
w_5	=	$w_3 + w_4$	=	2π
y	=	w_5	=	2π

Para calcular o gradiente utilizando a forma forward temos que utilizar o mesmo procedimento do exemplo anterior, mas neste caso duas vezes: uma por cada derivada parcial. A continuação mostramos o cálculo da derivada parcial $\frac{df}{dx_1}$.

Tabela A.4: Valor da derivada parcial $\frac{\partial f}{\partial x_1}$ em $x_1 = \pi, x_2 = 2$

$\dot{w}_1$	=	$\dot{x}_1$	=	1
$\dot{w}_2$	=	$\dot{x}_2$	=	0
$\dot{w}_3$	=	$\dot{w}_1 w_2 + w_1 \dot{w}_2$	=	2
$\dot{w}_4$	=	$\cos(w_1) \dot{w}_1$	=	-1
$\dot{w}_5$	=	$\dot{w}_3 + \dot{w}_4$	=	1
$\dot{y}$	=	$\dot{w}_5$	=	1

Neste caso, a variável $\dot{\omega}_2$ se inicia com valor 0 pois $\frac{\partial x_2}{\partial x_1} = 0$, $\omega_2 = x_2$ e $\dot{\omega}_2 = \frac{\partial \omega_2}{\partial x_1}$. Ao finalizar os cálculos, na variável $\dot{\omega}_5$ obtem-se o valor da derivada $\frac{\partial f(x)}{\partial x_1}$.

O procedimento apresentado nos exemplos anteriores pode ser generalizado da seguinte forma:

Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ uma função diferenciável. Denotamos as n variáveis reais de entrada como

$$w_{i-n} = x_i, \quad \dot{w}_{i-n} = \dot{x}_i, \quad i = 1, \dots, n, \quad \dot{x}_i = 1$$

As variáveis $\dot{x}_i$ se iniciam com o valor das derivadas parciais de cada variável x_i com respeito a variável original. Se deseja-se calcular a derivada parcial de f respeito a x_i , então $\dot{x}_i$ deveria ser 1 e $\dot{x}_j$ deveria ser 0 para toda j diferente de i .

Como f esta composta por funções elementares, é conveniente denota-as de alguma forma. A j -esima função elementar será denotada por ϕ_j .

Para descompor f em funções elementares, denotamos novas variáveis w_i e $\dot{w}_i$, como sendo

$$\begin{aligned} w_i &= \phi_i(w_j), \quad j < i \\ \dot{w}_i &= \sum_{j < i} \frac{\partial}{\partial w_j} \phi_i(w_j) \dot{w}_j, \end{aligned}$$

onde, neste caso $i = 1, \dots, \ell$ e ℓ é o número de operações elementares que compõem f . A notação $\phi_i(w_j)$ com $j < i$ significa que ϕ_i depende diretamente de w_j , isto é, para avaliar ϕ_i se usa explicitamente w_j .

Também é usual utilizar a seguinte notação

$$w_i = (\phi_i(w_j))_{j < i}$$

Com as notações anteriores podemos expressar o procedimento geral da seguinte forma:

Tabela A.5: Valor e derivadas de f

w_i	=	x_i	
$\dot{w}_{i-n}$	=	$\dot{x}_i$	$i = 1, \dots, n$
w_i	=	$\phi_i(w_j)$	$j < i$
$\dot{w}_i$	=	$\sum_{j < i} \frac{\partial}{\partial w_j} \phi_i(w_j) \dot{w}_j$	$i = 1, \dots, l$
y_{m-i}	=	w_{l-i}	
$\dot{y}_{m-i}$	=	$\dot{w}_{l-i}$	$i = m - 1, \dots, 0$

Como apreciamos na tabela anterior, o procedimento geral da forma forward está estruturado em três partes: primeiro iniciam-se as variáveis, depois executam-se as operações e se calculam as derivadas, e finalmente mostram-se os valores e as derivadas calculadas.

A.4 Diferenciação implícita

Seja $(\theta_0, u_0) \in \mathbb{R}^n \times \mathbb{R}^m$ um ponto satisfazendo a equação $F(\theta_0, u_0) = C$, onde C é uma constante e $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ é uma função diferenciável tal que $\frac{\partial F}{\partial u}$ seja não singular. Então existe uma função $u : N \rightarrow M$ continuamente diferenciável, onde $N \subset \mathbb{R}^n$ e $M \subset \mathbb{R}^m$ são abertos e $F(\theta, u(\theta)) = C$, para cada $\theta \in N$.

Este enunciado pertence ao teorema bem conhecido como teorema da função implícita. Neste caso, podemos derivar e obter

$$0 = \frac{\partial F}{\partial \theta}(\theta, u(\theta)) = \frac{\partial F}{\partial \theta} + \frac{\partial F}{\partial u} \frac{\partial u}{\partial \theta}$$

e consequentemente

$$\frac{\partial u}{\partial \theta} = - \left[\frac{\partial F}{\partial u} \right]^{-1} \frac{\partial F}{\partial \theta}.$$

Por exemplo, esta igualdade pode ser aplicada à função $F = \frac{\partial g(y_{1:n}, \Theta, \hat{\mathbf{x}}_\Theta)}{\partial u}$.

Apêndice B

B.1 Códigos computacionais

B.1.1 Códigos R: Modelo de simulação

```
#Gerador da Skew Normal

geraSN<-function(n,mu,sigma2,shape){
delta<-shape/sqrt(1+shape^2)
y<-mu+sqrt(sigma2)*(delta*abs(rnorm(n))+sqrt(1-delta^2)*rnorm(n))
return(y)}

#Gerador da Skew T

geraST<-function(n,mu,sigma2,shape,nu){
y<-mu+(rgamma(n,nu/2,nu/2))^(1/2)*geraSN(n,0,sigma2,shape)
return(y)}

#Parâmetros, constantes e vetores

T<-seq(10,100,10)

M=100

K=40
r=0.4
tau=3
q=10

sigma=12
lambda=3
nu=7

#####
#Modelo dos estados
#####

models<-list()
catches<-list()

for (k in 1:length(T))
{
```

```

C=rep(0.500,T[k])
B<-numeric()
B[1]=K
for(t in 2:T[k])
{
B[t]<-rnorm(1,B[t-1]+r*B[t-1]*(1-(B[t-1]/K))-C[t-1],tau)
}
models[[k]]<-B
catches[[k]]<-C
}

#####
#Modelo para as observações
#####

b2=sqrt(nu/pi)*gamma((nu-1)/2)/gamma(nu/2)
d2=lambda/sqrt(1+lambda^2)
delta=sigma*d2

datas<-list()

for (k in 1:length(T))
{
I<-matrix(NA,nrow=M,ncol=T[k])
for(i in 1:M)
{
for(t in 1:T[k])
{
I[i,t]<-geraST(1,(q*models[[k]][t])-delta*b2,sigma^2,lambda,nu)
}
}
datas[[k]]<-I
}

```

B.1.2 Códigos OpenBugs: Aplicação dados reais caso 1

```

model
{

# Ecuación de estado

Bmed[1] <- K
itau0<-itau2/100
B[1] ~ dnorm(Bmed[1],itau0);
Dep[1]<-B[1]/K
Hr[1]<-C[1]/B[1]
for (t in 2:N) {
Bmed[t] <- B[t-1] + r*B[t-1]*(1-B[t-1]/K) - C[t-1]
B[t] ~ dnorm(Bmed[t],itau2)
Dep[t]<-B[t]/K
Hr[t]<-C[t]/B[t]
}
}

```

```
#####
#Verossimilhança dos dados
#####

for (t in 1:N) {

#Modelamos a media

I[t] ~ dnorm(Imed[t],tau1[t])
Imed[t]<-q*B[t]+prec*delta*tt[t]
ttt[t]~ dnorm(0,1)I(0,)
tt[t]<-b+ttt[t]/sqrt(w[t])

#Modelamos a precisão
tau1[t]<- w[t]/Gamma
w[t]~ dgamma(a,a)

}

#####
delta<-lambda/sqrt(1+pow(lambda,2))
prec<-1/sqrt(isigma2)
Gamma<-sigma2*(1-pow(delta,2))

#####

k1<-sqrt(nu/2)*exp(loggam((nu-1)/2))/exp(loggam(nu/2))
b<--sqrt(2/3.1416)*k1

#####
#Distribuições a priori para os parâmetros
#####

#####
#K

sigm1<-0.25
prec1<-1/(sigm1*sigm1)
K ~ dlnorm(3.43,prec1)

#####
#r

sigm2<-0.25
prec2<-1/(sigm2*sigm2)
r ~ dlnorm(-1.17,prec2)

#####
# Priori para q

iq ~ dgamma(0.01,0.01)
q <- 1/iq
```

```
#####
#Erro de processo

itau2 ~ dgamma(0.01,0.01)
tau2 <- 1/itau2
tau <- sqrt(tau2)

#####
#nu

a<-0.5*nu
nu~dexp(aux1)I(1,)
aux1~dunif(0.02,0.5)

#####
#isigma

isigma2 ~ dgamma(0.01,0.01)
sigma2 <- 1/isigma2
sigma <- sqrt(sigma2)

#####
#Priori para lambda

lambda ~ dnorm(0,ww)
ww<-0.4052847*w
w ~ dgamma(0.25,0.25)

#####
#Uniforme en delta
#lambda ~ dnorm(0,ww)
#ww<-2*w
#w ~ dgamma(1,1)

#####
#Indicadores
MSY <- r*K/4
HMSY <- r/2
BMSY<-K/2
Bfor <- B[N]+r*B[N]*(1-B[N]/K) -C[N]
Ifor <-Imed[N]
}
```

B.1.3 Código ADMB

```
GLOBALS_SECTION
#include <admodel.h>
#include <stdio.h>
#include <time.h>
```



```

time_t start, finish;
long hour, minute, second;
double elapsed_time;
ofstream mcmc_report("mcmc.csv");

```

TOP_OF_MAIN_SECTION

```

time(&start);
arrmbldsize = 90000000;
gradient_structure::set_GRADSTACK_BUFFER_SIZE(1.e7);
gradient_structure::set_CMPDIF_BUFFER_SIZE(1.e7);
gradient_structure::set_MAX_NVAR_OFFSET(8196);
gradient_structure::set_NUM_DEPENDENT_VARIABLES(5000);

```

DATA_SECTION

```

init_int N
init_matrix data(1,N,1,2)

vector I(1,N)
vector C(1,N)
!!I=(column(data,1)); //CPUE
!!C=(column(data,2)); //Catch

```

```
int reporte_mcmc
```

PARAMETER_SECTION

```

init_bounded_number logK(2.3,6.21,1)
init_number logq
init_bounded_number logr(-2.3,0.69,1)
init_number logsigma //escala
init_number logVar_B //Process variance
init_number loglambda //

```

```
random_effects_vector B(1,N,5);
```

```

sdreport_number K
sdreport_number q
sdreport_number r
sdreport_number sigma
sdreport_number Var_B
sdreport_number lambda
sdreport_number Sigma_I

```

```

vector Bmed(1,N)
vector I_med(1,N)
vector c2(1,N)
vector xi_y(1,N)
vector c(1,N)
vector z(1,N)

```

```
number pi
```

```

number delta
number xi

objective_function_value f;

PROCEDURE_SECTION
  f=0;

  r=mfexp(logr);
  K=mfexp(logK);
  q=mfexp(logq);
  Var_B=mfexp(logVar_B);
  sigma=mfexp(logsigma);
  lambda=mfexp(loglambda);

  Sigma_I=sqrt(pow(sigma,2)*(1-(2/PI)*pow(delta,2)));

  //Equação de evolução
  Bmed(1)=K;
  f+=0.5*(log(2.0*PI*Var_B)+square(B(1)-Bmed(1))/100*Var_B);

  for(int i=2;i<=N;i++)
  {
    Bmed(i)=B(i-1)+r*B(i-1)*(1-B(i-1)/K)-C(i-1);
    f+=0.5*(log(2.0*PI*Var_B)+square(B(i)-Bmed(i))/Var_B);
  }

  //Equação de observação
  for(int i=1; i<=N; i++)
  {
    c2(i) = q*B(i);
    delta = lambda/sqrt(1+square(lambda));
    xi = -delta*sqrt(2/PI);
    xi_y(i) = c2(i)+xi;
    c(i) = lambda*(I(i)-xi_y(i))/sigma;
    z(i) = cumd_norm(c(i));
    f+=-log(2) +0.5*log(2*PI) + log(sigma) +
    0.5*square((I(i)-xi_y(i))/sigma) - log(z(i));
  }

REPORT_SECTION
  report << "Biomass RE" << endl;
  report << B << endl;
  report << "Mean Biomass" << endl;
  report << Bmed << endl;
  report << "Mean CPUE" << endl;
  report << xi_y << endl;
  report << "K r q Var_B sigma lambda Sigma_I" << endl;
  report << K <<" " <<r <<" " <<q <<" " <<Var_B<<" "
<<sigma<<" " << lambda<<" " <<Sigma_I <<endl;

```

B.2 Densidades condicionais completas

B.2.1 Modelo com erros de observação e proceso normais

a) Densidade condicional completa de X_t

Para $2 \leq t \leq n-1$ temos:

$$\begin{aligned} & p(X_t|K, r, \sigma^2, q, \tau^2, X_1, \dots, X_{t-1}, X_{t+1}, \dots, X_n, Y_1, \dots, Y_n) \\ & \propto p(X_t|X_{t-1}, K, r, \sigma^2) p(Y_t|X_t, q, \tau^2) p(X_{t+1}|X_t, K, r, \sigma^2) \\ & \propto \exp \left(- \frac{\{X_t - [X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}]\}^2}{2\sigma^2} \right. \\ & \quad \left. - \frac{(Y_t - qX_t)^2}{2\tau^2} - \frac{\{X_{t+1} - [X_t + rX_t(1 - X_t) - H_t]\}^2}{2\sigma^2} \right) \end{aligned}$$

Para $t = 1$ temos:

$$\begin{aligned} & p(X_t|K, r, \sigma^2, q, \tau^2, X_2, \dots, X_n, Y_1, \dots, Y_n) \\ & \propto p(X_t|K, \sigma^2) p(Y_t|X_t, q, \tau^2) p(X_{t+1}|X_t, K, r, \sigma^2) \\ & \propto \exp \left(- \frac{(X_t - K)^2}{2\sigma^2} - \frac{(Y_t - qX_t)^2}{2\tau^2} - \frac{\{X_{t+1} - [X_t + rX_t(1 - X_t) - H_t]\}^2}{2\sigma^2} \right) \end{aligned}$$

Para $t = n$ temos:

$$\begin{aligned} & p(X_t|K, r, \sigma^2, q, \tau^2, X_1, \dots, X_{t-1}, X_{t+1}, \dots, X_{n-1}, Y_1, \dots, Y_n) \\ & \propto p(X_t|X_{t-1}, K, r, \sigma^2) p(Y_t|X_t, q, \tau^2) \\ & \propto \exp \left(- \frac{\{X_t - [X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}]\}^2}{2\sigma^2} - \frac{(Y_t - qX_t)^2}{2\tau^2} \right) \end{aligned}$$

b) Densidade condicional completa de K

$$\begin{aligned} p(K|r, \sigma^2, q, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) & \propto p(K, r, \sigma^2, q, \tau^2) \prod_{t=2}^n p(X_t|X_{t-1}, K, r, \sigma^2) \\ & \propto p(K, r, \sigma^2, q, \tau^2) \exp \left(\frac{-1}{2\sigma^2} \sum_{t=2}^n \{X_t - [X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}]\}^2 \right) \end{aligned}$$

c) Densidade condicional completa de r

$$\begin{aligned} p(r|K, \sigma^2, q, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) & \propto p(K, r, \sigma^2, q, \tau^2) \prod_{t=2}^n p(X_t|X_{t-1}, K, r, \sigma^2) \\ & \propto p(K, r, \sigma^2, q, \tau^2) \exp \left(\frac{-1}{2\sigma^2} \sum_{t=2}^n \{X_t - [X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}]\}^2 \right) \end{aligned}$$

d) Densidade condicional completa de σ^2

$$\begin{aligned}
 & p(\sigma^2|K, r, q, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
 \propto & p(K, r, \sigma^2, q, \tau^2) p(X_1|\sigma^2) \prod_{t=2}^n p(X_t|X_{t-1}, K, r, \sigma^2) \\
 \propto & \frac{p(K, r, \sigma^2, q, \tau^2)}{(\sigma^2)^{n/2}} \exp \left\{ \frac{-1}{2\sigma^2} \left\{ X_1 + \sum_{t=2}^n \{X_t - [X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}]\}^2 \right\} \right\}
 \end{aligned}$$

e) Densidade condicional completa de q

$$\begin{aligned}
 & p(q|K, r, \sigma^2, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
 \propto & p(K, r, \sigma^2, q, \tau^2) \prod_{t=1}^n p(Y_t|X_t, q, \tau^2) \\
 \propto & \frac{p(K, r, \sigma^2, q, \tau^2)}{(\tau^2)^{n/2}} \exp \left[\frac{-1}{2\tau^2} \sum_{t=1}^n (Y_t - qX_t)^2 \right]
 \end{aligned}$$

f) Densidade condicional completa de τ^2

$$\begin{aligned}
 & p(\tau^2|K, r, \sigma^2, q, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
 \propto & p(K, r, \sigma^2, q, \tau^2) \prod_{t=1}^n p(Y_t|X_t, q, \tau^2) \\
 \propto & \frac{p(K, r, \sigma^2, q, \tau^2)}{(\tau^2)^{n/2}} \exp \left[\frac{-1}{2\tau^2} \sum_{t=1}^n (Y_t - qX_t)^2 \right]
 \end{aligned}$$

B.2.2 Modelo com erros de observação e proceso lognormais

a) Densidade condicional completa de X_t

Para $2 \leq t \leq n-1$ temos:

$$\begin{aligned}
 & p(X_t|K, r, \sigma^2, q, \tau^2, X_1, \dots, X_{t-1}, X_{t+1}, \dots, X_n, Y_1, \dots, Y_n) \\
 \propto & p(X_t|X_{t-1}, K, r, \sigma^2) p(Y_t|X_t, q, \tau^2) p(X_{t+1}|X_t, K, r, \sigma^2) \\
 \propto & \exp \left(- \frac{[\log(X_t) - \log\{X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}\}]^2}{2\sigma^2} \right. \\
 & \left. - \frac{\{\log(Y_t) - \log(qX_t)\}^2}{2\tau^2} - \frac{[\log(X_{t+1}) - \log\{X_t + rX_t(1 - X_t) - H_t\}]^2}{2\sigma^2} \right)
 \end{aligned}$$

Para $t = 1$ temos:

$$\begin{aligned}
& p(X_t|K, r, \sigma^2, q, \tau^2, X_2, \dots, X_n, Y_1, \dots, Y_n) \\
& \propto p(X_t|K, \sigma^2)p(Y_t|X_t, q, \tau^2)p(X_{t+1}|X_t, K, r, \sigma^2) \\
& \propto \exp\left(-\frac{[\log(X_t) - \log(K)]^2}{2\sigma^2} - \frac{\{\log(Y_t) - \log(qX_t)\}^2}{2\tau^2} - \frac{[\log(X_{t+1}) - \log\{X_t + rX_t(1 - X_t) - H_t\}]^2}{2\sigma^2}\right)
\end{aligned}$$

Para $t = n$ temos:

$$\begin{aligned}
& p(X_t|K, r, \sigma^2, q, \tau^2, X_1, \dots, X_{t-1}, X_{t+1}, \dots, X_{n-1}, Y_1, \dots, Y_n) \\
& \propto p(X_t|X_{t-1}, K, r, \sigma^2)p(Y_t|X_t, q, \tau^2) \\
& \propto \exp\left(-\frac{[\log(X_t) - \log\{X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}\}]^2}{2\sigma^2} - \frac{\{\log(Y_t) - \log(qX_t)\}^2}{2\tau^2}\right)
\end{aligned}$$

b) Densidade condicional completa de K

$$\begin{aligned}
& p(K|r, \sigma^2, q, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
& \propto p(K, r, \sigma^2, q, \tau^2) \prod_{t=2}^n p(X_t|X_{t-1}, K, r, \sigma^2) \\
& \propto p(K, r, \sigma^2, q, \tau^2) \exp\left(\frac{-1}{2\sigma^2} \sum_{t=2}^n \left[\log(X_t) - \log\{X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}\}\right]^2\right)
\end{aligned}$$

c) Densidade condicional completa de r

$$\begin{aligned}
& p(r|K, \sigma^2, q, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
& \propto p(K, r, \sigma^2, q, \tau^2) \prod_{t=2}^n p(X_t|X_{t-1}, K, r, \sigma^2) \\
& \propto p(K, r, \sigma^2, q, \tau^2) \exp\left(\frac{-1}{2\sigma^2} \sum_{t=2}^n \left[\log(X_t) - \log\{X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1}\}\right]^2\right)
\end{aligned}$$

d) Densidade condicional completa de σ^2

$$\begin{aligned}
& p(\sigma^2 | K, r, q, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
& \propto p(K, r, \sigma^2, q, \tau^2) p(X_1 | \sigma^2) \prod_{t=2}^n p(X_t | X_{t-1}, K, r, \sigma^2) \\
& \propto \frac{p(K, r, \sigma^2, q, \tau^2)}{(\sigma^2)^{n/2}} \exp \left\{ \frac{-1}{2\sigma^2} \left(\log(X_1) \right. \right. \\
& \quad \left. \left. + \log \sum_{t=2}^n \left[\log(X_t) - \log \left\{ X_{t-1} + rX_{t-1}(1 - X_{t-1}) - H_{t-1} \right\} \right]^2 \right) \right\}
\end{aligned}$$

e) Densidade condicional completa de q

$$\begin{aligned}
& p(q | K, r, \sigma^2, \tau^2, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
& \propto p(K, r, \sigma^2, q, \tau^2) \prod_{t=1}^n p(Y_t | X_t, q, \tau^2) \\
& \propto \frac{p(K, r, \sigma^2, q, \tau^2)}{(\tau^2)^{n/2}} \exp \left[\frac{-1}{2\tau^2} \sum_{t=1}^n \left\{ \log(Y_t) - \log(q) - \log(X_t) \right\}^2 \right]
\end{aligned}$$

f) Densidade condicional completa de τ^2

$$\begin{aligned}
& p(\tau^2 | K, r, \sigma^2, q, X_1, \dots, X_n, Y_1, \dots, Y_n) \\
& \propto p(K, r, \sigma^2, q, \tau^2) \prod_{t=1}^n p(Y_t | X_t, q, \tau^2) \\
& \propto \frac{p(K, r, \sigma^2, q, \tau^2)}{(\tau^2)^{n/2}} \exp \left[\frac{-1}{2\tau^2} \sum_{t=1}^n \left\{ \log(Y_t) - \log(q) - \log(X_t) \right\}^2 \right]
\end{aligned}$$

B.3 Trazas das cadeidas de Markov

Figura B.1: *Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dado de camarão da zona central do Chile, com erros de observação log-skew-normal (a).*

Figura B.2: *Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dados de camarão da zona central do Chile, com erros de observação log-skew-normal (b).*

Figura B.3: Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dado de camarão da zona central do Chile, com erros de observação log-skew- t (a).

Figura B.4: *Iterações do algoritmo de amostragem MCMC para os parâmetros do modelo logístico aplicado aos dado de camarão da zona central do Chile, com erros de observação log-skew-t (b).*