



Universidade Federal do Rio de Janeiro
Instituto de Matemática - CCMN

Pós-graduação em Estatística

**FUNÇÕES DE TRANSFERÊNCIA EM
MODELOS DINÂMICOS LINEARES
GENERALIZADOS**

Mariane Branco Alves

TESE DE DOUTORADO

Rio de Janeiro
junho de 2006

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Universidade Federal do Rio de Janeiro
Instituto de Matemática - CCMN

Mariane Branco Alves

**FUNÇÕES DE TRANSFERÊNCIA EM MODELOS
DINÂMICOS LINEARES GENERALIZADOS**

*Trabalho apresentado ao Programa de Pós-graduação em
Estatística do Instituto de Matemática - CCMN da Univer-
sidade Federal do Rio de Janeiro como requisito parcial
para obtenção do grau de Doutor em Estatística.*

Orientador: *Dani Gamerman*
Co-orientador: *Marco Antonio Rosa Ferreira*

Rio de Janeiro
junho de 2006

FUNÇÕES DE TRANSFERÊNCIA EM MODELOS DINÂMICOS LINEARES
GENERALIZADOS

Mariane Branco Alves

TRABALHO APRESENTADO AO PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA DO INSTITUTO DE MATEMÁTICA - CCMN DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO REQUISITO PARCIAL PARA OBTENÇÃO DO GRAU DE DOUTOR EM ESTATÍSTICA.

Aprovado por:

Prof. Dani Gamerman, Ph.D.

Prof. Marco Antonio Rosa Ferreira, Ph.D.

Profa. Alexandra Mello Schmidt, Ph.D.

Prof. Helio dos Santos Migon, Ph.D.

Prof. Reinaldo Castro de Souza, Ph.D.

Prof. Silvia Emiko Shimakura , Ph.D.

RIO DE JANEIRO, RJ - BRASIL
JUNHO DE 2006

ALVES, MARIANE BRANCO

Funções de Transferência
em Modelos Dinâmicos Lineares
Generalizados [Rio de Janeiro] 2006
VIII, 152p. 29,7 cm (IM/UFRJ, D.Sc.,
Estatística, 2006)

Tese - Universidade Federal do Rio
de Janeiro, IM

1.Funções de Transferência 2.Modelos
Dinâmicos Lineares Generalizados
3.Inferência Bayesiana 4.Monte Carlo
via Cadeais de Markov

I.IM/UFRJ II.Título (série)

*Há muito tempo que eu saí de casa
Há muito tempo que eu caí na estrada
Há muito tempo que eu estou na vida
Foi assim que eu quis e assim eu sou feliz*

*Principalmente por poder voltar
A todos os lugares onde já cheguei
Pois lá deixei um prato de comida
Um abraço amigo, um canto para dormir e sonhar*

*E aprendi que se depende sempre
De tanta, muita, diferente gente
Toda pessoa sempre é as marcas
Das lições diárias de outras tantas pessoas*

*E é tão bonito quando a gente entende
Que a gente é tanta gente onde quer que a gente vá
E é tão bonito quando a gente sente
Que nunca está sozinho por mais que pense estar*

*É tão bonito quando a gente pisa firme
Nessas linhas que estão nas palmas de nossas mãos
É tão bonito quando a gente vai à vida
Nos caminhos onde bate bem mais forte o coração*

Caminhos do Coração - Gonzaguinha

*A meus pais, que me deram raízes, mas também
souberam me dar asas; que me ensinaram o essencial:
amor, ética, dignidade. O restante eu sigo aprendendo
nos livros e na vida.*

*Ao Marcus, que conhece meus defeitos e ressalta
minhas virtudes; com quem eu sou, simplesmente, eu.
Com quem eu sou feliz.*

Agradecimentos

Ao meu orientador Dani Gamerman, minha admiração pela habilidade de aliar talento a organização. Obrigada pela paciência e dedicação ao longo deste trabalho. Foi inevitável, depois de todos esses anos, que eu passasse a lhe considerar um grande amigo e é ao *amigo* Dani, pessoa extremamente generosa, que eu quero registrar minha enorme gratidão por ter viabilizado meu retorno à vida acadêmica.

A meu orientador Marco Antonio Rosa Ferreira, pelas contribuições sempre relevantes que trouxe à orientação deste trabalho e por ter me cedido seu programa para ajuste de modelos dinâmicos lineares generalizados para o caso Poisson, o que certamente facilitou a implementação das extensões feitas neste trabalho.

Ao Laboratório de Poluição Atmosférica Experimental da Faculdade de Medicina da USP, pelos dados de poluição e óbitos em São Paulo e ao professor Julio da Motta Singer (IME-USP), pela disponibilização desses dados em sua página na internet.

À Secretaria Municipal de Meio Ambiente do Rio de Janeiro (SMAC-RJ), pelos dados de poluição utilizados em uma das aplicações deste trabalho e ao professor Antonio Carlos Ponce de Leon (IMS-UERJ), pela parceria no projeto de extensão Poluição versus Saúde, que viabilizou o uso desses dados.

A Washington Leite Junger, pela disponibilização de seu aplicativo para preenchimento de dados omissos.

A meus irmãos queridos, Magno e Fabiane, que não me faltam nunca. Qualquer agradecimento a vocês é pouco. Talvez eu deva, simplesmente, agradecer à vida, por vocês existirem e por se fazerem sempre presentes.

Ao Vitor, que se tornou de vez um membro da minha família durante o doutorado, pela dedicação à minha irmã.

Às queridas amigas Alcione e Alexandra. A primeira, uma pernambucana "da gema", com o dom de espalhar risos por onde passa: sua companhia no início dessa trajetória fez toda a diferença. A segunda, uma friburguense de gargalhada praiana, solar: nunca vou esquecer a força que você deu para o término dessa dissertação. Mas agradecer pela participação de vocês ao longo do doutorado é diminuir sua importância, porque amigas como vocês são a família que eu pude escolher.

À Cibele, mineira arretada (ô sô!), e com um coração enorme. Até te conhecer, eu me sentia completamente só no doutorado. Obrigada por ter dividido comigo essa jornada.

Ao professor e amigo Hedibert, que por trás de um temperamento explosivo guarda um coração tão grande quanto ele próprio e que sempre esteve pronto a conversar: fosse para resolver um problema de Estatística ou para falar dos caminhos a que a vida nos leva.

Felizmente, meus agradecimentos não cabem em uma página...

Ao professor Hélio Migon: suas aulas, no final da graduação, no mestrado e no doutorado, foram um divisor de águas para minha formação. Minha gratidão por ter me incentivado, há alguns anos atrás, a tentar um concurso para professora substituta no Departamento de Estatística. Meu retorno à Estatística começou com aquele concurso.

Ao Prof. Miguel de Simone (*in memoriam*), que espreitou uma de minhas aulas de professora substituta e inexperiente e veio me dizer, ao final, que eu "levava jeito para a coisa" e que vocação é coisa séria, à qual não se deve dar as costas. Pensei muito nessas palavras ao voltar à universidade.

A todos os professores que tive na UFRJ, pela formação sólida que me proporcionaram. Em particular, ao professor Guido Zapata, que se colocava em sala de aula com o cérebro e com a alma; que reconhecia em cada um de nós um indivíduo e que me ensinou, além de Análise Real e Teoria da Integração, a olhar os alunos nos olhos.

À Lia (na secretaria de pós-graduação) e à D. Margarida (na biblioteca), ambas agora ex-funcionárias e, mais recentemente, ao Eduardo (da secretaria de pós-graduação) que sempre me atenderam com enorme boa vontade.

Aos meus alunos da UERJ, que me têm trazido muitas alegrias e que, para mim, são um incentivo constante para aprender mais.

Finalmente, à própria UFRJ. Com o término do meu doutoramento, encerra-se um longo período na minha vida: fui aluna do "Fundão" de 1991 a 2006, com apenas uma breve interrupção, quando decidi que precisava experimentar a vida profissional fora das fronteiras acadêmicas, para em breve retornar, cheia de saudades. Comecei minha trajetória por essa universidade no IM, passei pela COPPE e, finalmente, retornei ao IM. Desnecessário dizer que, tendo feito graduação, mestrado e doutorado na instituição, a profissional que sou é fruto direto da formação que tive aqui. Também por isso, tenho enorme gratidão à UFRJ. Mas minha gratidão vai muito além do profissionalismo. A *pessoa* que sou tem muito a ver com a minha passagem por aqui: o Fundão permeia esses anos da minha vida. Aqui fiz grandes amigos. Na UFRJ, dei meus primeiros passos nos tablados das salas de aula universitárias e acabei constatando que esses tablados são o meu lugar. Aqui aprendi a estudar de verdade, a trabalhar numa idéia até senti-la íntima. E hoje a memória passeia e é sexta-feira de Grêmio - onde ouvir música brega na jukebox, além de não ser brega, era *cool* - e num instante é Quinta Musical no auditório do CT - dia de conhecer música diferente, feita pelo Brasil afora - e já, já é Quarta de cinema no Roxinho. E num piscar de olhos me pego aos pulos num dos shows inesquecíveis no bloco A, quando a gente ("a gente", então, éramos: eu, Elaine, Raquel e Rogério) decidia que, definitivamente, haveria oportunidade para fazer prova outro dia, mas show grátis e com quase uma hora de bis era imperdível! Ou então nos reuníamos para ir à passeata, vestidos de preto, com as caras pintadas de verde e amarelo. Saudades dos saraus lítero-musicais na Produção (onde eu descobri que não era a única a ouvir aqueles discos que ninguém conhecia) e cafezinhos no Anjinho e lanches no Batista - estes sempre regados a um bom papo, meu esporte preferido. E as horas intermináveis de biblioteca e provas... quantas provas! E, principalmente, quanta gente! Conheci mineiro, paulista, candango, potiguar, amazonense, cearense, matogrossense, gaúcho e outros fluminenses (e como são, em toda a sua diversidade, parecidas essas gentes!). Conheci pernambucano, baiano, alagoano, peruano, colombiano, equatoriano... tantos anos... anos que eu levo na memória e no coração. Valeu, Fundão!

Resumo

Modelos dinâmicos lineares constituem uma poderosa estrutura para a análise de séries temporais, uma vez que, ao permitir a evolução de seus parâmetros, explicitamente determinam a forma como estes parâmetros relacionam-se a seus valores passados. Uma hipótese bastante restritiva nessa classe de modelos é a de que a variável resposta deva ser gaussiana. Os modelos dinâmicos lineares generalizados estendem a classe dos modelos dinâmicos lineares, permitindo respostas cujas distribuições pertençam à família exponencial. Essa classe de modelos também pode ser vista como uma extensão aos modelos lineares generalizados, estruturando formalmente a autocorrelação serial. Uma outra vantagem dos modelos dinâmicos lineares generalizados é a construção de preditores formados por blocos estruturais com interpretação clara, tais como tendência, sazonalidade e efeitos de regressoras. Ao se especificar, em uma análise temporal, o efeito de uma regressora, é possível assumir-se que variações nessa variável não tenham impacto apenas imediato sobre a resposta esperada, mas que seu efeito se propague, de alguma forma, por instantes futuros. Adotamos, no presente trabalho, funções de transferência para modelagem de tais efeitos, representados por blocos estruturais E_t presentes nos modelos dinâmicos lineares generalizados propostos. Toda a inferência é realizada sob o paradigma bayesiano e, no contexto acima, tem-se duas fontes de dificuldades para a obtenção analítica de distribuições a posteriori: a natureza não gaussiana da resposta, associada a prioris não conjugadas e, caso haja parâmetros autoregressivos no bloco E_t , a não-linearidade do preditor em tais parâmetros. Durante a década de 80, diversos trabalhos trataram a não-normalidade / não-linearidade em modelos dinâmicos baseando o processo de inferência em primeiros e segundos momentos ou evitando integração por meio da obtenção de modas a posteriori. O propósito deste trabalho é a realização de inferência bayesiana completa sobre modelos dinâmicos lineares generalizados contendo funções de transferência em seus preditores, utilizando-se métodos de Monte Carlo via Cadeias de Markov para obtenção de amostras da distribuição a posteriori conjunta dos parâmetros envolvidos em tais modelos. Diversas estruturas de transferência são especificadas, associadas a respostas Poisson, Binomial, Gama e Gaussiana Inversa. Os modelos resultantes são aplicados a dados artificiais e, finalmente, os modelos Poisson e Gama são aplicados a dados reais. No primeiro caso, estima-se o efeito acumulado de poluentes atmosféricos sobre contagens de óbitos de crianças. O modelo Gama é usado para se estimar o efeito de volumes diários de chuva sobre níveis de poluição.

Palavras-chave: Funções de transferência, Modelos dinâmicos lineares generalizados, Inferência bayesiana, Monte Carlo via cadeias de Markov.

Abstract

Dynamic linear models constitute a powerful structure for the analysis of time series, once they allow parameters to evolve, formally determining the way these parameters relate to their past values. A sufficiently restrictive hypothesis in this class of models is that the dependent variable must be Gaussian. Dynamic generalized linear models extend the class of dynamic linear models, allowing responses whose distributions belong to the exponential family. This class of models can also be seen as an extension to the generalized linear models, formally structuring serial autocorrelation. Another advantage of the dynamic generalized linear models is the construction of predictors formed by structural blocks with clear interpretation, such as trend, seasonality and effects of regressor variables. As one specifies the effect of a regressor in a time series analysis it is possible to assume that variations in that variable do not have only immediate impact on the mean response, but that its effects somehow propagate to future times. We adopt, in the present work, transfer functions to model such effects, represented by structural blocks E_t considered in the proposed dynamic generalized linear models. All the inference is carried under the Bayesian paradigm and in the context above two sources of difficulties emerge for the analytical derivation of posterior distributions: non-gaussian nature of the response, associated to non-conjugate priors and, if there are autoregressive parameters in the block E_t , also non-linearity of the predictor on these parameters. During the 80's several works treated non-normality / non-linearity in dynamic models basing the inference process on first and second moments or avoiding integration by the determination of posterior modes. The purpose of this work is to produce complete Bayesian inference on generalized dynamic linear models with transfer functions in their predictors, using Monte Carlo Markov Chain methods to build samples of the posterior joint distribution of the parameters involved in such models. Several transfer structures are specified, associated to Poisson, Binomial, Gamma and Inverse Gaussian responses. The resulting models are applied to artificial data and, finally, Poisson and Gamma models are applied to real data. In the first case, we estimate the cumulative effect of atmospheric pollutants on infant death counts. The Gamma model is used to estimate the effect of daily rain volumes over pollutant levels.

Keywords: Transfer functions, Dynamic Generalized Linear Models, Bayesian inference, Monte Carlo Markov chain.

Sumário

1	Introdução	1
1.1	Contexto e Relevância do Trabalho	1
1.2	Organização do Trabalho	4
2	Revisão de Literatura	7
2.1	Modelos Dinâmicos Lineares	8
2.2	Modelos Dinâmicos Não-normais / Não-lineares	10
2.2.1	Modelos dinâmicos lineares generalizados	10
2.2.2	Aproximação da posteriori em modelos dinâmicos não-normais/ não-lineares: algumas alternativas	11
2.3	Funções de Transferência	13
2.4	Monte Carlo via Cadeias de Markov	15
2.4.1	O amostrador de Gibbs	17
2.4.2	O algoritmo Metropolis-Hastings	18
2.4.3	Estratégia de amostragem: movimentos individuais ou em blocos	20
2.4.4	Monitoração da convergência	20
2.5	Seleção de Modelos	21
2.5.1	EPD - <i>Expected predictive deviance</i>	22
2.5.2	DIC - <i>Deviance information criterion</i>	23
2.5.3	Verossimilhança preditiva	24
3	Funções de Transferência em Modelos Dinâmicos Lineares Generalizados	25
3.1	A Estrutura Geral dos Modelos Propostos	25
3.2	Estrutura Observacional	28
3.3	Funções de Transferência Adotadas	30
3.3.1	Modelos de defasagens polinomiais	30
3.3.2	Função de transferência de ordem ($r = 1, s = 0$)	32
3.3.3	Função de transferência de ordem ($r = 1, s > 0$)	34
3.3.4	Função de transferência de ordem ($r=1,s=0$) com fator de ganho dinâmico	34
3.3.5	Função de transferência de ordem ($r = 1, s \geq 0$) com erros aleatórios	35
3.4	Cálculos Envolvidos nos algoritmos MCMC	36
3.4.1	Resposta Poisson	37
3.4.2	Resposta Binomial	48
3.4.3	Resposta Gama	49
3.4.4	Resposta Gaussiana Inversa	50

3.5	Utilização das Amostras da Posteriori para Estimação e Predição	51
4	Análise de Dados Artificiais	53
4.1	Resposta Poisson	53
4.1.1	Função de transferência de ordem ($r = 1, s = 0$)	53
4.1.2	Função de transferência de ordem ($r = 1, s > 0$)	61
4.1.3	Função de transferência de ordem ($r = 1, s = 0$) com erro aleatório	69
4.1.4	Função de transferência de ordem ($r = 1, s = 0$) e nível dinâmico	72
4.1.5	Função de transferência de ordem ($r = 1, s = 0$) com fator de ganho dinâmico	77
4.2	Resposta Binomial	82
4.2.1	Função de transferência de ordem ($r = 1, s = 0$)	82
4.2.2	Função de transferência de ordem ($r = 1, s = 1$)	85
4.2.3	Função de transferência de ordem ($r = 1, s = 0$) com erro aleatório	88
4.2.4	Função de transferência de ordem ($r = 1, s = 0$) e nível dinâmico	91
4.3	Resposta Gama	94
4.3.1	Função de transferência de ordem ($r = 1, s = 0$)	94
4.3.2	Função de transferência de ordem ($r = 1, s > 0$)	96
4.4	Resposta Gaussiana Inversa	101
4.4.1	Função de transferência de ordem ($r = 1, s = 0$)	101
4.4.2	Função de transferência de ordem ($r = 1, s = 1$)	104
4.5	Alguns Comentários sobre os Exercícios com Dados Simulados	107
5	Aplicações a Dados Reais	109
5.1	Efeito de Monóxido de Carbono sobre Óbitos de Crianças em São Paulo	110
5.2	Efeito de Chuva sobre Níveis de Material Particulado no Rio de Janeiro	126
5.3	Conclusões	138
6	Conclusões e Trabalhos Futuros	139
6.1	Comentários Gerais	139
6.1.1	Aspectos Computacionais	139
6.1.2	Identificação dos Modelos Propostos	140
6.1.3	Alguns aspectos sobre as aplicações desenvolvidas	142
6.2	Tópicos para Trabalho Futuro	143
6.2.1	Modelagem de efeitos que apresentem pontos de inflexão	143
6.2.2	Modelagem espaço-temporal	144
6.2.3	Tratamento do efeito acumulado de mais que uma regressora	145

Lista de Figuras

2.1	<i>Exemplos de funções de resposta ao impulso e de transferência, para diferentes combinações de ordens $r = 0, 1, 2$ e $s = 0, 1, 2$.</i>	16
3.1	<i>Aproximação polinomial para os coeficientes β_j de um modelo de defasagens distribuídas.</i>	31
3.2	<i>Modelo de primeira ordem: Comportamento da função de resposta ao impulso (a) e da função de transferência (b), para $\rho < 1$ e $\rho > 1$, com $\beta > 0$.</i>	33
3.3	<i>Função de resposta ao impulso com fatores de ganho evoluindo segundo um passeio aleatório, simulada com $\rho = 0.9$ e $Q = 0.005$, nos instantes 100, 300 e 500.</i>	35
3.4	<i>Função de resposta ao impulso com fatores de ganho estocásticos, simulada com $\rho = 0.9$, $\bar{\beta} = 0.03$ e $Q = 0.005$, nos instantes 100, 300 e 500.</i>	36
4.1	<i>Regressora X_t e covariáveis Z_{1t} e Z_{2t} utilizadas nas simulações para respostas Poisson e Binomial.</i>	54
4.2	<i>Resposta Poisson: Traços das cadeias geradas e histogramas da amostra da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).</i>	55
4.3	<i>Resposta Poisson, modelo de ordem ($r=1, s=0$): Função de resposta ao impulso e de transferência - teóricas (linha sólida), médias e limites de credibilidade estimados a 95% (linhas tracejadas).</i>	56
4.4	<i>Resposta Poisson, modelo de ordem ($r=1, s=0$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.</i>	57
4.5	<i>Resposta Poisson, modelo de ordem ($r=1, s=0$): Histogramas das amostras das distribuições preditivas para horizontes variando de 1 a 30.</i>	58
4.6	<i>Resposta Poisson: Histogramas das amostras das posteriores de ρ, para dados simulados segundo o modelo de ordem ($r=1, s=0$), ajustado para $T=100, 300, 500$ e 700.</i>	59
4.7	<i>Resposta Poisson, modelo de ordem ($r=1, s=0$) com $\rho = 0,9$ e $\beta = 0,03$: (a) Função de resposta ao impulso; (b) Função de resposta média. Funções teóricas (linha em negrito), estimativas das médias a posteriori e de limites de credibilidade obtidos do ajuste dos modelos de ordem ($r=1, s=0$) e polinomial de grau 3.</i>	59
4.8	<i>Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de função de transferência de ordem ($r = 1, s = 1$).</i>	62

- 4.9 *Resposta Poisson, modelo de ordem ($r=1,s=1$): Função de resposta ao impulso e de transferência - teóricas (linha sólida), médias e limites de credibilidade estimados a 95% (linhas tracejadas).* 63
- 4.10 *Resposta Poisson, modelo de ordem ($r=1,s=1$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 64
- 4.11 *Resposta Poisson, modelo de ordem ($r=1,s=1$): Amostra das distribuições preditivas para horizontes variando de 1 a 30.* 65
- 4.12 *Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de ordem ($r=1,s=2$).* 66
- 4.13 *Resposta Poisson: verossimilhança para ρ e β_2 , para dados simulados de acordo com o modelo de ordem ($r=1,s=2$).* 68
- 4.14 *Resposta Poisson, modelo de ordem ($r=1,s=2$) com $\rho = 0,9$, $\beta_0 = 1,0$, $\beta_1 = 0,06$ e $\beta_2 = 0,03$: (a) Função de resposta ao impulso; (b) Função de resposta média. Funções teóricas (linha em negrito), estimativas das médias a posteriori e de limites de credibilidade obtidos do ajuste dos modelos de ordem ($r=1,s=2$) e polinomial de grau 3.* 69
- 4.15 *Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem ($r=1,s=0$) com erro aleatório.* 70
- 4.16 *Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem ($r=1,s=0$) com nível dinâmico .* 73
- 4.17 *Resposta Poisson, modelo de ordem ($r=1,s=0$) com nível dinâmico. Nível α_t : (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços de α_t teórico (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 74
- 4.18 *Resposta Poisson, modelo de ordem ($r=1,s=0$) com nível dinâmico: Função de resposta ao impulso e de transferência - teóricas (linha sólida), médias e limites de credibilidade estimados a 95% (linhas tracejadas).* 75
- 4.19 *Resposta Poisson, modelo de ordem ($r=1,s=0$) com nível dinâmico. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 75
- 4.20 *Resposta Poisson, modelo de ordem ($r=1,s=0$) com nível dinâmico : Histogramas das amostras das distribuições preditivas para horizontes variando de 1 a 30.* 76
- 4.21 *Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem ($r=1,s=0$) com fator de ganho dinâmico.* 78
- 4.22 *Resposta Poisson, modelo de ordem ($r=1,s=0$) com nível dinâmico. Fator de ganho β_t : (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços de β_t teórico (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 79

- 4.23 *Resposta Poisson, modelo de ordem ($r=1, s=0$) com fator de ganho dinâmico. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 80
- 4.24 *Resposta Poisson, modelo de ordem ($r=1, s=0$) com fator de ganho dinâmico, função de resposta ao impulso nos instantes $t = 124$, $t = 541$ e $t = 890$: função teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 80
- 4.25 *Resposta Poisson, modelo de ordem ($r=1, s=0$) com fator de ganho dinâmico: Histogramas das amostras das distribuições preditivas para horizontes variando de 1 a 30.* 81
- 4.26 *Resposta Binomial: Traços das cadeias geradas e histogramas da amostra da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).* 83
- 4.27 *Resposta Binomial, modelo de ordem ($r=1, s=0$) função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 84
- 4.28 *Resposta Binomial: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de função de transferência de ordem $r = 1, s = 1$.* 86
- 4.29 *Resposta Binomial, modelo de ordem ($r=1, s=1$) função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 87
- 4.30 *Resposta Binomial, modelo de ordem ($r=1, s=1$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 87
- 4.31 *Resposta Binomial, modelo de ordem ($r=1, s=0$) com erro aleatório. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 88
- 4.32 *Resposta Binomial($50, p_t$): Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de função de transferência de ordem ($r=1, s=0$) com erro aleatório.* 89
- 4.33 *Resposta Binomial($50, p_t$), modelo de ordem ($r=1, s=0$) com erro aleatório: função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 90
- 4.34 *Resposta Binomial ($50, p_t$), modelo de ordem ($r=1, s=0$) com erros aleatórios. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 91
- 4.35 *Resposta Binomial($5, p_t$): Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem ($r=1, s=0$) com nível dinâmico.* 92
- 4.36 *Resposta Binomial($5, p_t$), modelo de ordem ($r=1, s=0$) com nível dinâmico: função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95%.* 93

- 4.37 *Resposta Binomial(5, p_t), modelo de ordem ($r=1, s=0$) com nível dinâmico. Nível α_t : (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços de α_t teórico (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %. Função de resposta média: (c) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (d) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 93
- 4.38 *Regressora X_t e covariável Z_t , utilizadas nas simulações para respostas Gama e Gaussiana Inversa.* 94
- 4.39 *Resposta Gama: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).* 95
- 4.41 *Resposta Gama, modelo de ordem ($r=1, s=1$): Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).* 98
- 4.42 *Resposta Gama: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=2$).* 99
- 4.43 *Resposta Gama, modelo de ordem ($r=1, s=2$): Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).* 100
- 4.44 *Resposta Gama, modelo de ordem ($r=1, s=2$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 101
- 4.45 *Resposta Gaussiana Inversa: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).* 102
- 4.46 *Resposta Gaussiana Inversa, modelo de ordem ($r=1, s=0$): Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).* 103
- 4.47 *Resposta Gaussiana Inversa, modelo de ordem ($r=1, s=0$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 104
- 4.48 *Resposta Gaussiana Inversa: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=1$).* 104
- 4.49 *Resposta Gaussiana Inversa, modelo de ordem ($r=1, s=1$): Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).* 106
- 4.50 *Resposta Gaussiana Inversa, modelo de ordem ($r=1, s=1$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.* 106
- 5.1 *Níveis diários de CO (ppm) em SP - 01/Jan/1994 a 31/Dez/1997.* 111

5.2	<i>Contagem de óbitos por doenças respiratórias de menores de 5 anos em SP - 01/Jan/1994 a 31/Dez/1997.</i>	111
5.3	<i>Registros diários de temperatura mínima e umidade média em SP - 01/Jan/1994 a 31/Dez/1997.</i>	112
5.4	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 5: Histogramas das amostras da posteriori.</i>	114
5.5	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 5: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível do poluente.</i>	115
5.6	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 5: Evolução dos riscos relativos associados aos níveis máximos de CO observados a cada ano, quando comparados ao nível médio observado no período de análise.</i>	116
5.7	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 5: Contagens diárias de óbitos (linha tracejada) e função de resposta média, λ_t, estimada, com intervalos de credibilidade a 90%.</i>	116
5.8	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 5: Histogramas das amostras das distribuições preditivas para o número de óbitos, com horizontes variando de 1 a 30 dias.</i>	117
5.9	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 9: Média e intervalos de credibilidade a 90% para β_t.</i>	118
5.10	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 10: Histogramas das amostras da posteriori.</i>	119
5.11	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 10: Média e intervalos de credibilidade a 90% para β_t.</i>	120
5.12	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 10: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível do poluente, estimadas em quatro instantes - estimativas da média a posteriori e de limites de credibilidade a 90%.</i>	121
5.13	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 10: Evolução dos riscos relativos associados a 12 ppm de CO, comparado ao nível médio observado no período de análise, em diferentes instantes.</i>	122
5.14	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 10: Contagens diárias de óbitos (linha tracejada) e função de resposta média, λ_t, estimada, com intervalos de credibilidade a 90%.</i>	123
5.15	<i>Efeito de CO sobre óbitos de crianças em SP - modelo 5: Histogramas das amostras das distribuições preditivas para o número de óbitos, com horizontes variando de 1 a 30.</i>	124
5.16	<i>Níveis diários de PM_{10} ($\mu g/m^3$) no RJ - 01/Set/2000 a 31/Ago/2002.</i>	127
5.17	<i>Variáveis meteorológicas no RJ - 01/Set/2000 a 31/Ago/2002.</i>	128
5.18	<i>Efeito de chuva sobre PM_{10} no RJ - modelo 3: Histogramas das amostras da posteriori.</i>	131
5.19	<i>Efeito de chuva sobre PM_{10} no RJ - modelo 3: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível de chuva.</i>	132

- 5.20 *Efeito de chuva sobre PM_{10} no RJ - modelo 3: Evolução do risco relativo associado ao nível máximo de chuva observado no período de análise, quando comparado a dias sem chuva.* 132
- 5.21 *Efeito de CO sobre óbitos de crianças em SP - modelo 3: Histogramas das amostras das distribuições preditivas para o número de óbitos, com horizontes variando de 1 a 30.* 133
- 5.22 *Efeito de chuva sobre PM_{10} no RJ - modelo 8: Histogramas das amostras da posteriori.* 135
- 5.23 *Efeito de chuva sobre PM_{10} no RJ - modelo 7: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível de chuva.* 136
- 5.24 *Efeito de chuva sobre PM_{10} no RJ - modelo 7: Evolução do risco relativo associado ao nível máximo de chuva observado no período de análise, quando comparado a dias sem chuva.* 136
- 5.25 *Efeito de CO sobre óbitos de crianças em SP - modelo 7: Histogramas das amostras das distribuições preditivas para o número de óbitos, com horizontes variando de 1 a 30.* 137
- 6.1 *Estimação de um modelo contendo duas funções de transferência: Histogramas das amostras da posteriori.* 146

Lista de Tabelas

4.1	<i>Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem $(r = 1, s = 0)$, para dados simulados com $\rho = 0,9$.</i>	56
4.2	<i>Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r = 1, s = 0)$, para dados simulados com $\rho = 0,9$.</i>	56
4.3	<i>Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem $(r = 1, s = 0)$, para dados simulados com $\rho = 0,99$.</i>	60
4.4	<i>Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r = 1, s = 0)$, para dados simulados com $\rho = 0,99$.</i>	60
4.5	<i>Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $r=1, s=1$.</i>	63
4.6	<i>Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1, s=1)$.</i>	63
4.7	<i>Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $(r=1, s=2)$.</i>	67
4.8	<i>Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1, s=2)$.</i>	68
4.9	<i>Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $(r=1, s=0)$ com erro aleatório.</i>	71
4.10	<i>Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1, s=0)$ com erro aleatório.</i>	71
4.11	<i>Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $(r=1, s=0)$ com nível dinâmico.</i>	72
4.12	<i>Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1, s=0)$ com nível dinâmico.</i>	74
4.13	<i>Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $(r=1, s=0)$ com fator de ganho dinâmico.</i>	77
4.14	<i>Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1, s=0)$ com fator de ganho dinâmico.</i>	79
4.15	<i>Resposta Binomial: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem $(r = 1, s = 0)$.</i>	84
4.16	<i>Resposta Binomial: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r = 1, s = 0)$.</i>	84
4.17	<i>Resposta Binomial: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $(r=1, s=1)$.</i>	85

4.18	<i>Resposta Binomial: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1,s=1)$.</i>	87
4.19	<i>Resposta Binomial: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $(r=1,s=0)$ com erro aleatório.</i>	88
4.20	<i>Resposta Gama: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1,s=0)$ para dados simulados.</i>	96
4.21	<i>Resposta Gama: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1,s=0)$ para dados simulados.</i>	96
4.22	<i>Resposta Gama: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $r=1,s=2$.</i>	98
4.23	<i>Resposta Gama: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $r=1,s=2$.</i>	99
5.1	<i>Comparação dos modelos ajustados para órbitos de crianças, com avaliação do impacto cumulativo do poluente Monóxido de Carbono.</i>	113
5.2	<i>Efeito de CO sobre órbitos de crianças em SP - modelo 5: Sumário das amostras da posteriori.</i>	115
5.2	<i>Efeito de CO sobre órbitos de crianças em SP - modelo 10: Sumário das amostras da posteriori.</i>	120
5.3	<i>Comparação dos modelos ajustados para nível de PM_{10}, com avaliação do impacto cumulativo do volume de chuva.</i>	129
5.4	<i>Efeito de chuva sobre PM_{10} no RJ - modelo 3: Sumário das amostras da posteriori.</i>	130
5.5	<i>Efeito de chuva sobre PM_{10} no RJ - modelo 3: Sumário das amostras da posteriori.</i>	134

Introdução

1.1 Contexto e Relevância do Trabalho

Diversas são as aplicações em que se objetiva avaliar o impacto de um conjunto de regressoras sobre determinada variável resposta, levando-se em conta que todas as variáveis em questão são observadas temporalmente e considerando-se que variações em alguma dessas regressoras podem ter um impacto sobre a resposta que não se faz sentir apenas de forma imediata, propagando-se por determinado período. Denote-se por Y_t o processo de saída e, por X_t , o valor, no instante t , da regressora cujo impacto cumulativo encontra-se sob análise. A forma mais trivial possível para a modelagem do efeito E_t , de X_t sobre Y_t , é obtida por meio de defasagens distribuídas:

$$E_t = \sum_{j=0}^{\infty} \beta_j X_{t-j}. \quad (1.1)$$

A adoção de tal forma, entretanto, esbarra em algumas dificuldades: primeiramente, a expressão à direita em (1.1) deve ser truncada, de forma a se garantir a estabilidade do filtro, obtendo-se:

$$E_t = \sum_{j=0}^s \beta_j X_{t-j}. \quad (1.2)$$

Ainda assim, tem-se dificuldades na estimação dos coeficientes β_0, \dots, β_s devido à autocorrelação inerente à regressora em suas diversas defasagens. Segundo Klein (1958), embora somas e outras funções dos parâmetros β_j possam ser estimadas com um grau de precisão razoável, a autocorrelação entre as regressoras compromete a estimação dos efeitos instantâneos representados por cada coeficiente. É usual, então, que se imponham restrições sobre estes coeficientes, reduzindo-se substancialmente o número de parâmetros a estimar e obtendo-se melhores estimativas do efeito propagado de X . Ainda de acordo com Klein (1958), um trabalho precursor neste sentido é o de Fisher (1937), em que é proposto o decaimento aritmético dos coeficientes β_j . O uso de defasagens distribuídas sujeitas a restrições ganha vulto, entretanto, com o trabalho de Koyck (1954), no qual se postula, a partir de b defasagens, o decaimento geométrico dos coeficientes β_j , em (1.1), os quais passam a ser escritos como:

$$\beta_j = \rho^j \beta, \quad (j \geq b, \quad 0 \leq \rho < 1). \quad (1.3)$$

Se $b = 0$, então o efeito cumulativo de X , até o instante t , é dado por:

$$E_t = \rho E_{t-1} + \beta X_t.$$

Observe-se que a estimação dos efeitos de X sobre Y fica, então, restrita a dois parâmetros: β e ρ . Ainda, além da magnitude e da forma, também o período de persistência do efeito presumido de X passa a ser estimado, uma vez que tal formulação baseia-se na expressão não truncada (1.1), havendo a possibilidade de efeitos significativos para qualquer horizonte, embora a condição de estabilidade ($0 \leq \rho < 1$) sobre estes modelos implique que o efeito de X seja eventualmente dissipado.

Para se obter efeitos que cresçam até um ponto máximo para, então, decair, Koyck (1954) propõe que se especifique $b > 0$ em (1.3), obtendo-se:

$$E_t = \beta_0 X_t + \sum_{j=1}^b (\beta_j - \rho \beta_{j-1}) X_{t-j} + \rho E_{t-1},$$

ao custo de se estimar b parâmetros adicionais. Se o valor de b necessário à descrição adequada da trajetória dos efeitos de X for relativamente elevado, entretanto, recai-se nas dificuldades levantadas em relação a (1.2) devido à autocorrelação serial na regressora.

Uma forma alternativa para restrição dos coeficientes em um modelo de defasagens distribuídas truncado, como em (1.2), é proposta por Almon (1965), sendo a trajetória dos coeficientes de X no instante atual e em suas s defasagens aproximada por um polinômio de baixo grau d :

$$\beta_j = \sum_{k=0}^d \zeta_k j^k, \quad \text{para } j=0, \dots, s. \quad (1.4)$$

Aplicando-se (1.4) a (1.2), obtém-se um preditor em termos de ζ e de um conjunto de $d + 1$ novas regressoras pouco correlacionadas, cada uma das quais expressa em função de X_t, \dots, X_{t-s} . Os parâmetros a estimar passam a ser aqueles que governam o polinômio suavizador, mais especificamente, se os coeficientes β_j são adequadamente descritos por um polinômio de grau d , deve-se estimar os parâmetros ζ_0, \dots, ζ_d . Tal abordagem proporciona diversos formatos para a trajetória dos efeitos da regressora ao longo do tempo. Sua principal limitação reside no fato de basear-se na forma truncada (1.2), em que o horizonte s de influência de X sobre Y deve ser arbitrado, implicitamente assumindo-se que efeitos além desse horizonte sejam nulos.

Ainda que sujeitas às limitações já mencionadas, as duas abordagens supracitadas para restrição de coeficientes em um modelo de defasagens podem ser aplicadas com êxito na representação da dinâmica do efeito de uma regressora em diversas situações. Neste trabalho aplicam-se ambas as formas, tratando-as como casos particulares de funções de transferência, como descritas em Box et al. (1994). Uma função de transferência é uma formulação matemática descrevendo a relação entre a entrada e a saída de um sistema, obtida ao se aplicar, à regressora cuja inércia se deseja modelar, um filtro linear obtido da razão entre dois polinômios no operador defasagem: o primeiro de ordem $r \geq 0$ e, o segundo, de ordem $s \geq 0$. Como será visto ao longo deste trabalho, a abordagem de Koyck (1954) é obtida fazendo-se $r = 1$ e $s \geq 0$, enquanto na abordagem de Almon (1965) obtém-se uma forma particular de função de transferência em que $r = 0$. A adoção de funções de transferência constitui-se, dessa forma, em um instrumento

poderoso para análise de efeitos acumulados de uma regressora X sobre a resposta esperada, uma vez que, com base em parametrizações bastante parcimoniosas, pode-se ter uma enorme gama de formatos para a propagação de efeitos de X ao longo do tempo.

Embora extremamente relevante, a adoção de funções de transferência é, usualmente, restrita a modelos para os quais a resposta Y_t é suposta gaussiana, muitas vezes ao custo de transformações sobre a variável resposta, as quais podem levar a dificuldades para interpretação do modelo. Poucos são os registros encontrados na literatura que consideram funções de transferência associadas a respostas não-normais. No presente trabalho, funções de transferência são inseridas em preditores associados a observações cuja distribuição pertença à família exponencial. Uma outra característica desejável e aqui tratada é a flexibilidade dos modelos propostos, no sentido de se reconhecer formalmente incertezas associadas à passagem do tempo, o que é feito permitindo-se evolução estocástica aos parâmetros envolvidos. Neste sentido, a classe dos modelos dinâmicos lineares generalizados, definida por West et al. (1985), apresenta-se como a moldura ideal: tais modelos podem ser vistos tanto como uma generalização dos modelos lineares generalizados (Nelder e Wedderburn, 1972) - ao prescindir da hipótese de ausência de correlação serial entre observações - quanto como uma extensão da classe dos modelos dinâmicos lineares (Harrison e Stevens, 1976) - ao permitirem variáveis resposta não gaussianas.

Toda a inferência realizada no presente trabalho segue o paradigma bayesiano e, portanto, a incerteza associada aos parâmetros a estimar deve ser modelada probabilisticamente. Assim, a partir de distribuições a priori que reflitam as crenças iniciais sobre os parâmetros de interesse, combinadas a uma função de verossimilhança, obtida a partir da estrutura probabilística suposta adequada aos dados observados, objetiva-se determinar, a cada instante, distribuições a posteriori, atualizando-se as crenças iniciais sobre os parâmetros - expressas pela priori - à luz das informações trazidas pela verossimilhança. O processo de aprendizagem sobre os parâmetros fundamenta-se no teorema de Bayes, de acordo com a seguinte premissa: a cada instante tem-se uma distribuição a posteriori com respeito a t , a qual é atualizada por meio de uma equação de sistema, que dita a forma de evolução dos parâmetros, passando-se a ter uma distribuição a priori com relação a $t + 1$. Sob normalidade e linearidade, todo o processo de atualização pode ser feito de forma analítica, por meio do filtro de Kalman, como descrito por West e Harrison (1997), do ponto de vista bayesiano, e por Harvey (1989), sob o enfoque clássico.

No contexto da inserção de funções de transferência em modelos dinâmicos lineares generalizados, entretanto, dois fatores combinam-se, impedindo a obtenção analítica da densidade a posteriori:

- a natureza não gaussiana da resposta, associada a prioris não conjugadas;
- a não-linearidade dos preditores propostos nos parâmetros autoregressivos envolvidos nas funções de transferência.

Diversas formas têm sido propostas na literatura para a aproximação de distribuições a posteriori nesse contexto. Algumas abordagens baseiam-se em aproximações normais para a variável resposta e, sob linearidade do preditor, obtém-se a posteriori aplicando-se o filtro de Kalman. Outras preservam a natureza observacional, fixando-se na atualização de momentos das distribuições envolvidas no processo de atualização e outras, ainda, na obtenção de modas

a posteriori, evitando-se a integração necessária à obtenção dessa distribuição. A partir da década de 1990, com o avanço de métodos computacionais, passou a ser possível fazer inferência completa (obtendo-se estimativas para quaisquer quantidades de interesse associadas à posteriori) sobre modelos dinâmicos lineares generalizados. Diferentes métodos desenvolvidos a partir de então preconizam a obtenção de amostras das distribuições de interesse, ao invés da atualização analítica das mesmas. Em particular, para obtenção de amostras de distribuições a posteriori utilizam-se, no presente trabalho, os métodos de Monte Carlo via cadeias de Markov (MCMC), descritos detalhadamente por Gamerman (1997). Tais métodos fazem uso de algoritmos específicos para a construção de cadeias de Markov que tenham, como distribuição estacionária, a distribuição em que se tem interesse: neste caso, em particular, a distribuição a posteriori conjunta dos parâmetros envolvidos nos modelos propostos. O procedimento é iterativo e, uma vez suposta a convergência das cadeias geradas, tomam-se pontos destas como amostras da distribuição a posteriori, os quais podem ser utilizados para se obter estimativas de momentos de tal distribuição, bem como de quantis e quaisquer funções definidas sobre seus argumentos, por meio de manipulações bastante simples, como é usual em métodos de Monte Carlo.

Após ampla revisão bibliográfica, não foram encontrados trabalhos que objetivassem a realização de inferência bayesiana completa sobre modelos dinâmicos lineares generalizados envolvendo funções de transferência como um dos blocos estruturais em seus preditores. Assim, a relevância do presente trabalho situa-se na realização de inferência bayesiana, por meio de métodos de Monte Carlo via cadeias de Markov, sobre modelos que pressupõem:

- variáveis resposta com distribuição pertencente à família exponencial;
- preditores não necessariamente lineares, incluindo, entre outros blocos estruturais tais como tendência, sazonalidade e demais covariáveis, um bloco que trata especificamente de efeitos imediatos e acumulados de uma regressora, por meio de funções de transferência;
- o tratamento formal de autocorrelação serial ao permitir evolução estocástica dos parâmetros envolvidos.

1.2 Organização do Trabalho

A dissertação está organizada da seguinte forma: no capítulo 2 é apresentada uma revisão sucinta do ferramental teórico/metodológico envolvido. Modelos dinâmicos lineares e não-normais/ não-lineares são descritos, respectivamente, nas seções 2.1 e 2.2. Nessa última, é feita uma revisão dos métodos mais recentes para aproximação da posteriori em modelos não-normais/ não-lineares. Em seguida, na seção 2.3, apresenta-se a definição formal de funções de transferência, que, a cada instante, representam o efeito cumulativo de uma regressora, bem como de funções de resposta ao impulso, que representam o efeito sobre a resposta esperada

devido a um pulso em uma regressora, no instante em que ocorre tal pulso e nos instantes subsequentes. Na seção 2.4 descrevem-se os métodos de Monte Carlo via Cadeias de Markov, utilizados para aproximação da posteriori. Detalham-se, então, os algoritmos usados de forma a construir transições que garantam a convergência das cadeias geradas para a distribuição a posteriori: o amostrador de Gibbs e o algoritmo Metropolis-Hastings, os quais são utilizados de forma combinada nas aplicações desenvolvidas. Diferentes escolhas de distribuição para a variável resposta, aliadas a especificações alternativas para os preditores, determinam modelos que podem ser comparados com base nos critérios descritos na seção 2.5.

No capítulo 3, define-se de forma mais específica o problema a ser tratado. A estrutura geral dos modelos propostos é apresentada na seção 3.1 e casos particulares dessa estrutura são obtidos ao se variar a estrutura observacional - que, como descrito na seção 3.2, é obtida a partir da suposição de resposta distribuída segundo os modelos Poisson, Binomial, Gama e Gaussiano Inverso - bem como a forma de propagação dos efeitos da regressora de interesse. Assim, na seção 3.3 são apresentadas todas as formas de funções de transferência adotadas neste trabalho. Questões computacionais envolvidas na obtenção da posteriori via MCMC são detalhadas na seção 3.4. Finalmente, a forma de utilização das amostras geradas para estimação e previsão é descrita na seção 3.5.

No capítulo 4, tanto a identificabilidade dos modelos propostos quanto o desempenho dos procedimentos de estimação são avaliados a partir da aplicação dos algoritmos descritos na seção 3.4 a conjuntos de dados gerados artificialmente, a partir de valores arbitrados para as componentes do vetor paramétrico. O capítulo apresenta os resultados da estimação de diversas estruturas preditivas atribuídas a respostas Poisson (seção 4.1), Binomial (seção 4.2), Gama (seção 4.3) e Gaussiana Inversa (seção 4.4). Na seção 4.5 tecem-se alguns comentários sobre os resultados obtidos.

No capítulo 5, a metodologia desenvolvida é aplicada a dois conjuntos de dados reais. Na primeira aplicação, desenvolvida na seção 5.1, com resposta discreta, objetiva-se avaliar o impacto de poluentes atmosféricos sobre contagens diárias de óbitos de crianças com idades inferiores a 5 anos na região metropolitana de São Paulo. Trabalhos relacionando níveis de poluição atmosférica a desfechos epidemiológicos em grandes metrópoles têm sido frequentes nos últimos anos, mas muitos desses desconsideram efeitos de poluentes associados a períodos relativamente longos. Alguns fazem uso de defasagens distribuídas por até três dias ou, alternativamente, de médias móveis dos níveis de poluentes, usualmente compreendendo não mais que sete dias. O foco primordial de atenção na aplicação feita aqui recai sobre a estimação da forma e do tempo de persistência do efeito dos níveis de poluição sobre os desfechos analisados. O modelo adotado é Poisson, com diferentes estruturas preditivas. Na segunda aplicação, apresentada na seção 5.2, um modelo Gama é usado para quantificar o efeito cumulativo de volumes diários de chuva sobre níveis do poluente material particulado. Sabe-se que a chuva provoca o depósito do material particulado no solo e na água, "limpando" a atmosfera. Deseja-se avaliar o período de influência da chuva sobre os níveis de tal poluente. A modelagem de níveis de poluentes é bastante complexa, envolvendo características físico-químicas dos mesmos, bem como fatores meteorológicos. Ainda assim, esforços no sentido de se criar mecanismos de ajuste e predição dos níveis de poluentes são extremamente úteis, uma vez que diversos estudos envolvendo impacto de poluentes sobre diferentes desfechos esbar-

ram na ocorrência de períodos sem registro dos níveis de poluição. Devido à complexidade das respostas tratadas e dos mecanismos que as definem, os modelos propostos não pretendem dar conta de toda a incerteza a que estão sujeitos tais processos, mas lançar luz sobre algumas das questões relevantes nesse contexto. Na seção 5.3, resumem-se os resultados obtidos nas duas aplicações desenvolvidas.

Por fim, no capítulo 6 são traçadas conclusões sobre os resultados obtidos e enunciadas perspectivas sobre tópicos para pesquisa futura.

CAPÍTULO 2

Revisão de Literatura

Neste capítulo, uma revisão da teoria envolvida no presente trabalho é apresentada. Na seção 2.1, a classe de modelos dinâmicos lineares é descrita. Tais modelos são capazes de tratar formalmente correlação serial, tendo, como casos particulares, modelos tradicionais para séries temporais e modelos de regressão dinâmica, com coeficientes evoluindo estocasticamente. Uma de suas hipóteses básicas, entretanto, é a estrutura gaussiana para a variável resposta. A seção 2.2 apresenta diferentes aspectos sobre a violação das hipóteses básicas dos modelos dinâmicos lineares. Estes foram estendidos por West et al. (1985), de forma a tratar variáveis resposta com distribuições pertencentes à família exponencial: surgem, aí, os modelos dinâmicos lineares generalizados, descritos na subseção 2.2.1. As violações das hipóteses básicas de um modelo dinâmico linear levam a dificuldades para a obtenção analítica de distribuições a posteriori. Algumas abordagens para aproximação da posteriori utilizadas nas últimas décadas são revisadas na subseção 2.2.2.

Uma característica interessante nos modelos dinâmicos é o fato de o preditor poder ser particionado em blocos estruturais, com interpretação clara. Neste trabalho, tem-se interesse particular no bloco estrutural em que se modela a propagação dos efeitos de uma covariável sobre a resposta esperada. Neste sentido, a seção 2.3 apresenta uma breve revisão da teoria de funções de transferência.

A adoção de funções de transferência pode tornar o preditor não-linear, o que é mais um fator de dificuldade para obtenção analítica da posteriori, somado à natureza não-gaussiana da resposta. O método utilizado para aproximação da posteriori nos capítulos subsequentes é a simulação de Monte Carlo via Cadeias de Markov, que tem alguns de seus aspectos discutidos na seção 2.4.

Finalmente, na seção 2.5, são descritos os dois métodos para a escolha, entre diferentes modelos propostos, daquele(s) que se destaca(m), tanto do ponto de vista de qualidade de ajuste, quanto de parcimônia na parametrização.

2.1 Modelos Dinâmicos Lineares

Uma questão relevante na modelagem de dados observados temporalmente é o reconhecimento formal da estrutura de autocorrelação inerente a esse tipo de observações. A classe de modelos dinâmicos bayesianos, apresentada por Harrison e Stevens (1976), possibilita o tratamento de autocorrelação temporal, por meio da premissa de que os parâmetros (ou parte dos parâmetros) envolvidos em um modelo de regressão possam evoluir temporal e estocasticamente. Ao se adotar o paradigma bayesiano, admitem-se tanto informações observadas historicamente quanto opiniões subjetivas como fontes lícitas de informação para realização de inferência sobre os modelos propostos. Atribuem-se distribuições de probabilidade aos parâmetros, de forma a refletir a incerteza associada a estes.

A classe de modelos dinâmicos lineares (MDL) é discutida detalhadamente por West e Harrison (1997) - do ponto de vista bayesiano - e por Harvey (1989), que os denomina modelos de espaço de estados e aborda o problema sob o paradigma clássico de inferência. Tais modelos são definidos hierarquicamente, com base em duas equações. A primeira, denominada equação de observação, descreve o caráter estocástico da variável resposta e sua relação com as variáveis regressoras, por meio de um vetor paramétrico θ_t . A segunda, denominada equação de sistema ou de evolução, descreve a variação estocástica dos coeficientes da regressão. Tem-se, então, a seguinte estrutura:

$$Y_t = \mathbf{F}'_t \theta_t + v_t, \quad v_t \sim N(0, V_t) \quad (2.1)$$

$$\theta_t = \mathbf{G}_t \theta_{t-1} + \mathbf{u}_t, \quad \mathbf{u}_t \sim N(0, \mathbf{W}_t), \quad (2.2)$$

em que Y_t é a observação no instante t , $t = 1, 2, \dots, T$. Ao longo deste texto, assume-se Y_t escalar, mas a classe acima pode ser definida para um vetor de observações \mathbf{Y}_t . \mathbf{F}'_t é o transposto do vetor de p variáveis regressoras conhecidas a cada instante t , acomodando nível, tendência, sazonalidade, etc e θ_t é um vetor de parâmetros (ou estados) com dimensão $p \times 1$, que governa a relação entre a resposta e as regressoras; \mathbf{G}_t é uma matriz de dimensão $p \times p$ que descreve a evolução dos parâmetros θ_t e, finalmente, V_t e \mathbf{W}_t são a variância e a matriz de covariâncias dos erros de observação e evolução, v_t e \mathbf{u}_t , respectivamente. No contexto bayesiano, denotando-se a informação inicial disponível por D_0 , o modelo usualmente se completa com uma densidade a priori $(\theta_1 | D_0) \sim N(\mathbf{a}_1, \mathbf{R}_1)$ e com a hipótese de que os erros \mathbf{u}_t e v_t sejam serial e mutuamente independentes. Pode-se, ainda, acrescentar outros níveis à hierarquia do modelo e atribuir densidades a priori a hiperparâmetros.

Em suma, um modelo dinâmico linear fica completamente especificado pela quádrupla $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$. Alguns casos particulares relevantes são:

- modelos de regressão linear, obtidos ao se fazer $\mathbf{G} = \mathbf{I}_p$, a matriz identidade de ordem p , e $u_t = 0, \forall t$;
- modelos de séries temporais, se $\mathbf{F}_t = \mathbf{F}$ e $\mathbf{G}_t = \mathbf{G}, \forall t$;
- modelos de regressão dinâmica, se $\mathbf{G}_t = \mathbf{I}_p$.

Em um modelo de espaço de estados, blocos estruturais com interpretação bastante clara, retratando diferentes características da série observada - tais como tendência, sazonalidade, efeito passado e presente de regressoras (efeitos acumulados de regressoras são detalhados na seção 2.3) e outros - podem ser combinados, por meio do princípio da superposição, construindo-se um único modelo que tenta explicar os padrões subjacentes à série. Assim, uma grande vantagem de tais modelos é poder-se pensar no tratamento individual de diferentes características da série temporal de interesse e, então, reunir estes "submodelos", estimando-se conjuntamente os diferentes padrões observados. A superposição dos blocos estruturais é feita, simplesmente, empilhando-se, no vetor \mathbf{F} , as regressoras envolvidas em cada bloco e, na matriz \mathbf{G} , as respectivas submatrizes de evolução, de forma a compor uma matriz bloco-diagonal.

Tradicionalmente, a inferência sobre os modelos dinâmicos lineares é feita sequencialmente. Primeiramente, a informação inicial sobre os parâmetros de interesse é resumida em sua densidade a priori. Uma hipótese usual é a de que o modelo seja fechado a informações externas, ou seja, excetuando-se a informação contida na densidade a priori, toda a informação relevante é trazida somente pela verossimilhança. Assim, a informação disponível no instante t é $D_t = \{D_{t-1} \cup y_t\}$. Combinando-se, por meio do teorema de Bayes, a densidade a priori à verossimilhança, obtida através do modelo postulado para a variável resposta, obtém-se, a cada instante, a densidade atualizada para os parâmetros de interesse, denominada densidade a posteriori. Esta, por sua vez, é usada para a construção da densidade a priori no instante seguinte. Este processo de atualização é resumido a seguir. A cada instante t , obtém-se a densidade a priori:

$$\pi(\theta_t | D_{t-1}) = \int \pi(\theta_t | \theta_{t-1}) \pi(\theta_{t-1} | D_{t-1}) d\theta_{t-1}, \quad (2.3)$$

em que $\pi(\theta_t | \theta_{t-1})$ é obtida aplicando-se a equação de evolução (2.2) e $\pi(\theta_{t-1} | D_{t-1})$ é a densidade a posteriori do instante anterior.

O objetivo primordial em diversas análises e, em particular, em análises envolvendo séries temporais, é a previsão de valores futuros do processo modelado. Uma vez que se determine a densidade a priori para o instante t , a distribuição preditiva a um passo é obtida da seguinte forma:

$$p(y_t | D_{t-1}) = \int p(y_t | \theta_t) \pi(\theta_t | D_{t-1}) d\theta_t. \quad (2.4)$$

O processo de atualização dos parâmetros completa-se ao se obter a densidade a posteriori, aplicando-se o teorema de Bayes:

$$\pi(\theta_t | D_t) \propto p(y_t | \theta_t) \pi(\theta_t | D_{t-1}). \quad (2.5)$$

Previsões h passos à frente são, então, obtidas a partir de:

$$p(y_{t+h} | D_t) = \int p(y_{t+h} | \theta_{t+h}) \pi(\theta_{t+h} | D_t) d\theta_{t+h}, \quad (2.6)$$

em que

$$\pi(\theta_{t+h} | D_t) = \int \pi(\theta_{t+h} | \theta_t) \pi(\theta_t | D_t) d\theta_t. \quad (2.7)$$

e $\pi(\theta_{t+h} | \theta_t)$ é obtida aplicando-se a equação de evolução (2.2) recursivamente h vezes.

2.2 Modelos Dinâmicos Não-normais / Não-lineares

Sob a hipótese de normalidade dos erros em (2.1) e (2.2) e se a quádrupla $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$ é conhecida, as equações (2.3), (2.4), (2.5) e (2.6) podem ser solucionadas aplicando-se o filtro de Kalman. Essencialmente, o filtro propaga e atualiza os momentos de primeira e segunda ordens do vetor de estados. Sob normalidade e linearidade, tais momentos contêm toda a informação relevante sobre os estados. Entretanto, é usual ter-se violações destas hipóteses, como elementos das matrizes acima dependentes de quantidades desconhecidas ou distribuição não-gaussiana para Y_t .

De acordo com Durbin e Koopman (2000), violações da hipótese de normalidade podem ser divididas em duas categorias: observações vindas da família exponencial ou modelos com erros não-normais. Além disso, outras formas de violação do modelo dinâmico linear podem ocorrer ao se permitir variâncias desconhecidas na equação observacional e/ ou evolucionar, bem como parâmetros desconhecidos nas matrizes \mathbf{F}_t e \mathbf{G}_t . Nesse caso, é necessário propagar e atualizar toda a densidade dos parâmetros de estado e as integrais (2.3), (2.4), (2.5) e (2.6) não possuem solução analítica, devendo-se recorrer a métodos computacionais de aproximação para obtenção destas.

2.2.1 Modelos dinâmicos lineares generalizados

Nelder e Wedderburn (1972) introduziram a classe de modelos lineares generalizados, definidos por observações cuja distribuição de probabilidade seja membro da família exponencial, por um conjunto de covariáveis independentes, às quais é aplicada uma estrutura linear e, ainda, por uma função de ligação g monótona e diferenciável, que estabelece a relação entre a média da variável resposta e as covariáveis. Se Y_t é a variável resposta e \mathbf{F}_t é o vetor de covariáveis no instante t , tem-se então a seguinte estrutura para o modelo linear generalizado:

$$p(y_t | \chi_t) \propto \exp \left\{ \frac{y_t \chi_t - b(\chi_t)}{\phi_t} \right\}, \quad t = 1, \dots, T, \quad (2.8)$$

em que χ_t é o parâmetro canônico ou natural da distribuição de Y_t , ϕ_t é um parâmetro de escala e

$$\mu_t = E[Y_t | \chi_t] = \dot{b}(\chi_t) = \frac{db(\chi_t)}{d\chi_t}.$$

A relação entre μ_t e o vetor de regressoras no instante t , \mathbf{F}_t é dada por:

$$g(\mu_t) = \eta_t = \mathbf{F}_t' \boldsymbol{\theta}.$$

Embora essa classe de modelos seja bastante ampla, tem-se limitações inerentes a estes, principalmente no que tange à dependência entre as observações. Supõe-se, nos modelos lineares generalizados, que toda a estrutura de correlação entre as observações Y_1, Y_2, \dots, Y_T esteja

guardada no parâmetro (ou vetor paramétrico) θ ; assim, condicionalmente a θ , as observações são supostas independentes. Portanto, em situações nas quais haja dependência entre as observações, como ocorre em aplicações de séries temporais, o uso destes modelos pode comprometer a validade das inferências realizadas.

West et al. (1985) introduziram a classe de modelos dinâmicos lineares generalizados (MDLG), uma extensão tanto para os modelos lineares generalizados - no sentido de permitir a evolução temporal dos parâmetros - quanto para os modelos dinâmicos lineares - no sentido de permitir observações com distribuição na família exponencial. Esta classe é definida, então, pela equação observacional (2.8), mas o preditor η_t pode ser não linear e varia dinamicamente, segundo:

$$\begin{aligned} g(\mu_t) &= \eta_t = \mathbf{F}'_t \boldsymbol{\theta}_t \\ \boldsymbol{\theta}_t &= \mathbf{G}_t \boldsymbol{\theta}_{t-1} + u_t, \quad u_t \sim N(0, \mathbf{W}_t). \end{aligned}$$

2.2.2 Aproximação da posteriori em modelos dinâmicos não-normais/ não-lineares: algumas alternativas

Nesta subseção, é feita uma revisão de alguns métodos adotados nas duas últimas décadas para aproximação da distribuição a posteriori em modelos dinâmicos não-normais/ não-lineares. Além de se ilustrar a evolução dos métodos disponíveis nos últimos anos, são levantadas, a partir desta revisão, diferentes formas de violação das hipóteses básicas de um modelo dinâmico linear, além daquelas consideradas neste trabalho e descritas na subseção 2.2.1.

West et al. (1985) fazem inferência aproximada em modelos dinâmicos lineares generalizados usando *linear Bayes*. O processo de atualização baseia-se na especificação da distribuição a priori de tais parâmetros em termos de primeiros e segundos momentos, determinados por analogia aos modelos dinâmicos lineares. A partir daí, e utilizando uma priori conjugada para o parâmetro natural da família exponencial, os momentos de primeira e segunda ordens a posteriori para os parâmetros estruturais são obtidos, completando o processo de atualização.

Kitagawa (1987) apresenta um algoritmo para filtragem e suavização em modelos de espaço de estados não lineares e livres da suposição de normalidade, tanto para os erros de sistema quanto para os erros de evolução. O método proposto consiste na aproximação das densidades envolvidas no processo de filtragem e suavização de forma não paramétrica, por meio de *splines* de primeira ordem. O custo computacional envolvido, entretanto, torna-se elevado para vetores de estados com dimensão alta.

Pole (1988) considera modelos normais não-lineares. A distribuição a posteriori é obtida numericamente, por meio de uma variação dinâmica do método de quadratura gaussiana, baseado em grades atualizadas sequencialmente para o parâmetro não-linear e usando o fato de que, condicionalmente a este parâmetro, tem-se um modelo dinâmico linear. Integram-se, então, as quantidades de interesse em relação ao parâmetro de não-linearidade, obtendo-se inferências incondicionais. A atualização dinâmica das grades (construídas com poucos pontos) é um artifício para contornar a demanda computacional necessária ao se utilizar grades muito refinadas. O método da quadratura gaussiana também é utilizado por Moreira (1994), porém com grade fixa.

Singh e Roberts (1992) trabalham com modelos dinâmicos lineares generalizados, propondo uma aplicação iterativa do filtro de Kalman linear, para obtenção de modas a posteriori, substituindo a equação observacional (2.8) por:

$$\tilde{y}_t = \mathbf{F}'_t \boldsymbol{\theta}_t + \tilde{v}_t, \quad \tilde{v}_t \sim N(0, \tilde{V}_t), \quad (2.9)$$

em que \tilde{y}_t são observações modificadas, dadas por uma aproximação linear das observações reais, segundo:

$$\tilde{y}_t = \boldsymbol{\eta}_t + (y_t - \mu_t) \dot{g}(\mu_t) \quad (2.10)$$

e com variâncias dadas por:

$$\tilde{V}_t = \tilde{V}_t(\boldsymbol{\theta}_t) = \phi_t \ddot{b}(\chi_t) [\dot{g}(\mu_t)]^2, \quad (2.11)$$

com \dot{g} e \ddot{b} indicando a primeira e a segunda derivadas das funções g e b , respectivamente.

Fahrmeir (1992) considera extensões multivariadas de modelos dinâmicos lineares generalizados, mantendo a estrutura de transição linear e evita a integração numérica obtendo modas a posteriori dos parâmetros de estado, por meio de uma generalização do filtro de Kalman estendido. Outro exemplo de aplicação envolvendo inferência baseada em modas a posteriori em modelos dinâmicos lineares generalizados pode ser visto em Fahrmeir (1997).

A partir do início da década de 1990, diversos procedimentos para aproximação da distribuição a posteriori baseados em amostras foram desenvolvidos. A idéia é aproximar a distribuição dos parâmetros de estado por uma amostra desta. Descrevem-se, a seguir, algumas propostas nesse sentido.

Gordon et al. (1993) usam um esquema de reamostragem para aproximar a posteriori. O esquema baseia-se na idéia de que, como se pode ver na equação (2.5), a mensuração de y_t é usada para modificar a priori obtida com base na informação até o instante $t - 1$. Assim, uma vez disponível um conjunto de N amostras de $\pi(\boldsymbol{\theta}_{t-1} | D_{t-1})$, $\{\boldsymbol{\theta}_{t-1}(i) : i = 1, \dots, N\}$, utiliza-se a equação de sistema para se obter uma amostra de $\pi(\boldsymbol{\theta}_t | D_{t-1})$, como de costume e, ao se reamostrar os pontos de tal amostra com pesos $p(y_t | \boldsymbol{\theta}_t(i)) / \sum_j p(y_t | \boldsymbol{\theta}_t(j))$, tem-se convergência em distribuição para a posteriori, como mostram Smith e Gelfand (1992). Durbin e Koopman (2000) usam amostragem por importância com variáveis antitéticas e fazem inferência clássica e bayesiana em modelos de espaço de estados apresentando erros não normais, tanto na equação observacional quanto na de evolução.

Os métodos de Monte Carlo via Cadeias de Markov (MCMC) têm sido uma opção bastante freqüente na literatura, entre os métodos de aproximação da posteriori baseados em amostras. Tais métodos, utilizados nas aplicações feitas neste trabalho, são detalhados na seção 2.4. O algoritmo MCMC mais simples é o amostrador de Gibbs. Carlin et al. (1992) aplicam-no para obter aproximações da posteriori em modelos de espaço de estados nos quais as variâncias observacional e/ ou evolucionar são perturbadas por parâmetros desconhecidos. Diferentes especificações para as distribuições dos parâmetros de perturbação conduzem a uma variedade de distribuições não gaussianas para os erros, como, por exemplo, exponencial dupla ou t de Student. Permitem-se, ainda, estruturas não-lineares tanto na equação de observação quanto

na de sistema. Carter e Kohn (1994) utilizam MCMC em modelos dinâmicos, com erros distribuídos segundo uma mistura de normais. Shephard (1994) utiliza métodos MCMC para inferir sobre modelos de espaço de estados "parcialmente Gaussianos", nos quais as variâncias dos erros dependem de processos Markovianos. São tratados como casos particulares da estrutura apresentada: modelos para observações com distribuição *t* de Student, modelos de misturas normais, modelos multiplicativos e modelos de mudança de regime, entre outros. Outros exemplos de aplicações de algoritmos MCMC em modelos de espaço de estados não normais e/ou não lineares podem ser vistos em De Jong e Shephard (1995), Shephard e Pitt (1997), que apresentam, além da estimação bayesiana, um algoritmo de amostragem por importância para estimação clássica de modelos dinâmicos não-normais, Knorr-Held (1999), Geweke e Tanizaki (2001), e Durbin e Koopman (2002).

No presente trabalho, a estratégia adotada é aquela proposta por Gamerman (1998). As abordagens baseadas em algoritmos MCMC diferem, essencialmente, na forma de atualização do vetor de estados (movimentos individuais ou em blocos) e na escolha de densidades propostas para o algoritmo Metropolis-Hastings, nos casos em que este se faz necessário, além do amostrador de Gibbs. Estas questões são detalhadas na seção 2.4, na qual os métodos de Monte Carlo via Cadeias de Markov são discutidos e é apresentada a estratégia de atualização usada neste trabalho.

Revisões bastante abrangentes sobre métodos para aproximação da densidade a posteriori em modelos dinâmicos podem ser vistas em Migon et al. (2005) e Ferreira e Gamerman (2000), este último tratando especificamente de modelos dinâmicos lineares generalizados.

2.3 Funções de Transferência

Diversas aplicações relevantes em séries temporais envolvem a utilização de variáveis regressoras defasadas, com base na premissa de que o processo de interesse no tempo t , Y_t , não seja impactado apenas pelo valor concomitante de uma regressora X_t , mas também por seus valores passados, X_{t-1} , X_{t-2} , \dots , X_{t-s} , ou seja, admite-se que a relação entre Y e X esteja sujeita a algum tipo de inércia ou, equivalentemente, que eventos que afetam a resposta persistam ao longo do tempo. Nesta seção, apresenta-se uma breve revisão da teoria de funções de transferência, utilizadas neste trabalho de forma a modelar efeitos cumulativos de uma regressora.

Em Box et al. (1994), as funções de transferência, definidas a seguir, são construídas em termos de efeitos de uma regressora, X , sobre a variável resposta sob análise, Y . Aqui, uma pequena modificação é feita, visando-se a inserção de tais funções como blocos estruturais nos modelos dinâmicos lineares generalizados, descritos na seção 2.2.1. Segue-se, então, a formulação adotada por West e Harrison (1997), que se diferencia apenas por modificar a estrutura de regressão, considerando não a regressora X impactando diretamente a resposta Y , mas, ao invés disso, uma forma latente E construída de forma a representar a combinação de efeitos passados e presentes de X .

Suponha-se, portanto, que os efeitos passados e presentes da regressora X sobre o processo

de interesse possam ser representados da seguinte forma:

$$\begin{aligned} (1 - \rho_1 B - \dots - \rho_r B^r) E_t &= (\omega_0 - \omega_1 B - \dots - \omega_s B^s) X_{t-b} \\ &= (\omega_0 B^b - \omega_1 B^{b+1} - \dots - \omega_s B^{b+s}) X_t \end{aligned}$$

ou

$$\rho(B) E_t = \omega(B) B^b X_t = \beta(B) X_t, \quad (2.12)$$

em que B é o operador defasagem, ou seja, $BE_t = E_{t-1}$.

Alternativamente, pode-se representar o efeito da regressora X por meio de um filtro linear:

$$\begin{aligned} E_t &= \frac{\beta(B)}{\rho(B)} X_t \\ &= v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots \\ &= v(B) X_t. \end{aligned} \quad (2.13)$$

O polinômio $v(B) = v_0 + v_1 B + v_2 B^2 + \dots$ é denominado *função de transferência* e representa o efeito cumulativo da regressora X sobre o processo de interesse. Os pesos v_0, v_1, v_2, \dots são denominados *função de resposta ao impulso* e representam os efeitos de X sobre o processo de interesse no tempo presente e nos instantes subseqüentes.

Um processo ARIMA, por exemplo, pode ser visto como um modelo de função de transferência, tendo como variável de entrada um ruído branco. O filtro (2.13) é dito estável se, para $|B| \leq 1$, $v(B)$ é convergente, ou, equivalentemente, se a equação característica $\rho(B) = 0$, quando resolvida para B , tem suas raízes fora do círculo unitário. Tais condições implicam que incrementos finitos na variável regressora resultem em incrementos também finitos na resposta, ou, de outra forma, que a função de transferência (soma das respostas a impulso) eventualmente convirja para um estado de estabilidade. As condições para estabilidade da função de transferência são similares às condições de estacionariedade para modelos ARMA.

Ao se substituir a equação (2.13) em (2.12), obtém-se a seguinte identidade:

$$(1 - \rho_1 B - \dots - \rho_r B^r)(v_0 + v_1 B + v_2 B^2 + \dots) = (\omega_0 - \omega_1 B - \dots - \omega_s B^s) B^b. \quad (2.14)$$

Equacionando-se (2.14) em B tem-se:

$$v_j = \begin{cases} 0 & j < b \\ \rho_1 v_{j-1} + \rho_2 v_{j-2} + \dots + \rho_r v_{j-r} + \omega_0 & j = b \\ \rho_1 v_{j-1} + \rho_2 v_{j-2} + \dots + \rho_r v_{j-r} - \omega_{j-b} & j = b+1, b+2, \dots, b+s \\ \rho_1 v_{j-1} + \rho_2 v_{j-2} + \dots + \rho_r v_{j-r} & j > b+s. \end{cases}$$

Claramente, a forma gerada para a função de resposta ao impulso (e, conseqüentemente, para a função de transferência) depende do comportamento das soluções da equação a diferenças que a define. Detalhes sobre a solução de equações a diferenças podem ser vistos em Hamilton (1994). Assim, de acordo com Box et al. (1994), o comportamento da função de resposta ao impulso gerada pela estrutura de ordem (r, s) ou seja, autoregressão de ordem r , aliada a s defasagens em $B^b X_t$, consiste em:

- b valores nulos v_0, v_1, v_{b-1}
- $s - r + 1$ valores v_b, v_{b+1}, v_{b+s-r} sem padrão determinado (se $s < r$ tais valores não ocorrem)
- valores v_j , com $j \geq b + s - r + 1$ seguindo o padrão ditado pela equação a diferenças de ordem r , com valores iniciais dados por $v_{b+s}, \dots, v_{b+s-r+1}$.

Essencialmente, o padrão de comportamento da função de resposta ao impulso é ditado pelas raízes da equação característica $\rho(B) = 0$. Equivalentemente, pensando-se na inserção de funções de transferência como blocos estruturais em modelos de espaço de estados, tais comportamentos são ditados pelos autovalores do bloco da matriz de evolução \mathbf{G} correspondente aos efeitos da regressora cuja inércia está sendo modelada. De forma geral, tem-se os seguintes padrões: raízes distintas e reais do polinômio autoregressivo de ordem r geram termos de decaimento exponencial na função de resposta ao impulso, raízes distintas e complexas fornecem termos senoidais e raízes iguais fornecem termos polinomiais. Portanto, a função de resposta ao impulso será uma mistura de termos exponenciais, polinomiais e senoidais, dependendo do estudo das raízes do polinômio autoregressivo ou, equivalentemente, dos autovalores da matriz de evolução.

A figura 2.1 exhibe padrões das funções de resposta ao impulso e de transferência, para todas as combinações de $r = 0, 1, 2$ e $s = 0, 1, 2$.

2.4 Monte Carlo via Cadeias de Markov

Um dos focos principais do processo de inferência bayesiana é a determinação da densidade a posteriori dos parâmetros de interesse. Nos MDLG's, devido à verossimilhança construída com base na família exponencial, associada a prioris não conjugadas, não se obtém forma fechada para a densidade a posteriori de diversos parâmetros, ao contrário do que ocorre nos modelos dinâmicos normais. Além disso, como visto na seção 2.3, os modelos considerados neste trabalho incluem funções de transferência como blocos estruturais, e tais funções são não-lineares em seus parâmetros autoregressivos, o que constitui mais um fator de dificuldade para a obtenção analítica da densidade a posteriori. Faz-se necessária, portanto, a utilização de algum método para aproximação da posteriori. Na subseção 2.2.2, foram apontadas diferentes abordagens para aproximação da posteriori sob não-normalidade e/ou não-linearidade. No presente trabalho, em particular, a aproximação das densidades a posteriori (e preditivas) de interesse é feita aplicando-se os métodos de Monte Carlo via Cadeias de Markov (MCMC), descritos detalhadamente por Gamerman (1997).

Se uma cadeia de Markov homogênea é irredutível, recorrente positiva e aperiódica, então possui distribuição limite, a qual depende apenas da matriz de transição da cadeia. Além disso, uma vez que a cadeia atinja a distribuição limite, todos os estados subsequentes seguirão tal distribuição. Valendo-se de tais idéias, os métodos MCMC consistem na construção de uma cadeia de Markov que, por meio de escolhas adequadas de núcleos de transição, tenha como distribuição estacionária a densidade da qual se deseja gerar valores: nas aplicações bayesianas,

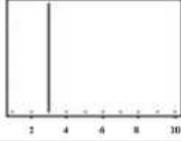
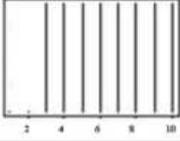
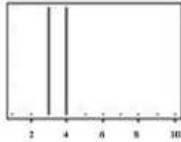
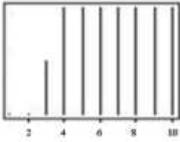
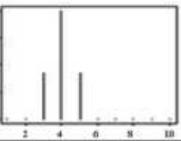
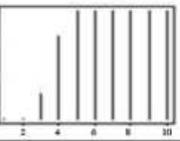
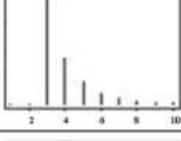
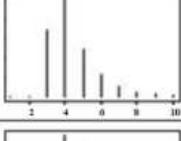
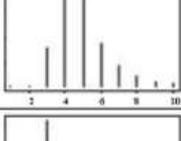
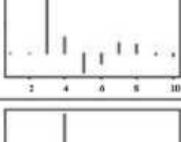
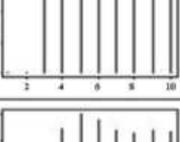
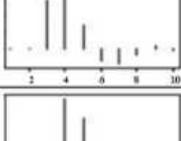
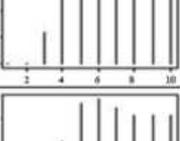
r, s, b	Modelo	Função de Resposta ao Impulso ψ_j	Função de Transferência $\sum_{i=0}^j \psi_i$
002	$E_t = B^2 X_t$		
012	$E_t = (0.5 + 0.5B)B^2 X_t$		
022	$E_t = (0.25 + 0.5B + 0.25B^2)B^2 X_t$		
102	$(1 - 0.5B)E_t = 0.5B^2 X_t$		
112	$(1 - 0.5B)E_t = (0.25 + 0.25B)B^2 X_t$		
122	$(1 - 0.5B)E_t = (0.125 + 0.25B + 0.125B^2)B^2 X_t$		
202	$(1 - 0.6B + 0.4B^2)E_t = 0.8B^2 X_t$		
212	$(1 - 0.6B + 0.4B^2)E_t = (0.4 + 0.4B)B^2 X_t$		
222	$(1 - 0.6B + 0.4B^2)E_t = (0.2 + 0.4B + 0.2B^2)B^2 X_t$		

Figura 2.1 Exemplos de funções de resposta ao impulso e de transferência, para diferentes combinações de ordens $r = 0, 1, 2$ e $s = 0, 1, 2$.

em particular, a densidade a posteriori. Assim, uma vez que se verifique que a convergência da cadeia gerada tenha sido atingida, podem-se utilizar os valores desta (aqueles obtidos após convergência) como amostra da distribuição estacionária.

O teorema ergódico (v. Gamerman (1997)) garante que médias de funções calculadas com base em amostras oriundas de procedimentos MCMC convergem, com probabilidade 1, para as médias teóricas destas funções. Assim, a amostragem via MCMC goza de boas propriedades. Toda a inferência sobre a distribuição estacionária é feita, então, de forma empírica. Por exemplo, para obtenção da estimativa do valor esperado a posteriori, toma-se a média da amostra gerada ou para obtenção de intervalos de credibilidade, tomam-se os percentis adequados na amostra gerada.

Os algoritmos MCMC mais usuais em aplicações de inferência bayesiana são o amostrador de Gibbs e o algoritmo Metropolis- Hastings, descritos, respectivamente, nas subseções 2.4.1 e 2.4.2. Algumas estratégias eficientes para implementação de tais algoritmos são discutidas na subseção 2.4.3. Na subseção 2.4.4, descrevem-se métodos para diagnóstico de convergência das cadeias simuladas.

2.4.1 O amostrador de Gibbs

O amostrador de Gibbs foi apresentado por Geman e Geman (1984), no contexto de processamento de imagens, em que a posteriori de interesse é a distribuição de Gibbs: daí, então, o nome dado ao algoritmo. Uma das dificuldades enfrentadas por Geman e Geman era justamente a determinação da constante de normalização da distribuição de Gibbs. Gelfand e Smith (1990) perceberam o potencial do amostrador de Gibbs para determinação de constantes de normalização em outras distribuições a posteriori e são os responsáveis pela difusão de tal algoritmo na comunidade estatística.

Suponha que se deseje amostrar da distribuição a posteriori do vetor paramétrico $\Theta = (\theta_1, \dots, \theta_d)'$. O amostrador de Gibbs é útil quando não se consegue gerar valores da posteriori conjunta dos parâmetros de interesse, mas as densidades condicionais completas¹ destes parâmetros são conhecidas e a geração de valores das mesmas é possível.

Para construção de uma cadeia de Markov cujas transições sejam definidas pelas condicionais completas, o amostrador de Gibbs procede da seguinte forma:

1. arbitram-se valores iniciais para todos os parâmetros de interesse:

$$\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_d^{(0)})';$$

2. geram-se, iterativamente, valores das densidades condicionais completas de cada um dos parâmetros, sempre se condicionando ao último valor gerado de cada um dos demais parâmetros. Assim, na i -ésima iteração do algoritmo, tem-se:

¹Se a densidade de interesse é $\pi(\theta_1, \dots, \theta_k)$, então a densidade condicional completa para θ_j é $\pi(\theta_j | \theta_{-j}) = \pi(\theta_j | \theta_1, \dots, \theta_{j-1}, \theta_{j+1}, \dots, \theta_d)$

$$\begin{aligned}
\theta_1^{(i)} &\sim \pi(\theta_1 | \theta_2^{(i-1)}, \dots, \theta_d^{(i-1)}) \\
\theta_2^{(i)} &\sim \pi(\theta_2 | \theta_1^{(i)}, \theta_3^{(i-1)}, \dots, \theta_d^{(i-1)}) \\
&\vdots \\
\theta_d^{(i)} &\sim \pi(\theta_d | \theta_1^{(i)}, \dots, \theta_{d-1}^{(i)})
\end{aligned}$$

- atualiza-se o contador para $i + 1$ e repete-se o passo 2 até que a convergência da cadeia tenha sido atingida.

À medida em que o número de iterações cresce, a cadeia de Markov simulada aproxima-se de sua distribuição de equilíbrio e, então, pode-se tomar o valor $\Theta^{(i)} = (\theta_1^{(i)}, \dots, \theta_d^{(i)})$ como um ponto amostral de $\pi(\theta_1, \dots, \theta_d)$.

Nos MDLG's, entretanto, não se consegue amostrar das densidades condicionais completas de todos os parâmetros de interesse, como descrito no passo 2 acima. Nestes casos, passos Metropolis-Hastings podem ser inseridos no algoritmo de Gibbs.

2.4.2 O algoritmo Metropolis-Hastings

O nome Metropolis-Hastings deve-se aos trabalhos de Metropolis et al. (1953) - no contexto de definição de propriedades químicas de substâncias - e Hastings (1970), que generalizou o algoritmo de Metropolis para utilização no contexto estatístico. O algoritmo Metropolis-Hastings é útil para a geração de valores daqueles parâmetros cujas densidades condicionais completas não tenham forma analítica fechada (e, portanto, não estejam disponíveis para amostragem).

Suponha que se deseje gerar valores de uma densidade alvo $\pi(\cdot)$, não disponível para amostragem. A cada iteração do algoritmo Metropolis-Hastings, valores para cada parâmetro são gerados de uma distribuição proposta arbitrária, q . Tais valores são aceitos com uma probabilidade que depende da qualidade do movimento proposto, avaliada com base na densidade proposta, q e na densidade alvo, π . Este algoritmo procede da seguinte forma para geração de um valor θ da densidade-alvo $\pi(\theta)$ de um parâmetro qualquer, na i -ésima iteração:

- gera-se um valor θ_* de uma densidade proposta $q(\theta^{(i-1)}, \theta_*)$;
- aceita-se o valor gerado com probabilidade

$$\alpha(\theta, \theta_*) = \min \left\{ 1, \frac{\pi(\theta_*)/q(\theta^{(i-1)}, \theta_*)}{\pi(\theta^{(i-1)})/q(\theta_*, \theta^{(i-1)})} \right\}. \quad (2.15)$$

A cadeia resultante tem núcleo de transição $q(\theta, \theta_*)\alpha(\theta, \theta_*)$ e converge para $\pi(\theta)$. Pode-se inserir este procedimento no passo 2 do amostrador de Gibbs, de forma a gerar valores para os parâmetros que não apresentem densidades condicionais completas disponíveis para

amostragem, bastando considerar a densidade condicional completa de θ_j como a densidade alvo π presente nos passos do algoritmo Metropolis-Hastings.

Ao se utilizar os algoritmos MCMC para a estimação de modelos de espaço de estados, surgem algumas questões. Uma delas é a escolha de densidades propostas para o algoritmo Metropolis-Hastings. Os passeios aleatórios, tais que $q(\theta, \theta_*) = \phi(|\theta_* - \theta|)$ são uma escolha freqüente, sendo usual tomar ϕ como a densidade normal. Como este tipo de proposta é simétrica, a razão de aceitação em $\alpha(\theta, \theta_*)$ reduz-se à razão de densidades alvo, $\pi(\cdot)$.

Outra simplificação pode ser obtida ao se amostrar das densidades a priori, sendo a razão de aceitação em $\alpha(\theta, \theta_*)$ reduzida, nesse caso, à razão de verossimilhanças.

Embora a escolha de densidades propostas seja arbitrária, claramente escolhas adequadas resultam em maior eficiência para o algoritmo MCMC. Observe-se, em (2.15) que os movimentos da cadeia de Markov simulada dependem de um comprometimento entre densidade proposta, q , e densidade alvo, π . Assim, se a densidade proposta for uma boa aproximação para a densidade alvo, bons movimentos para a primeira também serão bons movimentos para a segunda e a convergência do algoritmo é obtida rapidamente.

Em Knorr-Held (1999), por exemplo, os parâmetros são gerados de densidades a priori autoregressivas gaussianas. Segundo Knorr-Held, a escolha dessa densidade proposta é bastante vantajosa, visto que esta reflete a estrutura de autocorrelação existente entre os parâmetros; além disso, como as gerações são feitas de densidades a priori, as razões de aceitação reduzem-se a razões de verossimilhança, o que facilita os cálculos.

Como descrito na subseção 2.2.2, Singh e Roberts (1992) substituem a equação observacional de um modelo dinâmico linear generalizado por um modelo de trabalho, baseado em observações artificiais obtidas a partir de uma expansão linear em torno da média e dadas por (2.10). Gamerman (1998) faz uso desta idéia, usando o modelo de trabalho definido por (2.9), (2.10) e (2.11) para a construção de densidades-proposta, q , para o algoritmo Metropolis-Hastings. Assim, sugere-se que as densidades propostas utilizadas nos passos Metropolis-Hastings sejam as densidades condicionais completas obtidas a partir do modelo de trabalho proposto por Singh e Roberts. Pretende-se, com isso, que a densidade proposta dependa da verossimilhança, aproximando-se desta. Observe-se que o modelo de trabalho é usado apenas para construção de núcleos de transição do algoritmo Metropolis-Hastings, sendo a estrutura observacional (2.8) mantida.

No presente trabalho, alterna-se entre o uso da densidade proposta sugerida por Gamerman, à qual passaremos a nos referir como "proposta baseada na verossimilhança", e de propostas "passeio aleatório": normais centradas no valor corrente do parâmetro de interesse (ou de alguma transformação do parâmetro que o leve à reta real). Com a proposta passeio aleatório, elimina-se a necessidade de se calcular as quantidades envolvidas no modelo de trabalho. Entretanto, deve-se calibrar a variância da proposta de forma a garantir taxas de aceitação que conduzam à convergência rapidamente. Esta tarefa pode mostrar-se bastante árdua sob a presença de parâmetros de estado em modelos dinâmicos, quando a dimensão do vetor paramétrico é elevada. A densidade proposta sugerida por Gamerman, em contrapartida, tem variância (e, conseqüentemente, taxa de aceitação) determinada diretamente por (2.11) e mostrou-se, nas aplicações desenvolvidas neste trabalho, extremamente eficiente.

2.4.3 Estratégia de amostragem: movimentos individuais ou em blocos

Outra questão relevante na implementação dos algoritmos MCMC é a forma de atualização dos parâmetros estruturais. A estratégia mais natural é a geração individual de valores destes, como, por exemplo, em Carlin et al. (1992). Entretanto, a correlação entre valores vizinhos dos estados em modelos dinâmicos faz com que a convergência neste tipo de estratégia torne-se lenta. Existem diversas propostas para se contornar este problema. Carter e Kohn (1994) e Frühwirth-Schnatter (1994) propõem a geração simultânea de todos os parâmetros de estado, utilizando o amostrador de Gibbs e obtêm resultados eficientes. Knorr-Held (1999) propõe uma solução intermediária entre a geração individual e a geração simultânea de todo o vetor de parâmetros estruturais, particionando-o em blocos cujos tamanhos podem variar (determinística ou estocasticamente) de 2 a T . A metodologia empregada faz uso de gerações dos parâmetros estruturais via algoritmo Metropolis-Hastings e Knorr-Held relata bons resultados ao permitir variação estocástica do tamanho dos blocos a cada iteração do algoritmo.

A abordagem aqui adotada é aquela proposta por Gamerman (1998). Para contornar o problema de correlação inerente aos parâmetros de estado e otimizar a convergência da cadeia, faz-se uso da relação bem definida entre estes parâmetros e os erros de evolução (2.2), usando reparametrização em termos destes distúrbios. A amostragem é feita sobre estes últimos (u_t), que, a priori, não são correlacionados, usando movimentos individuais. O custo computacional envolvido reside justamente na necessidade de, a cada iteração, calcular os valores dos parâmetros de estado a partir do acúmulo dos valores gerados para os erros de evolução. Espera-se, entretanto, que este custo seja compensado pela aceleração na convergência da cadeia.

2.4.4 Monitoração da convergência

Existem diversos métodos para avaliação de convergência dos algoritmos MCMC. Nenhum deles, entretanto, é conclusivo. Estes fornecem apenas indícios de convergência. Assim, para que se tenha confiança de que as cadeias geradas tenham, de fato, atingido a distribuição estacionária de interesse, o ideal é que esta conclusão seja tomada com base em mais que um procedimento. A seguir, três destes métodos são descritos.

- **Inspeção Visual**

Um procedimento informal bastante simples e confiável consiste na inspeção visual das cadeias de Markov por meio de gráficos, verificando-se, por exemplo, se cadeias iniciadas de pontos diferentes mesclam-se, à medida em que o número de iterações cresce, apresentando comportamento de convergência em distribuição.

- **Geweke**

Geweke (1992) sugere um procedimento para teste de convergência a partir da avaliação de médias ergódicas de uma única cadeia gerada, com base na idéia de que, após convergência, diferentes intervalos da cadeia gerada devam apresentar comportamentos semelhantes.

Seja $\psi = t(\theta)$ uma função real de θ . Após um período suficientemente longo de iterações, em que se suspeite já ter havido convergência da cadeia, constroem-se as médias $\bar{\psi}_a$ e $\bar{\psi}_b$, de n_a e n_b iterações (considera-se um intervalo razoável entre estes dois grupos de iterações, de forma a garantir independência entre as amostras). A diferença padronizada entre as médias de ψ nas duas subamostras é tomada como estatística de teste, com distribuição assintótica normal padrão. Assim, diferenças grandes indicam ausência de convergência.

- **Gelman e Rubin**

Gelman e Rubin (1992) sugerem um método baseado na idéia de que cadeias paralelas, geradas de diferentes pontos de partida, devem misturar-se após convergência, desta forma diminuindo as chances de obtenção de distribuições ótimas locais. Uma alternativa, para tanto, é realizar um teste inspirado em técnicas de análise de variância: sob convergência, a variância entre cadeias deve ser pequena em relação à variação interna das cadeias.

2.5 Seleção de Modelos

Diferentes naturezas de variáveis respostas combinadas a especificações variadas para o preditor η_t (e, no contexto deste trabalho, diferentes especificações de funções de transferência) preconizam modelos que podem ser comparados.

Quanto à escolha da forma de função de transferência adotada, supondo-se respostas gaussianas, existem métodos para identificação prévia das ordens r , s e b em (2.12). Nesse contexto e para o caso em que há apenas uma regressora com impacto defasado, Box et al. (1994) propõem um procedimento de identificação de funções de transferência baseado na função de correlação cruzada entre a regressora X_t e a resposta Y_t , após suavizadas pelo mesmo filtro ARMA, escolhido a partir dos procedimentos usuais de identificação de modelos ARMA. A identificação do modelo ARMA a ser adotado é feita sobre a variável de entrada, X_t . Edlund (1987) faz uma revisão de métodos de identificação de funções de transferência e estende o procedimento de identificação para o caso em que há mais que uma regressora envolvida no problema, mas ainda assim considerando apenas respostas gaussianas.

No presente trabalho, não é feita identificação prévia de funções de transferência. Ao invés disso, ajustam-se diversas formulações, propondo-se variadas formas de impacto de X_t sobre Y_t e, com base em procedimentos de seleção de modelos, escolhe-se aquela que melhor se ajusta aos dados. Observe-se, entretanto, que a escolha é feita sobre o modelo que proporciona o melhor ajuste, levando-se em conta, para tal escolha, toda a estrutura do preditor η_t e não somente o bloco E_t de efeitos da regressora. Para escolher a forma estrutural que melhor reflete a relação entre os dados analisados, são aplicados dois métodos, que baseiam-se na idéia de seleção de modelos que tenham boa capacidade de aderência aos dados, porém de forma parcimoniosa, contando, para isso, com um termo de medida de ajuste e um termo de penalidade pela complexidade do modelo. Ambos podem ser aplicados facilmente após obtenção de amostras da posteriori via MCMC e são descritos nas subseções seguintes.

2.5.1 EPD - *Expected predictive deviance*

Gelfand e Ghosh (1998) propõem um critério cujo objetivo é a boa predição de uma réplica dos dados observados, combinada à fidedignidade ao vetor de observados. Denota-se uma réplica do vetor de observações, y_{obs} , por y_{rep} . A perda ao se tomar a ação a quando y_{obs} foi observado e y_{rep} é obtido é denotada por $L(y_{rep}, a; y_{obs})$. A ação a é uma estimativa que busca acomodar tanto as quantidades observadas y_{obs} quanto as previstas, y_{rep} . Para o modelo m , minimiza-se $E \{L(y_{rep}, a; y_{obs}) | y_{obs}\}$ com relação a a , sendo a esperança tomada com base na distribuição preditiva a posteriori para y_{rep} , sob o modelo m . Escolhe-se, então, o modelo para o qual se obtém o menor destes mínimos. Para cada modelo m , deve-se obter, portanto:

$$\min_a E_{y_{rep}|y_{obs},m} \{L(y_{rep}, a; y_{obs})\}. \quad (2.16)$$

Define-se a perda para o l -ésimo componente de y_{rep} e de a por:

$$L(y_{l,rep}, a_l; y_{obs}) = L(y_{l,rep}, a_l) + kL(y_{l,obs}, a_l), \quad k \geq 0. \quad (2.17)$$

O fator k indica a penalidade por devios de a em relação a y_{obs} , comparados a desvios com relação a y_{rep} .

Agregando-se as perdas para todas as componentes do vetor de réplica em (2.17) e substituindo-se em (2.16), tem-se:

$$\begin{aligned} D_k(m) &= \sum_{l=1}^n \min_{a_l} E_{y_{rep}|y_{obs},m} L(y_{l,rep}, a_l; y_{obs}) \\ &= \sum_{l=1}^n \min_{a_l} \left\{ E_{y_{rep}|y_{obs},m} L(y_{l,rep}, a_l) + kL(y_{l,obs}, a_l) \right\}. \end{aligned} \quad (2.18)$$

Para modelos com y_l pertencente à família exponencial (2.8), Gelfand e Ghosh (1998) propõem o uso da *deviance* como função de perda:

$$\begin{aligned} L(y_l, a_l) &= 2\phi \log \frac{p(y_l | \chi(y_l); \phi)}{p(y_l | \chi(a_l); \phi)} \\ &= 2 \{y_l [\chi(y_l) - \chi(a_l)] - [b(\chi(y_l)) - b(\chi(a_l))]\} \end{aligned}$$

e, nesse caso, o critério de seleção (2.18) é dado por:

$$D_k(m) = 2 \sum_{l=1}^n [t_l^{(m)} - t(\mu_l^{(m)})] + 2(k+1) \sum_{l=1}^n \left\{ \frac{t(\mu_l^{(m)}) + kt(y_{l,obs})}{k+1} - t\left(\frac{\mu_l^{(m)} + ky_{l,obs}}{k+1}\right) \right\}, \quad (2.19)$$

em que $\mu_l^{(m)} = E[y_{l,rep}|y_{l,obs}, m]$ e $t(y) = y\chi(y) - b(\chi(y))$.

O primeiro termo à direita da equação (2.19) é visto como uma penalidade, sendo, segundo os autores, aproximadamente uma soma ponderada de variâncias preditivas. O segundo termo é visto como uma medida de qualidade do ajuste, anulando-se quando $\mu_l^{(m)} = y_{l,obs}$ e aumentando à medida em que $\mu_l^{(m)}$ se afasta de $y_{l,obs}$.

2.5.2 DIC - Deviance information criterion

Spiegelhalter et al. (2002) propõem um critério bayesiano para escolha entre modelos hierárquicos. Suponha-se que se deseje comparar diferentes modelos, com foco no parâmetro (ou vetor de parâmetros) Θ . A *deviance* é denotada por:

$$d_{\Theta} \{y, \theta, \tilde{\theta}\} = 2 \{ -\log[p(y|\theta)] + \log[p(y|\tilde{\theta})] \},$$

em que $\tilde{\theta}$ é um estimador para o parâmetro θ . No contexto bayesiano, uma escolha usual é tomar a média a posteriori, $\tilde{\theta} = E[\theta|y] = \bar{\theta}$.

Uma medida de complexidade, proposta para avaliação do número efetivo de parâmetros no modelo é dada por:

$$p_D \{y, \theta, \tilde{\theta}\} = E_{\theta|y}[d_{\Theta} \{y, \theta, \tilde{\theta}\}] = E_{\theta|y} \{ -2\log[p(y|\theta)] \} + 2\log[p(y|\tilde{\theta})] \quad (2.20)$$

$$= \overline{D(\theta)} - D(\tilde{\theta}), \quad (2.21)$$

em que, $D(\cdot)$ é a forma geral da deviance bayesiana, dada por:

$$D(\theta) = -2\log[p(y|\theta)] + 2\log[f(y)],$$

sendo $f(y)$ um fator de padronização. Segundo Spiegelhalter et al. (2002), para comparações de modelos, é suficiente tomar $f(y) = 1$.

Rearranjando a equação (2.21), tem-se:

$$\overline{D(\theta)} = D(\tilde{\theta}) + p_D.$$

Finalmente, o critério de seleção é dado por:

$$\begin{aligned} DIC &= D(\bar{\theta}) + 2p_D \\ &= \overline{D(\theta)} + p_D, \end{aligned} \quad (2.22)$$

e pode ser visto como uma medida bayesiana de adequação somada a um termo de penalidade.

2.5.3 Verossimilhança preditiva

Uma terceira forma adotada neste trabalho para seleção de um modelo baseia-se em sua capacidade preditiva para um conjunto de valores futuros $\mathbf{y}_f = (y_{T+1}, \dots, y_{T+h})$, por meio da comparação de sua verossimilhança preditiva às de outros modelos. Entre os propostos, aquele que apresente maior verossimilhança preditiva deve ser o escolhido de acordo com este critério. A verossimilhança preditiva é dada por:

$$\begin{aligned} p(\mathbf{y}_f | M, D_T) &= \int p(\mathbf{y}_f, \boldsymbol{\lambda}_f | M, D_T) d\boldsymbol{\lambda}_f \\ &= \int p(\mathbf{y}_f | \boldsymbol{\lambda}_f, M, D_T) \pi(\boldsymbol{\lambda}_f | M, D_T) d\boldsymbol{\lambda}_f \\ &= E_{\boldsymbol{\lambda}_f | M, D_T} [p(\mathbf{y}_f | \boldsymbol{\lambda}_f, M, D_T)], \end{aligned} \quad (2.23)$$

em que $\boldsymbol{\lambda}_f$ é obtido após atualização, por meio da equação de sistema, dos parâmetros dinâmicos que compõem o preditor.

Partindo-se de uma amostra $(\boldsymbol{\lambda}_f^{(1)}, \dots, \boldsymbol{\lambda}_f^{(N)})$ da posteriori $\pi(\boldsymbol{\lambda}_f | M, D_T)$, uma estimativa de Monte Carlo para a verossimilhança preditiva (2.23) é dada por:

$$\begin{aligned} \hat{E}_{\boldsymbol{\lambda}_f | M, D_T} [p(\mathbf{y}_f | \boldsymbol{\lambda}_f, M, D_T)] &= \sum_{n=1}^N \frac{p(\mathbf{y}_f | \boldsymbol{\lambda}_f^{(n)}, M, D_T)}{N} \\ &= \sum_{n=1}^N \frac{\prod_{i=1}^h p(y_{T+i} | \boldsymbol{\lambda}_{T+i}^{(n)}, M, D_T)}{N}. \end{aligned} \quad (2.24)$$

Funções de Transferência em Modelos Dinâmicos Lineares Generalizados

A principal contribuição deste trabalho é a realização de inferência bayesiana completa sobre modelos pertencentes à classe dos modelos dinâmicos lineares generalizados, originados da combinação entre diferentes propostas para a estrutura observacional - compostas por distribuições pertencentes à família exponencial - e diferentes preditores, construídos de acordo com especificações diversas de funções de transferência e do caráter estático ou dinâmico dos parâmetros envolvidos nos mesmos.

Neste capítulo, os modelos explorados são descritos, apresentando-se detalhamentos sobre aspectos computacionais envolvidos na estimação dos mesmos. O capítulo está organizado da seguinte forma: na seção 3.1, a forma geral dos modelos sobre os quais se deseja inferir é apresentada; a seção 3.2 descreve as diferentes estruturas observacionais consideradas, ditadas por distribuições pertencentes à família exponencial, em particular: Poisson, Binomial, Gama e Gaussiana Inversa; na seção 3.3, apresentam-se as formas propostas para funções de transferência utilizadas. Na seção 3.4, são detalhados os cálculos envolvidos nos algoritmos MCMC para obtenção de amostras das posteriores referentes aos modelos compostos pela combinação das estruturas observacionais e de transferência. Finalmente, os métodos para utilização das amostras obtidas via MCMC para estimação de quantidades associadas à posteriori e predição são brevemente descritos na seção 3.5.

3.1 A Estrutura Geral dos Modelos Propostos

O propósito deste trabalho é a estimação de modelos dinâmicos lineares generalizados envolvendo funções de transferência em seus preditores. Modelos de funções de transferência são bastante usuais na literatura de séries temporais e são diversas as aplicações dos mesmos, principalmente aqueles que pressupõem resposta gaussiana e que fazem uso da abordagem clássica de inferência, como em Box et al. (1994).

Do ponto de vista bayesiano, o uso de funções de transferência em modelos dinâmicos normais é descrito por West e Harrison (1997). A principal dificuldade na estimação de tais modelos reside no fato de que o preditor é não-linear nos parâmetros autoregressivos da função de transferência e, portanto, não se obtém forma analítica fechada para a posteriori. Ainda assim, uma vez que a resposta seja gaussiana, condicionando-se aos parâmetros autoregres-

sivos, tem-se um modelo dinâmico linear. Pole (1988) faz uso dessa idéia e estima modelos dinâmicos normais com funções de transferência de primeira ordem, aplicando o método da quadratura gaussiana, com grade dinâmica no parâmetro autoregressivo. Mais recentemente, Ravines et al. (2006) utilizam métodos de Monte Carlo via cadeias de Markov para inferir sobre modelos dinâmicos também com resposta normal, em que estruturas de defasagens distribuídas são comparadas a uma função de transferência de primeira ordem. Dentre as exceções aos modelos de função de transferência com resposta normal, Migon e Harrison (1985) estimam o impacto de anúncios de TV sobre a percepção de consumidores quanto ao produto anunciado, usando modelos dinâmicos lineares generalizados, com resposta binomial, estimados via *linear Bayes*.

Como visto na subseção 2.2.2, até o início da década de 90, a inferência bayesiana sobre modelos dinâmicos não-normais e/ ou não-lineares era feita em termos de obtenção de primeiros e segundos momentos ou modas a posteriori e, algumas vezes, envolvendo transformações na estrutura observacional. Atualmente, graças aos avanços computacionais e de métodos de simulação, é possível fazer inferência completa sobre modelos dinâmicos não-normais/não-lineares, a partir do paradigma de obtenção de amostras das distribuições envolvidas no processo de inferência bayesiana. Nesse contexto, no presente trabalho realiza-se toda a inferência sobre modelos dinâmicos lineares generalizados (envolvendo em seus preditores funções de transferência) por meio de métodos de Monte Carlo via cadeias de Markov. A utilização dos métodos de simulação dá-se em função de duas fontes de dificuldades para obtenção analítica da distribuição a posteriori (e, conseqüentemente, de preditivas): não-linearidade nos parâmetros autoregressivos das funções de transferência e não-normalidade da resposta, cuja distribuição pertence à família exponencial, associada a prioris não conjugadas.

A forma geral dos modelos propostos neste trabalho é dada por:

$$\begin{aligned}
 Y_t &\sim \mathcal{F}(\chi_t) \\
 g(\mu_t) &= \eta_t = \alpha_t + \delta' \mathbf{Z}_t + E_t \\
 \alpha_t &= \alpha_{t-1} + u_t, \quad u_t \sim N(0, W), \\
 E_t &= \frac{\beta(B)}{\rho(B)} X_t = v(B) X_t.
 \end{aligned} \tag{3.1}$$

em que :

- \mathcal{F} denota uma distribuição para a resposta Y_t pertencente à família exponencial, com função de probabilidade ou densidade dada por (2.8) e parâmetro natural χ_t ;
- g é uma função de ligação monótona e diferenciável;
- $\mu_t = E[Y_t | \chi_t]$;
- η_t é um preditor, não necessariamente linear, cuja estrutura básica é dada por:

α_t , um nível dinâmico, evoluindo segundo um passeio aleatório (o caso particular em que o nível é estático é obtido fazendo-se $W = 0$);

\mathbf{Z}_t , um vetor de variáveis de controle, incluindo, por exemplo, regressoras e termos seno/ cosseno para ajuste de sazonalidade;

E_t , um termo de efeitos passados e presentes de uma regressora cujo impacto sobre a resposta se deseje avaliar através do tempo.

Observe-se que os parâmetros de estado, $\alpha_1, \dots, \alpha_T$, podem ser escritos em termos dos erros de evolução u_1, \dots, u_T , para os quais pressupõe-se independência a priori. Esta parametrização traz vantagens computacionais, como mencionado na subseção 2.4.3. O vetor paramétrico a ser estimado, Θ , é composto, portanto, por u_1, \dots, u_T, W e δ , além dos parâmetros Ψ que compõem o bloco E_t . O modelo completa-se com as distribuições a priori para os parâmetros Ψ , $\pi(\Psi)$, além de $u_1 \sim N(a_1, R_1)$, $W \sim GI(\frac{n}{2}, \frac{ns}{2})$, $\delta \sim NM(\mathbf{m}_\delta, \mathbf{C}_\delta)$, NM denotando a densidade normal multivariada e GI , a gama inversa. Pressupõe-se, a priori, independência entre todos os parâmetros. Busca-se, então, aproximar a seguinte densidade a posteriori:

$$\begin{aligned} \pi(\delta, u_1, u_2, \dots, u_T, W, \Psi | D_T) &\propto \prod_{t=1}^T \exp \left\{ \frac{y_t \chi_t(u_t) - b[\chi_t(u_t)]}{\phi_t} \right\} \\ &\times \exp \left\{ -\frac{1}{2} (\delta - \mathbf{m}_\delta)' \mathbf{C}_\delta^{-1} (\delta - \mathbf{m}_\delta) \right\} \\ &\times \prod_{t=2}^T W^{-1/2} \exp \left\{ -\frac{1}{2W} u_t^2 \right\} \times \exp \left\{ -\frac{1}{2R_1} (u_1 - a_1)^2 \right\} \\ &\times W^{-(\frac{n}{2}+1)} \exp \left\{ -\frac{ns}{2} W^{-1} \right\} \pi(\Psi). \end{aligned} \quad (3.2)$$

Caso o nível α seja mantido estático, com priori $N(m_\alpha, C_\alpha)$, a expressão (3.2) reduz-se a:

$$\begin{aligned} \pi(\delta, \alpha, \Psi | D_T) &\propto \prod_{t=1}^T \exp \left\{ \frac{y_t \chi_t - b(\chi_t)}{\phi_t} \right\} \\ &\times \exp \left\{ -\frac{1}{2} (\delta - \mathbf{m}_\delta)' \mathbf{C}_\delta^{-1} (\delta - \mathbf{m}_\delta) \right\} \\ &\times \exp \left\{ -\frac{1}{2C_\alpha} (\alpha - m_\alpha)^2 \right\} \pi(\Psi). \end{aligned} \quad (3.3)$$

Mais especificamente, os modelos tratados neste trabalho são obtidos compondo-se diferentes escolhas para a estrutura observacional \mathcal{F} e variadas formas de funções de transferência, representando os efeitos E_t .

3.2 Estrutura Observacional

São consideradas as seguintes distribuições pertencentes à família Exponencial:

1. $Y_t \sim \text{Poisson}(\lambda_t)$.

No modelo Poisson, tem-se função de probabilidade dada por:

$$p(y_t | \lambda_t) = \frac{e^{-\lambda_t} \lambda_t^{y_t}}{y_t!}, \quad y_t = 0, 1, 2, \dots,$$

em que, de acordo com a notação em (2.8): $\chi_t = \log(\lambda_t)$, $b(\chi_t) = e^{\chi_t}$ e $\phi_t = 1$. Adota-se a função de ligação canônica:

$$\log(\lambda_t) = \eta_t,$$

sendo η_t um preditor, não necessariamente linear, descrito em termos do vetor paramétrico Θ , cujas componentes têm interpretação estrutural. Assim, λ_t pode ser escrito em termos de Θ - o que se denota por $\lambda_t(\theta)$ - e a função de verossimilhança é dada por:

$$p(y_1, \dots, y_T | \lambda_t(\theta)) \propto \exp \left\{ - \sum_{t=1}^T \lambda_t(\theta) \right\} \prod_{t=1}^T [\lambda_t(\theta)]^{y_t}. \quad (3.4)$$

2. $Y_t \sim \text{Binomial}(n, p_t)$

Nesse caso, definindo-se Y_t como sendo a proporção de sucessos em n ensaios de Bernoulli independentes, tem-se função de probabilidade para Y_t dada por:

$$p(y_t | p_t) = \binom{n}{ny_t} p_t^{ny_t} (1 - p_t)^{n - ny_t}, \quad y_t = 0, \frac{1}{n}, \frac{2}{n}, \dots, 1,$$

que se inclui na família exponencial com: $\chi_t = \log\left(\frac{p_t}{1-p_t}\right)$, $b(\chi_t) = \log[1 + e^{\chi_t}]$ e $\phi_t = 1/n$. Adota-se a função de ligação canônica:

$$\log\left(\frac{p_t}{1-p_t}\right) = \eta_t.$$

Escrevendo-se p_t em termos de Θ , $p_t(\theta)$, tem-se a função de verossimilhança:

$$p(y_1, \dots, y_T | p_t(\theta)) \propto \prod_{t=1}^T [p_t(\theta)]^{ny_t} [1 - p_t(\theta)]^{n - ny_t}. \quad (3.5)$$

3. $Y_t \sim \text{Gama}(\varphi, \lambda_t)$

No modelo Gama, tem-se função densidade de probabilidade:

$$f(y_t | \varphi, \lambda_t) = \frac{\lambda_t^\varphi}{\Gamma(\varphi)} e^{-\lambda_t y_t} y_t^{\varphi-1}, \quad y_t > 0,$$

para a qual, de acordo com (2.8): $\chi_t = -\frac{\lambda_t}{\varphi}$, $b(\chi_t) = -\log(-\chi_t)$ e $\phi_t = 1/\varphi$

De forma a garantir resposta esperada positiva, a função de ligação aplicada sobre $\mu_t = \frac{\varphi}{\lambda_t}$ é a logarítmica, ao invés da ligação canônica $-\mu_t^{-1}$:

$$\log\left(\frac{\varphi}{\lambda_t}\right) = \eta_t.$$

Pode-se escrever λ_t em termos do parâmetro de escala φ e dos parâmetros estruturais Θ , $\lambda_t(\varphi, \theta)$. Assim, a função de verossimilhança é dada por:

$$p(y_1, \dots, y_T | \varphi, \lambda_t(\varphi, \theta)) \propto \prod_{t=1}^T \frac{[\lambda_t(\varphi, \theta)]^\varphi}{\Gamma(\varphi)} \exp\left\{-\sum_{t=1}^T y_t \lambda_t(\varphi, \theta)\right\}. \quad (3.6)$$

4. $Y_t \sim \text{Gaussiana Inversa}(\mu_t, \sigma^2)$

A função densidade de probabilidade Gaussiana Inversa é dada por:

$$f(y_t | \mu_t, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} y_t^{-3/2} \exp\left\{-\frac{(y_t - \mu_t)^2}{2\sigma^2 \mu_t^2 y_t}\right\}, \quad y_t > 0,$$

em que: $\chi_t = -\frac{1}{2\mu_t^2}$, $b(\chi_t) = -\sqrt{-2\chi_t}$ e $\phi_t = \sigma^2$. A função de variância de Y_t é dada por $\sigma^2 \mu_t^3$ e, para garantir valores positivos para a função de resposta média, μ_t , adota-se, mais uma vez, a função de ligação logarítmica, ao invés da ligação canônica $-\frac{1}{2\mu_t^2}$:

$$\log(\mu_t) = \eta_t.$$

Escrevendo-se μ_t em termos dos parâmetros estruturais, Θ , como $\mu_t(\theta)$, tem-se a função de verossimilhança dada por:

$$p(y_1, \dots, y_T | \mu_t(\theta), \sigma^2) \propto (\sigma^2)^{-T/2} \prod_{t=1}^T y_t^{3/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=1}^T \frac{[y_t - \mu_t(\theta)]^2}{[\mu_t(\theta)]^2 y_t}\right\}. \quad (3.7)$$

3.3 Funções de Transferência Adotadas

Uma maneira óbvia para se modelar o efeito acumulado de uma regressora, através do tempo, é o emprego de defasagens, ou seja, pode-se assumir que o efeito da regressora, no tempo t , seja dado por:

$$E_t = \beta_0 X_t + \beta_1 X_{t-1} + \cdots + \beta_s X_{t-s}. \quad (3.8)$$

Observe-se que a estrutura acima é obtida como um caso particular da formulação geral de funções de transferência, vista na equação (2.12), assumindo-se $r = b = 0$. Assim constituem-se os chamados modelos de defasagens distribuídas ou modelos de função de transferência com forma livre. Entretanto, o uso da regressora defasada implica o estabelecimento do horizonte de influência, s , da mesma. Muitas vezes é difícil determinar previamente o número adequado de defasagens e, além disso, é possível que o efeito de X seja bastante persistente, o que conduziria a modelos sobreparametrizados. Considerem-se, ainda, as dificuldades de estimação devido à autocorrelação temporal inerente à regressora X nos diversos momentos de observação: $X_t, X_{t-1}, \dots, X_{t-s}$.

Nas propostas a seguir, o foco de atenção recai sobre a parametrização parcimoniosa de funções de transferência. Uma possibilidade para contornar a sobreparametrização é a adoção de restrições sobre os coeficientes da regressora em suas diversas defasagens. Este tipo de formulação pode reduzir drasticamente a quantidade de parâmetros a estimar, como se pode ver na subseção 3.3.1. Outra alternativa, descrita nas demais subseções, é obtida ao se adotar parâmetros autoregressivos na estrutura preditiva, ou seja, construindo-se funções de transferência com forma funcional definida. Assim, o período de persistência do efeito da regressora passa a ser estimado - ao invés de pré-fixado - e tem-se, ainda uma parametrização extremamente parcimoniosa. Em contrapartida, os preditores envolvidos em tais modelos são não-lineares.

3.3.1 Modelos de defasagens polinomiais

É usual restringir os coeficientes em um modelo de defasagens distribuídas, de modo que estes variem suavemente de acordo com o número de defasagens. Estas restrições proporcionam uma redução considerável no número de parâmetros a estimar. Para tanto, uma alternativa bastante usual é a construção de um modelo que imponha padrão polinomial aos coeficientes $\beta_0, \beta_1, \dots, \beta_s$. Este tipo de abordagem foi apresentada, originalmente, por Almon (1965), em um contexto econométrico. Na última década, diversos trabalhos em epidemiologia (e.g., Schwartz, 2000; Zanobetti et al., 2002) vêm utilizando defasagens polinomiais, primordialmente em modelos aditivos generalizados (Hastie e Tibshirani, 1990), em que parte do preditor (não-linear) é não-paramétrica. Nos modelos de defasagens polinomiais, o efeito da regressora em (3.8) é reescrito em termos de coeficientes $\zeta_0, \zeta_1, \dots, \zeta_d$, associados aos coeficientes originais por meio de um polinômio de grau $d \ll s$ nas defasagens, da seguinte forma:

$$\beta_j = \sum_{k=0}^d \zeta_k j^k, \quad \text{para } j=0, \dots, s, \quad (3.9)$$

em que s é o número máximo de defasagens consideradas e d é o grau do polinômio de restrição. A figura 3.1 ilustra o princípio acima.

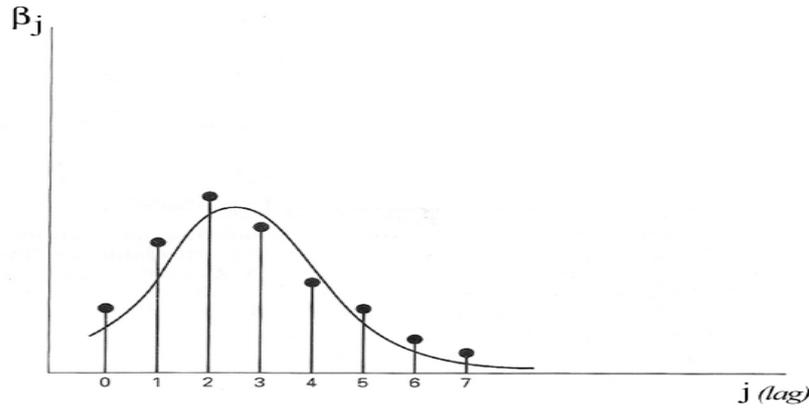


Figura 3.1 Aproximação polinomial para os coeficientes β_j de um modelo de defasagens distribuídas.

Aplicando-se as restrições (3.9) à equação (3.8), o efeito da regressora passa a ser escrito como:

$$\begin{aligned} E_t &= \zeta_0 X_t \\ &+ (\zeta_0 + \zeta_1 + \zeta_2 + \dots + \zeta_d) X_{t-1} \\ &+ (\zeta_0 + 2\zeta_1 + 2^2\zeta_2 + \dots + 2^d\zeta_d) X_{t-2} \\ &\quad \vdots \\ &+ (\zeta_0 + s\zeta_1 + s^2\zeta_2 + \dots + s^d\zeta_d) X_{t-s}. \end{aligned}$$

Finalmente, evidenciando-se os coeficientes ζ_j , $j = 0, 1, \dots, d$, tem-se:

$$\begin{aligned} E_t &= \zeta_0 (X_t + X_{t-1} + X_{t-2} + \dots + X_{t-s}) \\ &+ \zeta_1 (X_{t-1} + 2X_{t-2} + 3X_{t-3} + \dots + sX_{t-s}) \\ &+ \zeta_2 (X_{t-1} + 4X_{t-2} + 9X_{t-3} + \dots + s^2X_{t-s}) \\ &\quad \vdots \\ &+ \zeta_d (X_{t-1} + 2^d X_{t-2} + 3^d X_{t-3} + \dots + s^d X_{t-s}). \end{aligned}$$

Assim, definindo-se:

$$S_{tj} = \sum_{i=0}^s i^j X_{t-i}, \quad j = 0, 1, 2, \dots, d, \quad (3.10)$$

tem-se que o efeito da regressora X , acumulado por s períodos de tempo, é dado por:

$$E_t = \zeta_0 S_{t0} + \zeta_1 S_{t1} + \zeta_2 S_{t2} + \dots + \zeta_d S_{td}. \quad (3.11)$$

Portanto, para se aplicar restrições polinomiais aos coeficientes de um modelo de defasagens distribuídas, basta que se ajuste o modelo com parâmetros ζ_j e regressoras S_j , $j = 0, 1, 2, \dots, d$, obtidas por meio da transformação (3.10).

O vetor paramétrico sobre o qual se deseja inferir é, então: $\Theta = (u_1, \dots, u_T, W, \delta, \Psi)$, em que $\Psi = (\zeta_0, \zeta_1, \dots, \zeta_d)$. Adotando-se priori $NM(\mathbf{m}_\zeta, \mathbf{C}_\zeta)$ para Ψ , tem-se a expressão da distribuição a posteriori (3.2) conhecida a menos de uma constante de normalização. Como mencionado anteriormente, neste trabalho os efeitos (3.11) são inseridos como blocos estruturais em modelos dinâmicos lineares generalizados. Assim, embora o preditor seja linear nos parâmetros $\zeta_0, \zeta_1, \dots, \zeta_d$, se a resposta não tem distribuição gaussiana, é necessário que se utilizem métodos computacionais para determinação da constante de normalização, como os métodos de Monte Carlo via Cadeias de Markov (MCMC), descritos na seção 2.4. Uma vez que se obtenha, a partir do algoritmo MCMC, uma amostra da distribuição a posteriori de $\zeta_0, \zeta_1, \dots, \zeta_d$, aplica-se a transformação (3.9) a cada ponto da amostra, obtendo-se facilmente uma amostra da distribuição a posteriori de $\beta_0, \beta_1, \dots, \beta_s$, a partir da qual se pode determinar quantidades de interesse, tais como média, moda e quaisquer quantis.

Uma dificuldade na utilização dos modelos de defasagens distribuídas com restrição polinomial é a escolha do grau do polinômio restritivo, d . De acordo com Zanobetti et al. (2002), a imposição de muita restrição (d pequeno) pode introduzir tendência no modelo, enquanto pouca restrição produz estimativas que são ruidosas demais para serem informativas. Ainda, segundo Greene (1997), a ordem do polinômio, em geral, é bastante pequena, raramente excedendo os graus três ou quatro. No presente trabalho, o grau do polinômio restritivo é escolhido a partir dos métodos de seleção de modelos descritos na seção 2.5. Permanece, entretanto, a dificuldade para a determinação do horizonte de influência da regressora, s , o que conduz aos modelos apresentados a seguir, baseados em efeitos autoregressivos.

3.3.2 Função de transferência de ordem ($r = 1, s = 0$)

A função de transferência de primeira ordem é obtida a partir da estrutura geral (2.12), fazendo-se $r = 1$, $b = 0$ e $s = 0$, como a seguir:

$$E_t = \rho E_{t-1} + \beta X_t. \quad (3.12)$$

Resolvendo a equação a diferenças acima por substituições recursivas, tem-se:

$$E_t = \beta X_t + \rho \beta X_{t-1} + \rho^2 \beta X_{t-2} + \dots$$

Assim, a função de resposta ao impulso é dada por:

$$v_j = (\rho B)^j \beta, \quad j = 0, 1, 2, \dots$$

Apenas de forma a ilustrar os padrões possíveis a partir da estrutura acima, suponha-se $\beta > 0$. O formato das funções de resposta ao impulso e de transferência depende, então, apenas dos valores do parâmetro autoregressivo ρ . Tem-se os seguintes comportamentos para a função de resposta ao impulso:

- $0 < \rho < 1$: decaimento geométrico;
- $-1 < \rho < 0$: decaimento geométrico com alternância de sinal;
- $\rho > 1$: crescimento geométrico;
- $\rho < -1$: crescimento geométrico com alternância de sinal.

A figura 3.2 ilustra tais comportamentos, para $\beta = 1$ e diferentes valores de ρ .

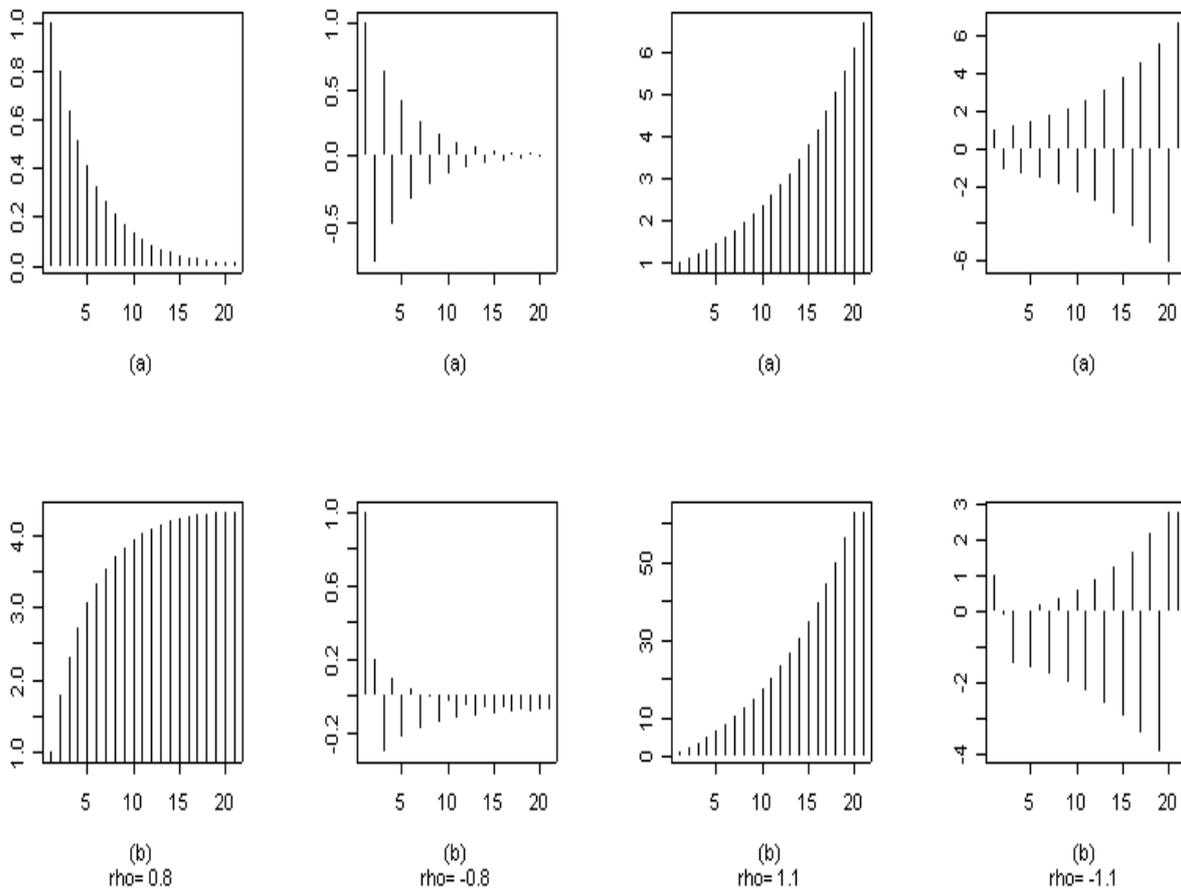


Figura 3.2 Modelo de primeira ordem: Comportamento da função de resposta ao impulso (a) e da função de transferência (b), para $|\rho| < 1$ e $|\rho| > 1$, com $\beta > 0$.

Observe-se que o modelo é estável para $|\rho| < 1$. Nesse caso, é denominado modelo de defasagens geométricas e está associado, originalmente, a Koyck (1954). O modelo geométrico preconiza o decaimento gradativo do efeito da regressora X , até que, eventualmente, tal efeito desapareça. O parâmetro autoregressivo, ρ , controla a velocidade do decaimento. Quanto mais próximo a 1 for o valor absoluto de ρ , mais persistente é o efeito de X .

Como visto, a introdução de parâmetros autoregressivos na expressão de E_t contorna os problemas levantados na subseção anterior, referentes à especificação do horizonte de influência da regressora. Em contrapartida, cria-se uma nova dificuldade, uma vez que preditores

que utilizem a forma (3.12) são não-lineares em ρ . Na subseção 2.2.2 foram levantadas algumas formas de aproximação para este tipo de problema. A abordagem adotada no presente trabalho é a obtenção de amostras da posteriori utilizando métodos de simulação estocástica via Cadeias de Markov. Como já mencionado, por meio do uso destes métodos, pode-se contornar a não normalidade da resposta e não-linearidade nos parâmetros autoregressivos das funções de transferência, fazendo-se inferência completa sobre todos parâmetros de interesse, a saber: $\Theta = (u_1, \dots, u_T, W, \delta, \Psi)$, em que $\Psi = (\rho, \beta, E_0)$. A posteriori a ser obtida via MCMC assume a forma (3.2), completada ao se assumir as prioris: $\rho \sim U(0, 1)$, $\beta \sim N(m_\beta, C_\beta)$ e $E_0 \sim N(m_E, C_E)$.

3.3.3 Função de transferência de ordem ($r = 1, s > 0$)

O decaimento no efeito da regressora, postulado pelo modelo de defasagens geométricas, pode ser postergado ao se adotar defasagens na regressora X , tendo-se, então:

$$E_t = \rho E_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_s X_{t-s}. \quad (3.13)$$

A forma acima determina funções de resposta ao impulso que apresentam um ponto de inflexão. Na figura 2.1 são exibidos alguns exemplos, para $s = 1$ e $s = 2$. A correlação intrínseca à regressora em suas diversas defasagens pode gerar dificuldades para a estimação. Em geral, entretanto, um pequeno número de defasagens em X é suficiente para se obter bons ajustes. O vetor paramétrico Θ conta, então, com as componentes do bloco de função de transferência: $\Psi = (\rho, \beta_0, \dots, \beta_s, E_0)$. Para obtenção da posteriori, assumem-se as prioris: $\rho \sim U(0, 1)$, $\beta \sim NM(\mathbf{m}_\beta, \mathbf{C}_\beta)$ e $E_0 \sim N(m_E, C_E)$. Nas aplicações desenvolvidas neste trabalho, pressupõe-se independência a priori entre os parâmetros (embora tal suposição não seja necessária), de tal forma que \mathbf{C}_β é uma matriz diagonal.

3.3.4 Função de transferência de ordem ($r=1, s=0$) com fator de ganho dinâmico

Pressupõe-se, no modelo geométrico, que o impacto imediato de uma elevação unitária na regressora, em qualquer instante t , seja constante e dado por β . Pode-se estender esta formulação, tornando o modelo bem mais flexível, admitindo-se que o impacto da regressora sobre a resposta média varie ao longo do tempo e, portanto, assumindo-se:

$$E_t = \rho E_{t-1} + \beta_t X_t. \quad (3.14)$$

Faz-se necessária, então, uma equação que governe a dinâmica de β_t . Admita-se que β_t evolua segundo um passeio aleatório:

$$\beta_t = \beta_{t-1} + v_t, \quad v_t \sim N(0, Q). \quad (3.15)$$

Tem-se, então, uma função de resposta ao impulso cujo formato e magnitude podem variar a cada instante t . A figura 3.3 exibe uma função de resposta ao impulso com fatores de ganho dinâmicos, simulada com $\rho = 0.9$ e $Q = 0.005$, nos instantes $t = 100, 300$ e 500 . O caso estático (3.12) é obtido como um caso particular deste último, fazendo-se a variância de evolução, Q , nula.

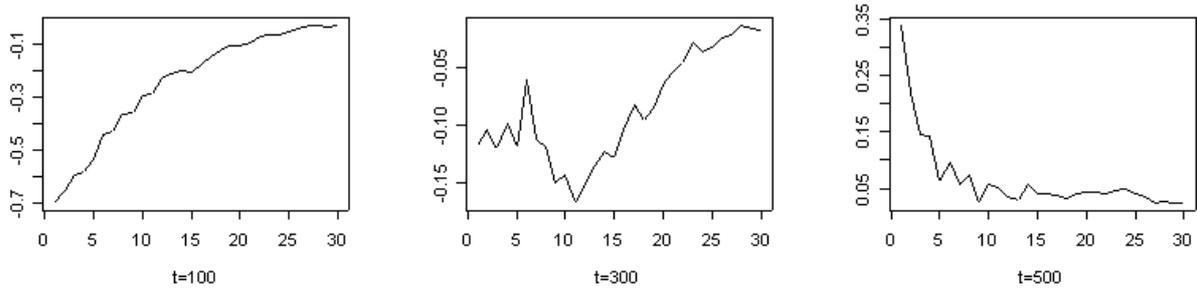


Figura 3.3 Função de resposta ao impulso com fatores de ganho evoluindo segundo um passeio aleatório, simulada com $\rho = 0.9$ e $Q = 0.005$, nos instantes 100, 300 e 500.

No processo de estimação, novamente se faz uso da proposta de Gamerman (1998), com reparametrização em termos dos erros de evolução. Assim, o vetor paramétrico correspondente ao bloco de efeitos cumulativos da regressora X é: $\Psi = (\rho, E_0, v_1, \dots, v_T, Q)$, para o qual são postuladas as seguintes distribuições a priori: $\rho \sim U(0, 1)$, $E_0 \sim N(m_E, C_E)$, $v_1 \sim N(a_v, R_v)$, $Q \sim GI(\frac{n_v}{2}, \frac{n_v s_v}{2})$.

3.3.5 Função de transferência de ordem ($r = 1, s \geq 0$) com erros aleatórios

Em praticamente todos os modelos construídos para ajuste e previsão de um processo aleatório, diversas variáveis influentes são omitidas, seja por dificuldade de mensuração destas ou mesmo por ser uma tarefa praticamente impossível enumerar todas as variáveis que influenciam a resposta de interesse. Migon (1984) propõe, para o tratamento de variáveis omitidas na modelagem, que os fatores de ganho β_t sejam estocásticos, da seguinte forma:

$$\beta_t = \bar{\beta} + \varepsilon_t, \quad \varepsilon_t \sim N(0, Q_\varepsilon).$$

Tem-se, então, a cada instante, fatores de ganho que oscilam em torno de um nível fixo $\bar{\beta}$. Na figura 3.4, pode-se ver uma função de resposta ao impulso com fatores de ganho estocásticos, simulada com $\rho = 0.9$, $\bar{\beta} = 0.03$ e $Q_\varepsilon = 0.005$, em três instantes.

Para avaliação do efeito de variáveis não incluídas no modelo, utiliza-se, no presente trabalho, um termo de erro aditivo no bloco estrutural referente à função de transferência de primeira ordem:

$$E_t = \rho E_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_s X_{t-s} + \varepsilon_t, \quad \varepsilon_t \sim N(0, Q_\varepsilon). \quad (3.16)$$

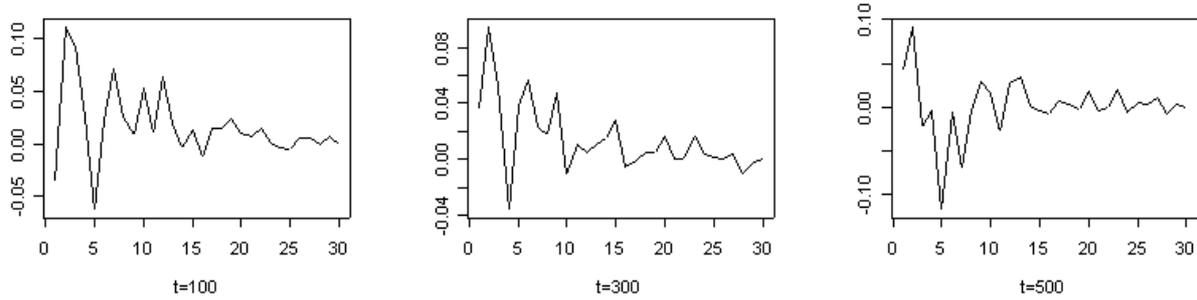


Figura 3.4 Função de resposta ao impulso com fatores de ganho estocásticos, simulada com $\rho = 0.9$, $\bar{\beta} = 0.03$ e $Q = 0.005$, nos instantes 100, 300 e 500.

A avaliação da magnitude dos erros ε_t serve como um indicativo da qualidade dos ajustes feitos com base nas regressoras inseridas formalmente no modelo, uma vez que erros "grandes" forneceriam indícios de baixo poder de explicação da estrutura proposta.

Nesse caso, embora o preditor η_t do modelo dinâmico linear generalizado torne-se mais flexível, as funções de transferência e de resposta ao impulso não variam ao longo do tempo, uma vez que os coeficientes ρ e β são fixos. Os parâmetros a estimar são: $\Psi = (\rho, \beta_0, \dots, \beta_s, E_0, \varepsilon_1, \dots, \varepsilon_T, Q_\varepsilon)$, cujas componentes têm as seguintes distribuições a priori: $\rho \sim U(0, 1)$, $\beta \sim NM(\mathbf{m}_\beta, \mathbf{C}_\beta)$, $E_0 \sim N(m_E, C_E)$, $Q_\varepsilon \sim GI(\frac{n_\varepsilon}{2}, \frac{n_\varepsilon s_\varepsilon}{2})$. Pressupõe-se independência a priori e, portanto, \mathbf{C}_β e \mathbf{C}_ε são matrizes diagonais.

3.4 Cálculos Envolvidos nos algoritmos MCMC

Como mencionado anteriormente, na classe de modelos tratada no presente trabalho, não se obtém forma analítica fechada para as densidades a posteriori dos parâmetros de interesse, recorrendo-se a métodos de Monte Carlo via Cadeias de Markov para obtenção de amostras destas. Nesta seção, detalha-se a estratégia para atualização de cada componente do vetor paramétrico Θ .

Com exceção de variâncias de evolução, os parâmetros envolvidos nos modelos tratados neste trabalho não possuem densidades condicionais completas com forma analítica fechada, sendo necessário, portanto, inserir-se passos Metropolis-Hastings no amostrador de Gibbs, o que implica a escolha de densidades propostas. Caso sejam utilizadas propostas passeio aleatório, as razões de aceitação (2.15) reduzem-se a razões de densidades a posteriori, uma vez que tais propostas são simétricas. Já para o uso das propostas baseadas na verossimilhança, sugeridas por Gamerman (1998), deve-se determinar a densidade condicional completa de cada parâmetro, a partir do modelo de trabalho (2.9).

As subseções a seguir descrevem o esquema de amostragem de todos os parâmetros envolvidos nos modelos tratados, para cada uma das estruturas observacionais adotadas: Poisson, Binomial, Gama e Gaussiana Inversa.

3.4.1 Resposta Poisson

Cada iteração do algoritmo MCMC é composta pelos seguintes passos:

1. atualiza-se o nível:
 - (a) Se o nível α é fixo, atualiza-se α ;
 - (b) Se o nível α_t é dinâmico:
 - atualiza-se $\alpha_1, \alpha_2, \dots, \alpha_T$;
 - atualiza-se a variância evolucional W ;
2. atualiza-se δ ;
3. atualiza-se cada um dos parâmetros envolvidos no vetor Ψ , que compõem o bloco de função de transferência.

Os passos acima são detalhados a seguir. Para aplicação das densidades propostas sugeridas por Gamerman (1998) no algoritmo Metropolis-Hastings, como descrito na subseção 2.4.2, observe-se que, no modelo Poisson, tem-se $\chi_t = \log(\lambda_t) = \eta_t$ e $b(\chi_t) = e^{\chi_t}$, portanto:

$$\ddot{b}(\chi_t) = e^{\chi_t} = \lambda_t;$$

$$g(\mu_t) = \log(\lambda_t) \Rightarrow \dot{g}(\mu_t) = \lambda_t^{-1}$$

Aplicando-se (2.10) e (2.11) a respostas Poisson, tem-se, então:

$$\tilde{y}_t = \log(\lambda_t) + \lambda_t^{-1}(y_t - \lambda_t); \quad (3.17)$$

$$\tilde{V}_t = \lambda_t^{-1}. \quad (3.18)$$

Atualização do Nível e dos Parâmetros do Bloco Estrutural de Controle

1. Atualização de α (nível estático)

Caso o modelo adotado tenha nível fixo, α , com priori $N(m_\alpha, C_\alpha)$, a distribuição a posteriori a ser obtida é dada por (3.3), com verossimilhança (3.4). Seja

$$\hat{y}_t = \tilde{y}_t - E_t - \delta' \mathbf{Z}_t.$$

A densidade proposta baseada na verossimilhança, para o algoritmo Metropolis-Hastings, é dada por¹:

¹ $\theta_{-\alpha}$ denota todas as componentes do vetor paramétrico θ , com exceção de α .

$$q(\alpha, \alpha_*) = \pi(\alpha_* | \alpha, \hat{y}_1, \dots, \hat{y}_T, \theta_{-\alpha}) \propto \exp \left\{ - \sum_{t=1}^T \frac{(\hat{y}_t - \alpha_*)^2}{2\tilde{V}_t} \right\} \exp \left\{ - \frac{(\alpha_* - m_\alpha)^2}{2C_\alpha} \right\}.$$

Então

$$q(\alpha, \alpha_*) \propto \exp \left\{ - \frac{1}{2} \left[\alpha_*^2 \left(C_\alpha^{-1} + \sum_{t=1}^T \frac{1}{\tilde{V}_t} \right) - 2\alpha_* \left(C_\alpha^{-1} m_\alpha + \sum_{t=1}^T \frac{\hat{y}_t}{\tilde{V}_t} \right) \right] \right\},$$

concluindo-se que

$$(\alpha_* | \alpha, \hat{y}_1, \dots, \hat{y}_T, \theta_{-\alpha}) \sim N(m_{\alpha A}, C_{\alpha A}),$$

em que:

$$C_{\alpha A} = \left(C_\alpha^{-1} + \sum_{t=1}^T \frac{1}{\tilde{V}_t} \right)^{-1};$$

$$m_{\alpha A} = C_{\alpha A} \cdot \left(C_\alpha^{-1} m_\alpha + \sum_{t=1}^T \frac{\hat{y}_t}{\tilde{V}_t} \right).$$

Denote-se por $LDN(m, C, x)$ o logaritmo do núcleo da densidade normal com média m e variância C , avaliada no ponto x , ou seja: $LDN(m, C, x) = -\frac{(x-m)^2}{2C}$. Então o logaritmo da razão de propostas² para α é dado por:

$$\log \left[\frac{q(\alpha_*, \alpha)}{q(\alpha, \alpha_*)} \right] = LDN(m_{\alpha A_*}, C_{\alpha A_*}, \alpha) - LDN(m_{\alpha A}, C_{\alpha A}, \alpha_*), \quad (3.19)$$

em que a presença ou ausência do subscrito * indica, respectivamente, se as quantidades \tilde{y}_t , \tilde{V}_t e \hat{y}_t envolvidas no cálculo de $m_{\alpha A}$ e $C_{\alpha A}$ são obtidas com base no valor proposto, α_* , ou no valor corrente, α .

Denote-se por D_T toda a informação disponível no instante T , incluindo valores passados da série temporal e informação a priori. Então o logaritmo da razão de densidades-alvo para α é dado por:

$$\log \left[\frac{\pi(\alpha_* | D_T, \theta_{-\alpha})}{\pi(\alpha | D_T, \theta_{-\alpha})} \right] = \sum_{t=1}^T (\lambda_t - \lambda_{t_*}) + \sum_{t=1}^T y_t (\alpha_* - \alpha) + LDN(m_\alpha, C_\alpha, \alpha_*) - LDN(m_\alpha, C_\alpha, \alpha). \quad (3.20)$$

Substituindo-se (3.19) e (3.20) no logaritmo de (2.15), tem-se a log-probabilidade de aceitação do valor proposto α_* .

²Os algoritmos descritos foram implementados em linguagem FORTRAN e, por razões computacionais, todas as probabilidades de aceitação são calculadas no logaritmo.

Outra alternativa que se mostra eficaz para modelos com nível fixo é a adoção de uma proposta passeio aleatório: $(\alpha_*|\alpha) \sim N(\alpha, V_\alpha)$. Com este tipo de proposta, tem-se uma simplificação computacional, uma vez que $q(\alpha_*, \alpha) = q(\alpha, \alpha_*)$ e, portanto, a razão de aceitação reduz-se à razão de condicionais completas a posteriori, cujo logaritmo é dado por (3.20). Em contrapartida, deve-se calibrar V_α , de forma a se obter taxa de aceitação em torno de 50%: variâncias muito pequenas implicariam valores propostos semelhantes aos correntes e, conseqüentemente, probabilidade de aceitação próxima a 1. Como a cadeia deve ser capaz de percorrer todo o espaço paramétrico até atingir a distribuição de equilíbrio, taxas de aceitação muito elevadas podem implicar muitas iterações até que a convergência seja obtida. Com variâncias muito grandes, por outro lado, corre-se o risco de gerar valores nas caudas da densidade-alvo, o que conduziria a valores baixos para a razão de aceitação e, conseqüentemente, a convergência lenta.

2. Atualização da variância e dos erros de evolução para modelos com nível α_t evoluindo temporalmente

Caso o nível α_t varie ao longo do tempo, devem-se obter amostras de (3.2), com verossimilhança dada por (3.4). Diretamente de (3.2), tem-se que a variância dos erros de evolução, W , possui densidade condicional completa gama inversa :

$$(W|D_T, \theta_{-W}) \sim GI\left(\frac{T+n-1}{2}, \frac{\sum_{t=2}^T u_t^2 + ns}{2}\right).$$

Conforme mencionado anteriormente, ao invés de gerar valores para os parâmetros estruturais α_t , que são extremamente correlacionados, o que torna a convergência do algoritmo MCMC lenta, procede-se de acordo com Gamerman (1998): a geração é feita sobre os erros de evolução, u_t , que, a priori, são independentes. Como o vetor de erros (u_1, \dots, u_T) em geral tem dimensão elevada, o uso de propostas passeio aleatório torna-se ineficaz, uma vez que, nesse tipo de abordagem, deve-se especificar a variância da proposta para cada ponto u_t . Da equação de evolução em (3.1), observa-se que

$$\alpha_t = \sum_{i=1}^t u_i.$$

Portanto, dependem de u_t os termos $\alpha_t, \alpha_{t+1}, \dots, \alpha_T$. A geração dos erros de evolução é feita, então, com movimentos individuais e utilizando-se como proposta a densidade condicional completa de u_t no modelo (2.9). Assim, para $t > 1$, tem-se que a geração de um valor proposto u_{t*} para u_t é feita da seguinte forma:

$$q(u_t, u_{t*}) = \pi(u_{t*}|u_t, \tilde{y}_1, \dots, \tilde{y}_T, \theta_{-u_t}) \propto \prod_{i=t}^T \exp\left\{-\frac{(\tilde{y}_i - \alpha_i - E_i - \delta' \mathbf{Z}_i)^2}{2\tilde{V}_i}\right\} \exp\left\{-\frac{u_{t*}^2}{2W}\right\}.$$

Fazendo-se

$$\hat{y}_i = \tilde{y}_i - \gamma_i - E_i - \delta' \mathbf{Z}_i,$$

com

$$\gamma_i = \alpha_i - u_{t^*} = u_1 + \cdots + u_{t-1} + u_{t+1} + \cdots + u_i,$$

a expressão acima reduz-se a:

$$\pi(u_{t^*} | u_t, \hat{y}_1, \dots, \hat{y}_T, \theta_{-u_t}) \propto \exp \left\{ \frac{-u_{t^*}^2}{2} \left(\sum_{i=t}^T \frac{1}{\tilde{V}_i} + W^{-1} \right) - u_{t^*} \sum_{i=t}^T \frac{\hat{y}_i}{\tilde{V}_i} \right\},$$

com \tilde{V}_i e \hat{y}_i calculados com base no valor corrente, u_t , e concluindo-se que, para $t > 1$, tem-se que a densidade proposta $q(u_t, u_{t^*}) = \pi(u_{t^*} | u_t, \hat{y}_1 \cdots, \hat{y}_T, \theta_{-u_t})$ é a densidade $N(m_{uA}, W_{uA})$, em que:

$$W_{uA} = \left(\sum_{i=t}^T \frac{1}{\tilde{V}_i} + W^{-1} \right)^{-1};$$

$$m_{uA} = W_{uA} \sum_{i=t}^T \frac{\hat{y}_i}{\tilde{V}_i}.$$

Para $t = 1$, tem-se:

$$W_{uA} = \left(\sum_{i=1}^T \frac{1}{\tilde{V}_i} + R_1^{-1} \right)^{-1};$$

$$m_{uA} = W_{uA} \left(\sum_{i=1}^T \frac{\hat{y}_i}{\tilde{V}_i} + R_1^{-1} a_1 \right).$$

Os valores propostos, u_{t^*} , para u_t , gerados da densidade acima, são aceitos com probabilidade

$$\min \left\{ 1, \frac{\pi(u_{t^*} | D_T, \theta_{-u_t}) / q(u_t, u_{t^*})}{\pi(u_t | D_T, \theta_{-u_t}) / q(u_{t^*}, u_t)} \right\}, \quad (3.21)$$

$$\pi(u_t | D_T, \theta_{-u_t}) \propto \prod_{i=t}^T \exp(-\lambda_i) \lambda_i^{y_i} \exp \left\{ -\frac{u_t^2}{2W} \right\}, \quad t \geq 2$$

e, para $t = 1$:

$$\pi(u_1 | D_T, \theta_{-u_1}) \propto \prod_{i=1}^T \exp(-\lambda_i) \lambda_i^{y_i} \exp \left\{ -\frac{(u_t - a_1)^2}{2R_1} \right\}.$$

Assim, a log-razão de aceitação em (3.21) é dada por:

$$\begin{aligned} & LDN(m_{uA^*}, W_{uA^*}, u_t) - LDN(m_{uA}, W_{uA}, u_{t^*}) \\ & + \sum_{i=t}^T (\lambda_i - \lambda_{i^*}) + \sum_{i=t}^T y_i [\log(\lambda_{i^*}) - \log(\lambda_i)] \\ & + LDN(0, W, u_{t^*}) - LDN(0, W, u_t), \quad t \geq 2 \end{aligned} \quad (3.22)$$

e, para $t = 1$, substitui-se $LDN(0, W, u_{t*}) - LDN(0, W, u_t)$ por:

$$LDN(a_1, R_1, u_{1*}) - LDN(a_1, R_1, u_1).$$

Uma vez atualizado o valor de u_t , este é utilizado para determinar α_t , o parâmetro estrutural no instante t .

Nos itens subseqüentes, relativos à geração dos demais parâmetros, denota-se o nível por α_t , mas deve-se ter em mente que este parâmetro pode ser estático ou dinâmico e, de acordo com sua natureza, opta-se pela geração descrita no item 1 ou no item 2 desta subseção, de forma a completar cada iteração do amostrador de Gibbs.

3. Atualização de δ

Por analogia aos itens anteriores, fazendo-se:

$$\hat{y}_t = \tilde{y}_t - \alpha_t - E_t,$$

tem-se a seguinte densidade proposta, que constituirá o núcleo de transição: $q(\delta, \delta_*) = \pi(\delta_* | \delta, \hat{y}_1, \dots, \hat{y}_T, \theta_{-\delta})$, sendo

$$(\delta_* | \delta, \hat{y}_1, \dots, \hat{y}_T, \theta_{-\delta}) \sim NM(\mathbf{m}_{\delta A}, \mathbf{C}_{\delta A}),$$

com:

$$\mathbf{C}_{\delta A} = \left(\mathbf{C}_{\delta}^{-1} + \sum_{t=1}^T \frac{\mathbf{Z}'_t \mathbf{Z}_t}{\tilde{V}_t} \right)^{-1}$$

e

$$\mathbf{m}_{\delta A} = \mathbf{C}_{\delta A} \left(\mathbf{C}_{\delta}^{-1} \mathbf{m}_{\delta} + \sum_{t=1}^T \frac{\hat{y}_t \mathbf{Z}_t}{\tilde{V}_t} \right).$$

Assim, denotando-se $LDNM(\mathbf{m}, \mathbf{C}, \mathbf{x}) = -\frac{1}{2}(\mathbf{x} - \mathbf{m})' \mathbf{C}^{-1}(\mathbf{x} - \mathbf{m})$, o logaritmo da razão de propostas para δ é dado por:

$$\log \left[\frac{q(\delta_*, \delta)}{q(\delta, \delta_*)} \right] = LDNM(\mathbf{m}_{\delta A^*}, \mathbf{C}_{\delta A^*}, \delta) - LDNM(\mathbf{m}_{\delta A}, \mathbf{C}_{\delta A}, \delta_*). \quad (3.23)$$

e o logaritmo da razão de densidades-alvo para δ é dado por:

$$\begin{aligned} \log \left[\frac{\pi(\delta_* | D_T, \theta_{-\delta})}{\pi(\delta | D_T, \theta_{-\delta})} \right] &= \sum_{t=1}^T (\lambda_t - \lambda_{t*}) + \sum_{t=1}^T y_t \mathbf{Z}'_t (\delta_* - \delta) \\ &+ LDNM(\mathbf{m}_{\delta}, \mathbf{C}_{\delta}, \delta_*) - LDNM(\mathbf{m}_{\delta}, \mathbf{C}_{\delta}, \delta). \end{aligned} \quad (3.24)$$

O valor δ_* é aceito com log-probabilidade dada pelo mínimo entre 0 e a soma das parcelas (3.23) e (3.24).

Atualização dos Parâmetros Envolvidos nas Funções de Transferência

1. Função de Transferência de ordem ($r=1, s$)

Se os efeitos de X são determinados por (3.13), então o vetor paramétrico a ser atualizado é $\Psi = (E_0, \rho, \beta_0, \beta_1, \dots, \beta_s)$.

(a) Atualização de E_0

Observando-se que:

$$E_t = \rho^t E_0 + \beta' \mathbf{F}_{\mathbf{X}_t},$$

com

$$\beta' = (\beta_0, \beta_1, \dots, \beta_s)$$

e

$$\mathbf{F}_{\mathbf{X}_t} = \sum_{i=1}^t \rho^{t-i} \mathbf{X}_i, \quad \text{com } \mathbf{X}_i' = (X_i, X_{i-1}, \dots, X_{i-s}), \quad (3.25)$$

tem-se:

$$\eta_t = \log(\lambda_t) = \alpha_t + \rho^t E_0 + \beta' \mathbf{F}_{\mathbf{X}_t} + \delta' \mathbf{Z}_t.$$

Portanto, fazendo-se

$$\hat{y}_t = \tilde{y}_t - \alpha_t - \beta' \mathbf{F}_{\mathbf{X}_t} - \delta' \mathbf{Z}_t,$$

tem-se a densidade proposta baseada na verossimilhança para E_0 , dada por $q(E_0, E_{0*}) = \pi(E_{0*} | E_0, \hat{y}_1, \dots, \hat{y}_T, \theta_{-E_0})$, em que, por analogia aos itens anteriores:

$$(E_{0*} | E_0, \hat{y}_1, \dots, \hat{y}_T, \theta_{-E_0}) \sim N(m_{EA}, C_{EA}),$$

com

$$C_{EA} = \left(C_E^{-1} + \sum_{t=1}^T \frac{\rho^{2t}}{\tilde{V}_t} \right)^{-1};$$

$$m_{EA} = C_{EA} \left(C_E^{-1} m_E + \sum_{t=1}^T \frac{\hat{y}_t \rho^t}{\tilde{V}_t} \right).$$

Assim, o logaritmo da razão de propostas para E_0 é dado por:

$$\log \left[\frac{q(E_{0*}, E_0)}{q(E_0, E_{0*})} \right] = LDN(m_{EA*}, C_{EA*}, E_0) - LDN(m_{EA}, C_{EA}, E_{0*}). \quad (3.26)$$

O logaritmo da razão de densidades-alvo é dado por:

$$\log \left[\frac{\pi(E_{0*}|D_T, \theta_{-E_0})}{\pi(E_0|D_T, \theta_{-E_0})} \right] = \sum_{t=1}^T (\lambda_t - \lambda_{t*}) + \sum_{t=1}^T y_t \rho^t (E_{0*} - E_0) + LDN(m_E, C_E, E_{0*}) - LDN(m_E, C_E, E_0). \quad (3.27)$$

O valor E_{0*} é aceito com log-probabilidade dada pelo mínimo entre 0 e a soma das parcelas (3.26) e (3.27).

Alternativamente, pode-se adotar uma proposta passeio aleatório: $(E_{0*}|E_0) \sim N(E_0, V_E)$ e, nesse caso, como se tem uma proposta simétrica, o logaritmo da razão de aceitação é dado apenas por (3.27).

(b) *Atualização de β*

Fazendo-se

$$\hat{y}_t = \tilde{y}_t - \alpha_t - \rho^t E_0 - \delta' \mathbf{Z}_t,$$

tem-se que:

$$(\beta_* | \beta, \hat{y}_1, \dots, \hat{y}_T, \theta_{-\beta}) \sim NM(\mathbf{m}_{\beta A}, \mathbf{C}_{\beta A}),$$

onde

$$\mathbf{C}_{\beta A} = \left(\mathbf{C}_{\beta}^{-1} + \sum_{t=1}^T \frac{\mathbf{F}_{\mathbf{X}}'(t) \mathbf{F}_{\mathbf{X}}(t)}{\tilde{V}_t} \right)^{-1};$$

$$\mathbf{m}_{\beta A} = \mathbf{C}_{\beta A} \left(\mathbf{C}_{\beta}^{-1} \mathbf{m}_{\beta} + \sum_{t=1}^T \frac{\hat{y}_t \mathbf{F}_{\mathbf{X}}(t)}{\tilde{V}_t} \right),$$

concluindo-se que o logaritmo da razão de propostas para β é dado por:

$$\log \left[\frac{q(\beta_*, \beta)}{q(\beta, \beta_*)} \right] = LDNM(\mathbf{m}_{\beta A*}, \mathbf{C}_{\beta A*}, \beta) - LDNM(\mathbf{m}_{\beta A}, \mathbf{C}_{\beta A}, \beta_*) \quad (3.28)$$

e logaritmo da razão de densidades-alvo para β é dado por:

$$\log \left[\frac{\pi(\beta_* | D_T, \theta_{-\beta})}{\pi(\beta | D_T, \theta_{-\beta})} \right] = \sum_{t=1}^T (\lambda_t - \lambda_{t*}) + \sum_{t=1}^T y_t \mathbf{F}_{\mathbf{X}}'(t) (\beta_* - \beta) + LDNM(\mathbf{m}_{\beta}, \mathbf{C}_{\beta}, \beta_*) - LDNM(\mathbf{m}_{\beta}, \mathbf{C}_{\beta}, \beta), \quad (3.29)$$

podendo-se então determinar a log-razão de aceitação para o valor proposto β_* somando-se (3.28) e (3.29).

Se $s = 0$, então $\beta = \beta_0$ é escalar e, nesse caso, a proposta passeio aleatório, $(\beta_* | \beta) \sim N(\beta, V_{\beta})$, conduz rapidamente à convergência, desde que tenha variância V_{β} bem calibrada, como discutido anteriormente. A razão de aceitação reduz-se, então, a (3.29).

(c) *Atualização de ρ*

Seja $\xi = \text{logito}(\rho) = \log\left(\frac{\rho}{1-\rho}\right)$. Utilizando-se a proposta passeio aleatório para ξ , tem-se: $(\xi_*|\xi) \sim N(\xi, V_\xi)$, com V_ξ arbitrada de forma a se obter taxa de aceitação para ξ em torno de 50%.

Como a proposta é simétrica, a razão de aceitação para ξ reduz-se à razão entre a densidade-alvo avaliada no ponto proposto e no ponto corrente. Como a densidade-alvo é a condicional completa de ξ , dada por:

$$\pi(\xi|D_T, \theta_{-\rho}) = \pi(\rho|D_T, \theta_{-\rho}) \frac{d\rho}{d\xi}$$

e

$$\frac{d\rho}{d\xi} = \frac{e^\xi}{(1+e^\xi)^2},$$

tem-se o logaritmo da razão de densidades-alvo dado por:

$$\begin{aligned} \log \left[\frac{\pi(\xi_*|D_T, \theta_{-\rho})}{\pi(\xi|D_T, \theta_{-\rho})} \right] &= \sum_{t=1}^T (\lambda_t - \lambda_{t*}) + \sum_{t=1}^T y_t (E_{t*} - E_t) \\ &+ \xi_* - \xi + 2\log(1+e^\xi) - 2\log(1+e^{\xi_*}). \end{aligned} \quad (3.30)$$

2. Modelo de Defasagens Polinomiais

O modelo de defasagens polinomiais, definido pela expressão (3.8), sujeita às restrições (3.9), pode ser tratado como um caso particular do modelo de ordem $(r = 1, s)$: basta fixar $\rho = E_0 = 0$ e gerar os parâmetros ζ exatamente como os parâmetros β , substituindo-se as regressoras X_t, X_{t-s} por S_{t0}, \dots, S_{td} , dadas por (3.10). Uma vez obtida uma amostra dos parâmetros ζ , determina-se uma amostra de β aplicando-se (3.9) a cada ponto amostral de ζ .

3. Modelo de Primeira Ordem com Fatores de Ganho Dinâmicos

Se o bloco estrutural de efeitos passados e presentes de X é determinado por (3.14) e (3.15), procede-se da seguinte forma: define-se uma matriz \mathbf{F}_X , cujos elementos são dados por

$$F_X(t, i) = \begin{cases} \sum_{j=i}^t \rho^{t-j} X_j, & i \leq t \\ 0, & i > t \end{cases} = \begin{pmatrix} X_1 & 0 & 0 & \cdots & 0 \\ \rho X_1 + X_2 & X_2 & 0 & \cdots & 0 \\ \rho^2 X_1 + \rho X_2 + X_3 & \rho X_2 + X_3 & X_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \end{pmatrix}$$

Aplicando-se (3.15) a (3.14), observa-se que o efeito de X no tempo t pode ser escrito em termos dos erros de evolução v_1, \dots, v_t como:

$$E_t = \rho^t E_0 + \sum_{i=1}^t v_i F_X(t, i). \quad (3.31)$$

(a) *Atualização de E_0*

A atualização de E_0 é feita como descrito no item 1a, referente a funções de transferência de ordem ($r = 1, s$), modificando-se apenas o cálculo de E_t , que deve ser feito como em (3.31) e, assim, \hat{y}_t , deve ser expresso por:

$$\hat{y}_t = \tilde{y}_t - \alpha_t - \sum_{i=1}^t v_i F_X(t, i) - \delta' \mathbf{Z}_t.$$

(b) *Atualização de ρ*

A atualização de ρ permanece como descrito no item 1c, bastando calcular-se λ_t utilizando os efeitos dados por (3.31).

(c) *Atualização de v_t , $t = 1, 2, \dots, T$*

Seguindo a proposta de Gamerman (1998), os parâmetros estruturais β_t são escritos em termos de seus erros de evolução: $\beta_t = v_1 + \dots + v_t$. Define-se, então:

$$\hat{y}_i = \tilde{y}_i - \alpha_i - E_i + v_t F_X(i, t) - \delta' \mathbf{Z}_i$$

e obtém-se a seguinte distribuição proposta: para $t > 1$, $q(v_t, v_{t*})$ é a densidade $N(m_{vA}, Q_{vA})$, com

$$Q_{vA} = \left(\sum_{i=t}^T \frac{F_X^2(i, t)}{\tilde{V}_i} + Q^{-1} \right)^{-1},$$

$$m_{vA} = Q_{vA} \sum_{i=t}^T \frac{\hat{y}_i F_X(i, t)}{\tilde{V}_i}$$

e, para $t = 1$, tem-se:

$$Q_{vA} = \left(\sum_{i=1}^T \frac{F_X^2(i, 1)}{\tilde{V}_i} + R_v^{-1} \right)^{-1},$$

$$m_{vA} = Q_{vA} \left(\sum_{i=1}^T \frac{\hat{y}_i F_X(i, 1)}{\tilde{V}_i} + R_v^{-1} a_v \right).$$

Os valores propostos, v_{t*} , para v_t , gerados da densidade acima, são aceitos com probabilidade

$$\min \left\{ 1, \frac{\pi(v_{t*}|D_T, \theta_{-v_t})/q(v_t, v_{t*})}{\pi(v_t|D_T, \theta_{-v_t})/q(v_{t*}, v_t)} \right\},$$

em que, de acordo com o modelo de Poisson, tem-se, para $t > 1$:

$$\pi(v_t|D_T, \theta_{-v_t}) \propto \prod_{i=t}^T \exp(-\lambda_i) \lambda_i^{y_i} \exp \left\{ -\frac{v_t^2}{2Q} \right\}$$

e para $t = 1$:

$$\pi(v_1|D_T, \theta_{-v_1}) \propto \prod_{i=1}^T \exp(-\lambda_i) \lambda_i^{y_i} \exp \left\{ -\frac{(v_1 - a_v)^2}{2R_v} \right\}.$$

(d) Atualização de Q

Como, a priori, $Q \sim GI(\frac{n_v}{2}, \frac{n_v s_v}{2})$ e $v_2, \dots, v_T \stackrel{iid}{\sim} N(0, Q)$, conclui-se que:

$$(Q|D_T, \theta_{-Q}) \sim GI \left(\frac{T + n_v - 1}{2}, \frac{\sum_{t=2}^T v_t^2 + n_v s_v}{2} \right).$$

4. Função de Transferência de ordem ($r=1, s$) com erros aleatórios

Suponha-se que a forma escolhida para refletir efeitos passados e presentes da regressora X seja dada por (3.16). Nesse caso, deve-se estimar o vetor $\Psi = (\rho, \beta_0, \dots, \beta_s, E_0, \varepsilon_1, \dots, \varepsilon_T, Q_\varepsilon)$.

(a) Atualização de E_0

Procede-se como no item 1a, alterando-se apenas o cálculo de E_t , segundo (3.16), e de \hat{y}_t , agora dado por:

$$\hat{y}_t = \tilde{y}_t - \alpha_t - \beta' \mathbf{F}_{X_t} - \delta' \mathbf{Z}_t - \varepsilon_t.$$

(b) *Atualização de β*

Segue-se o procedimento descrito no item 1b, modificando-se o cálculo de E_t de acordo com (3.16) e de \hat{y}_t , que passa ser expresso por:

$$\hat{y}_t = \tilde{y}_t - \alpha_t - \rho^t E_0 - \delta' \mathbf{Z}_t - \varepsilon_t.$$

(c) *Atualização de ρ*

A atualização de ρ segue o procedimento descrito no item 1c, modificando-se apenas, para obtenção de λ_t , o cálculo dos efeitos de X , dados por (3.16).

(d) *Atualização de ε_t , $t = 1, 2, \dots, T$*

Observe-se que cada erro ε_t encontra-se presente apenas no instante t e que:

$$E_t = \rho^t E_0 + \beta' \mathbf{F}_{\mathbf{X}_t} + \varepsilon_t,$$

com $\beta' = (\beta_0, \beta_1, \dots, \beta_s)$ e $\mathbf{F}_{\mathbf{X}_t}$ como em (3.25). Definindo-se, então:

$$\hat{y}_t = \tilde{y}_t - \alpha_t - \rho^t E_0 - \beta' \mathbf{F}_{\mathbf{X}_t} - \delta' \mathbf{Z}_t,$$

tem-se, por analogia aos itens anteriores, que a densidade proposta baseada na verossimilhança é dada por:

$$q(\varepsilon_t, \varepsilon_{t*}) = \pi(\varepsilon_{t*} | \varepsilon_t, \hat{y}_1, \dots, \hat{y}_T, \theta_{-\varepsilon_t}) \propto \exp \left\{ -\frac{(\hat{y}_t - \varepsilon_{t*})^2}{2\tilde{V}_t} \right\} \exp \left\{ -\frac{\varepsilon_{t*}^2}{2Q_\varepsilon} \right\},$$

concluindo-se que

$$(\varepsilon_{t*} | \varepsilon_t, \hat{y}_1, \dots, \hat{y}_T, \theta_{-\varepsilon_t}) \sim N(m_{\varepsilon A}, Q_{\varepsilon A}),$$

com

$$Q_{\varepsilon A} = \left(\frac{1}{\tilde{V}_t} + \frac{1}{Q_\varepsilon} \right)^{-1},$$

$$m_{\varepsilon A} = Q_{\varepsilon A} \left(\frac{\hat{y}_t}{\tilde{V}_t} \right).$$

Portanto, o logaritmo da razão de propostas para ε_t é dado por:

$$\log \left[\frac{q(\varepsilon_{t*}, \varepsilon_t)}{q(\varepsilon_t, \varepsilon_{t*})} \right] = LDN(m_{\varepsilon A*}, Q_{\varepsilon A*}, \varepsilon_t) - LDN(m_{\varepsilon A}, Q_{\varepsilon A}, \varepsilon_{t*})$$

e o logaritmo da razão de densidades-alvo para ε_t é dado por:

$$\begin{aligned} \log \left[\frac{\pi(\varepsilon_{t*} | D_T, \theta_{-\varepsilon})}{\pi(\varepsilon_t | D_T, \theta_{-\varepsilon})} \right] &= \lambda_t - \lambda_{t*} + y_t(\varepsilon_{t*} - \varepsilon_t) \\ &+ LDN(0, Q_\varepsilon, \varepsilon_{t*}) \\ &- LDN(0, Q_\varepsilon, \varepsilon_t). \end{aligned} \quad (3.32)$$

(e) *Atualização de Q_ε*

Como, a priori, $Q_\varepsilon \sim GI\left(\frac{n_\varepsilon}{2}, \frac{n_\varepsilon s_\varepsilon}{2}\right)$ e $\varepsilon_1, \dots, \varepsilon_T \stackrel{iid}{\sim} N(0, Q_\varepsilon)$, conclui-se que:

$$(Q_\varepsilon | D_T, \theta_{-Q_\varepsilon}) \sim GI\left(\frac{T + n_\varepsilon}{2}, \frac{n_\varepsilon s_\varepsilon + \sum_{t=1}^T \varepsilon_t^2}{2}\right).$$

3.4.2 Resposta Binomial

Se a resposta tem distribuição Binomial(n, p_t), tem-se a seguinte estrutura preditiva:

$$g(\mu_t) = \log\left(\frac{p_t}{1-p_t}\right) = \eta_t = \alpha_t + \delta' \mathbf{Z}_t + E_t, \quad (3.33)$$

com nível dinâmico, α_t , ou estático, $\alpha_t = \alpha \forall t$, e E_t assumindo as diferentes formas descritas na seção 3.3. Para implementação do algoritmo MCMC, procede-se como na subseção 3.4.1, em que os cálculos para respostas Poisson são detalhados, efetuando-se apenas algumas modificações, descritas a seguir.

Inicialmente, devem-se alterar as parcelas referentes à verossimilhança no algoritmo Metropolis-Hastings, segundo (3.5). Especificamente, nas log-razões de densidades-alvo (3.20), (3.22), (3.24), (3.27), (3.29), (3.30), (3.32), as parcelas referentes à log-verossimilhança devem ser substituídas por:

$$n \sum_{t=1}^T y_t (\eta_{t*} - \eta_t) + n \sum_{t=1}^T [\log(1 + e^{\eta_t}) - \log(1 + e^{\eta_{t*}})],$$

a presença ou ausência do subscrito * indicando, mais uma vez, se o cálculo do preditor η_t baseia-se no valor proposto para cada um dos parâmetros de interesse ou em seu valor corrente.

Além disso, para a adoção de densidades propostas baseadas na verossimilhança, observa-se que, no modelo binomial: $\chi_t = \log\left(\frac{p_t}{1-p_t}\right)$, $b(\chi_t) = \log[1 + e^{\chi_t}]$ e $\phi_t = 1/n$. Mantém-se então a estrutura de regressão, alterando-se \tilde{y}_t e \tilde{V}_t no modelo de trabalho (2.9), que passa a ser obtido a partir de:

$$\mu_t = E[Y_t | \chi_t] = \dot{b}(\chi_t) = \frac{e^{\chi_t}}{1 + e^{\chi_t}} = p_t \Rightarrow \ddot{b}(\chi_t) = \frac{e^{\chi_t}}{(1 + e^{\chi_t})^2} = p_t(1 - p_t).$$

De (3.33), tem-se que:

$$\dot{g}(\mu_t) = \frac{1}{p_t(1 - p_t)}.$$

Portanto, seguindo-se (2.10) e (2.11), tem-se:

$$\tilde{y}_t = \log\left(\frac{p_t}{1-p_t}\right) + \frac{y_t - p_t}{p_t(1-p_t)}; \quad (3.34)$$

$$\tilde{V}_t = \frac{p_t(1-p_t)}{n[p_t(1-p_t)]^2} = \frac{1}{np_t(1-p_t)}. \quad (3.35)$$

3.4.3 Resposta Gama

Para modelos com resposta Gama(φ, λ_t), as propostas baseadas no modelo de trabalho para \tilde{y}_t , (2.9), são obtidas observando-se que $\chi_t = -\frac{\lambda_t}{\varphi}$, $b(\chi_t) = -\log(-\chi_t)$, $\phi_t = 1/\varphi$ e, portanto, uma vez que se adote função de ligação logarítmica, tem-se:

$$\mu_t = E[Y_t|\chi_t] = \dot{b}(\chi_t) = -\frac{1}{\chi_t} = \frac{\varphi}{\lambda_t} \Rightarrow \ddot{b}(\chi_t) = \frac{1}{\chi_t^2} = \frac{\varphi^2}{\lambda_t^2};$$

$$\dot{g}(\mu_t) = \frac{1}{\mu_t} = \frac{\lambda_t}{\varphi}.$$

Portanto, de (2.10) e (2.11) tem-se:

$$\tilde{y}_t = \log\left(\frac{\varphi}{\lambda_t}\right) + \left(y_t - \frac{\varphi}{\lambda_t}\right) \frac{\lambda_t}{\varphi};$$

$$\tilde{V}_t = \frac{\varphi^2 \lambda_t^2}{\lambda_t^2 \varphi^3} = \frac{1}{\varphi}.$$

A função de verossimilhança é dada por (3.6) e, portanto, ao se atualizar qualquer das componentes do vetor $\Theta' = (\alpha_1, \dots, \alpha_T, W, \delta, \Psi)$, a log-razão de verossimilhanças é dada por:

$$\begin{aligned} \log \left[\frac{p(y_1, \dots, y_T | \varphi, \lambda_{t^*}(\varphi, \theta_*))}{p(y_1, \dots, y_T | \varphi, \lambda_t(\varphi, \theta))} \right] &= \varphi \sum_{t=1}^T \log(y_t \lambda_{t^*}(\varphi, \theta_*)) - \varphi \sum_{t=1}^T \log(y_t \lambda_t(\varphi, \theta)) \\ &+ \sum_{t=1}^T y_t (\lambda_t(\varphi, \theta) - \lambda_{t^*}(\varphi, \theta_*)). \end{aligned} \quad (3.36)$$

Assim, para a geração de Θ , utilizam-se os procedimentos descritos na subseção 3.4.1, fazendo-se as substituições supracitadas.

Uma vez atualizado Θ , cada iteração do amostrador de Gibbs é completada ao se atualizar o parâmetro de forma, φ , o que é feito da seguinte maneira:

Atualização de φ

Seja $\varpi = \log(\varphi)$. Utiliza-se uma proposta passeio-aleatório para ϖ : $(\varpi_*|\varpi) \sim N(\varpi, V_\varpi)$, com V_ϖ arbitrada de forma a se obter taxa de aceitação em torno de 50% para ϖ . Como tal proposta é simétrica, a razão de aceitação para φ reduz-se à razão de condicionais completas a posteriori. Assumindo-se priori $\varphi \sim \text{Gama}(a, l)$ e considerando-se o Jacobiano da transformação, $\frac{d\varphi}{d\varpi} = e^\varpi$, obtém-se a log-razão de aceitação fazendo-se $\varphi = e^\varpi$ em (3.36) e somando-se àquela expressão a log-razão de prioris. Portanto:

$$\begin{aligned} \log \left[\frac{\pi(\varpi_*|D_T, \theta)}{\pi(\varpi|D_T, \theta)} \right] &= e^{\varpi_*} \sum_{t=1}^T \log(y_t \lambda_{t*}(\varpi_*, \theta)) - e^\varpi \sum_{t=1}^T \log(y_t \lambda_t(\varpi, \theta)) \\ &+ T[\log \Gamma(e^\varpi) - \log \Gamma(e^{\varpi_*})] + \sum_{t=1}^T y_t (\lambda_t(\varpi, \theta) - \lambda_{t*}(\varpi_*, \theta)) \\ &+ l(e^\varpi - e^{\varpi_*}) + a(\varpi_* - \varpi). \end{aligned} \quad (3.37)$$

Uma vez atualizado o valor de ϖ , obtém-se o valor atualizado $\varphi = e^\varpi$.

3.4.4 Resposta Gaussiana Inversa

Se a resposta tem distribuição Gaussiana Inversa(μ_t, σ^2), tem-se que $\chi_t = -\frac{1}{2\mu_t^2}$, $b(\chi_t) = -\sqrt{-2\chi_t}$ e $\phi_t = \sigma^2$. Portanto, as propostas baseadas no modelo de trabalho para \tilde{y}_t , (2.9), são obtidas observando-se que, ao se adotar função de ligação logaritmica, tem-se:

$$\begin{aligned} \mu_t = E[Y_t|\chi_t] = \dot{b}(\chi_t) &= \frac{1}{\sqrt{(-2\chi_t)}} = \mu_t \Rightarrow \ddot{b}(\chi_t) = (-2\chi_t)^{-\frac{3}{2}} = \mu_t^3; \\ \dot{g}(\mu_t) &= \frac{1}{\mu_t}. \end{aligned}$$

Portanto, de (2.10) e (2.11) tem-se:

$$\begin{aligned} \tilde{y}_t &= \log(\mu_t) + \frac{y_t - \mu_t}{\mu_t}; \\ \tilde{V}_t &= \frac{\sigma^2 \mu_t^3}{\mu_t^2} = \sigma^2 \mu_t. \end{aligned}$$

A função de verossimilhança é dada por (3.7) e, portanto, a log-razão de verossimilhanças, ao se atualizar componentes de $\Theta' = (\alpha_1, \dots, \alpha_T, W, \delta, \Psi)$ é dada por:

$$\log \left[\frac{p(y_1, \dots, y_T | \sigma^2, \mu_{t*}(\theta_*))}{p(y_1, \dots, y_T | \sigma^2, \mu_t(\theta))} \right] = \frac{1}{2\sigma^2} \sum_{t=1}^T \frac{[y_t - \mu_t(\theta)]^2}{y_t \mu_t^2(\theta)} - \frac{[y_t - \mu_{t*}(\theta_*)]^2}{y_t \mu_{t*}^2(\theta_*)} \quad (3.38)$$

Consideram-se as alterações acima e gera-se $\Theta' = (\alpha_1, \alpha_T, \delta, \Psi)$ como descrito na subseção 3.4.1. Além disso, deve-se, a cada iteração do amostrador de Gibbs, atualizar o parâmetro de escala, σ^2 , como a seguir.

Atualização de σ^2

Seja $\vartheta = \log(\sigma^2)$. Utilizando-se uma proposta passeio-aleatório para ϑ : $(\vartheta_* | \vartheta) \sim N(\vartheta, V_\vartheta)$, a razão de aceitação para ϑ reduz-se à razão de condicionais completas a posteriori. Assumindo-se priori $\sigma^2 \sim \text{Gama}(a_\sigma, l_\sigma)$ e considerando-se o Jacobiano da transformação, $\frac{d\sigma^2}{d\vartheta} = e^\vartheta$, obtém-se a log-razão de aceitação fazendo-se $\sigma^2 = e^\vartheta$ em (3.38) e somando-se àquela expressão a log-razão de priors:

$$\log \left[\frac{\pi(\vartheta_* | D_T, \theta)}{\pi(\vartheta | D_T, \theta)} \right] = \frac{T}{2} \left[\log(e^\vartheta) - \log(e^{\vartheta_*}) \right] + \left(\frac{1}{2e^\vartheta} - \frac{1}{2e^{\vartheta_*}} \right) \sum_{t=1}^T \frac{[y_t - \mu_t(\theta)]^2}{y_t \mu_t^2(\theta)} + l_\sigma(e^{\bar{\omega}} - e^{\bar{\omega}_*}) + a_\sigma(\vartheta_* - \vartheta).$$

Finalmente, obtém-se $\sigma^2 = e^\vartheta$.

3.5 Utilização das Amostras da Posteriori para Estimação e Predição

Denote-se por θ um componente do vetor paramétrico de interesse. Após convergência dos procedimentos MCMC, suponha-se que sejam selecionados N pontos entre os valores gerados para θ , compondo-se, então, a amostra: $\theta' = (\theta_1, \dots, \theta_N)$. Devido à autocorrelação intrínseca às cadeias de Markov, é usual tomar-se os pontos referentes a iterações em intervalos espaçados, de forma a se ter uma amostra de valores aproximadamente independentes.

Intervalos de credibilidade para θ podem ser estimados facilmente a partir da obtenção de quantis adequados na amostra ordenada $(\theta_{(1)}, \dots, \theta_{(N)})$, $\theta_{(1)} = \min(\theta_1, \dots, \theta_N)$ e $\theta_{(N)} = \max(\theta_1, \dots, \theta_N)$. Intervalos de credibilidade para uma função real $t(\theta)$ são construídos a partir de uma amostra de $t(\theta)$, obtida, simplesmente, avaliando-se t em cada ponto θ_n , $n = 1, \dots, N$.

A esperança a posteriori de $t(\theta)$ é dada por

$$E[t(\theta) | D_T] = \int t(\theta) \pi(\theta | D_T) d\theta$$

e, como é usual nos procedimentos de Monte Carlo, tal quantidade é estimada por:

$$\hat{E}[t(\theta) | D_T] = \bar{t}_N = \frac{\sum_{n=1}^N t(\theta_n)}{N}. \quad (3.39)$$

Já para a avaliação do erro-padrão de simulação associado à estimativa acima, devem-se considerar as autocorrelações inerentes a cadeias de Markov, definindo-se:

$$\gamma_k = \text{Cov}[t(\theta_n), t(\theta_{n+k})]; \quad \rho_k = \frac{\gamma_k}{\gamma_0}, \text{ com } \gamma_0 = V[\theta_n];$$

sendo γ_k obtido a partir da amostra da distribuição a posteriori de θ . A variância de \bar{t}_N é dada, então, por τ_N^2/N , com:

$$\tau_N^2 = \gamma_0 \left(1 + 2 \sum_{k=1}^{N-1} \frac{N-k}{N} \rho_k \right). \quad (3.40)$$

Quando $N \rightarrow \infty$, $\tau_N^2 \rightarrow \tau^2$, em que:

$$\tau^2 = \gamma_0 \left(1 + 2 \sum_{k=1}^{\infty} \rho_k \right),$$

desde que a soma de autocorrelações seja convergente. Sob independência, as expressões acima reduzem-se a γ_0 e, como ressalta Gamerman (1997), é importante diferenciar entre $\gamma_0 = V[\theta_n]$ - uma medida de incerteza da posteriori de θ - e τ_N^2/N , esta última uma medida de variabilidade associada à estimativa (3.39).

O erro-padrão estimado a partir da amostra, $\hat{\tau}^2/N$, é obtido substituindo-se, em (3.40), γ_0 pela variância amostral de $t(\theta_n)$ e a autocorrelação teórica de ordem k , ρ_k , por sua estimativa:

$$r_k = \frac{\sum_{n=1}^{N-k} (t(\theta_n) - \bar{t}_N)(t(\theta_{n+k}) - \bar{t}_N)}{\sum_{n=1}^N (t(\theta_n) - \bar{t}_N)^2}.$$

Para a obtenção de uma amostra de valores preditivos, $(y_{T+h}|D_T)$, cujos valores teóricos seriam dados por (2.6), procede-se da seguinte forma: os parâmetros de estado α_t e β_t são atualizados aplicando-se as equações de evolução em (3.1) ou (3.15), respectivamente, com erros futuros u_{T+1}, \dots, u_{T+h} e v_{T+1}, \dots, v_{T+h} gerados de forma independente, de distribuições $N(0, W)$ e $N(0, Q)$, para cada ponto amostral de W e Q . Assim, amostras de α_{T+h} e β_{T+h} são obtidas aplicando-se as equações de evolução a cada ponto das amostras de $(\alpha_T|D_T)$ e $(\beta_T|D_T)$:

$$\alpha_{T+h} = \alpha_T + \sum_{i=1}^h u_{T+i};$$

$$\beta_{T+h} = \beta_T + \sum_{i=1}^h v_{T+i}.$$

Uma vez atualizados os parâmetros de estado, obtém-se uma amostra do preditor η_{T+h} e, finalmente, valores para y_{T+h} são gerados facilmente da distribuição \mathcal{F} , especificada em (3.1).

No capítulo a seguir, a metodologia aqui descrita é avaliada a partir de exercícios de ajuste dos modelos propostos a dados simulados.

Análise de Dados Artificiais

No presente capítulo, são apresentados os resultados obtidos no ajuste dos modelos propostos a dados simulados, com respostas seguindo as distribuições Poisson (seção 4.1), Binomial (seção 4.2), Gama (seção 4.3) e Gaussiana Inversa (seção 4.4). Busca-se, então, avaliar a capacidade dos mesmos para identificar os valores dos parâmetros utilizados na geração dos dados artificiais (aos quais passamos a nos referir como valores teóricos), bem como o desempenho do procedimento MCMC. Na seção 4.5 tecem-se algumas conclusões sobre os resultados obtidos com os conjuntos de dados simulados.

4.1 Resposta Poisson

A estrutura geral do modelo considerado nesta seção é dada por:

$$\begin{aligned} y_t &\sim \text{Poisson}(\lambda_t) \\ \log(\lambda_t) &= \eta_t, \end{aligned} \tag{4.1}$$

em que o preditor η_t assume diferentes formas, determinadas pelas variadas funções de transferência postuladas e pela natureza dos parâmetros que o compõem: estática ou dinâmica.

4.1.1 Função de transferência de ordem ($r = 1, s = 0$)

O modelo estático de primeira ordem apresenta o seguinte preditor:

$$\begin{aligned} \eta_t &= \alpha + E_t + \delta_1 Z_{1t} + \delta_2 Z_{2t} \\ E_t &= \rho E_{t-1} + \beta X_t, \end{aligned} \tag{4.2}$$

em que α é um nível estático, Z_1 e Z_2 são covariáveis e X é a variável cujo impacto cumulativo, a cada instante t (medido por E_t), se deseja avaliar. ρ é um parâmetro autoregressivo, que mede a "memória" do efeito de X , enquanto β mede o efeito imediato de X sobre a resposta esperada.

Com base na estrutura acima, foram gerados dois conjuntos de dados artificiais, ambos compostos por 1000 pontos, sendo 970 destes utilizados para estimação do modelo e os 30

restantes, para avaliação da capacidade preditiva. Nesta seção e ao longo de todo o presente capítulo, os valores adotados para as regressoras X e Z na simulação dos dados artificiais são valores reais de variáveis sobre as quais se tem interesse nas aplicações desenvolvidas no capítulo 5. Portanto, não há um mecanismo de simulação para X e Z , mas sim um mecanismo de geração de Y condicionalmente a X e Z . A figura 4.38 exibe as trajetórias da regressora X_t e das covariáveis Z_{1t} e Z_{2t} utilizadas ao longo das seções 4.1 e 4.2.

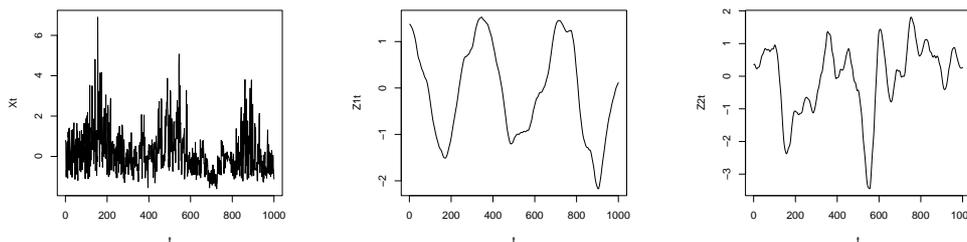


Figura 4.1 Regressora X_t e covariáveis Z_{1t} e Z_{2t} utilizadas nas simulações para respostas Poisson e Binomial.

O primeiro conjunto foi simulado com os seguintes valores para o vetor paramétrico: $\alpha = 1, 0$, $\rho = 0, 9$, $\beta = 0, 03$, $E_0 = 0, 5$, $\delta_1 = -0, 1$, $\delta_2 = 0, 13$. A estimação foi feita usando o algoritmo MCMC descrito no capítulo 3, com geração de valores a partir de uma proposta passeio aleatório para $\log(\rho)$ e proposta baseada na verossimilhança para os demais parâmetros. A amostragem de δ_1 e δ_2 foi feita em conjunto e os demais parâmetros foram amostrados individualmente. As distribuições a priori para todos os parâmetros foram vagas: $\alpha, \beta, E_0, \delta_1, \delta_2 \stackrel{iid}{\sim} N(0, 10^5)$, $\rho \sim U(0, 1)$. Observou-se convergência bastante rápida do algoritmo MCMC (as cadeias geradas já apresentam comportamento convergente com algo em torno de 2000 iterações). Ao todo, o procedimento contou com 30000 iterações. Optou-se por gerar cadeias longas, de forma a se ter mais segurança de sua convergência. A figura 4.2 exibe as trajetórias de duas cadeias paralelas para cada um dos parâmetros, com saltos de 20 iterações) e observa-se que as duas mesclam-se bem. Na mesma figura também estão dispostos os histogramas das amostras das distribuições a posteriori. Os histogramas foram construídos tomando-se as 10000 iterações finais de cada cadeia paralela, a intervalos de 20 observações, de forma a reduzir a autocorrelação serial, sendo a amostra final, portanto, composta por 1000 observações. O procedimento de estimação foi capaz de recuperar bastante bem os valores teóricos dos parâmetros.

A tabela 4.1 exibe os valores das componentes do vetor paramétrico utilizados para simular os dados e estatísticas sumárias das amostras da posteriori - valores mínimo e máximo, quartis, média e seu erro-padrão estimado, $\hat{\tau}^2/N$ - obtidas como descrito na seção 3.5. Exceto pelo parâmetro E_0 , que mede o efeito acumulado da regressora no instante inicial e sobre o qual existe grande incerteza, de forma geral tanto a média amostral quanto a mediana apresentam-se como estimativas pontuais acuradas, sendo a média um estimador bastante preciso, como se pode ver pelos erros-padrão calculados. A tabela 4.2 apresenta as correlações a posteriori entre

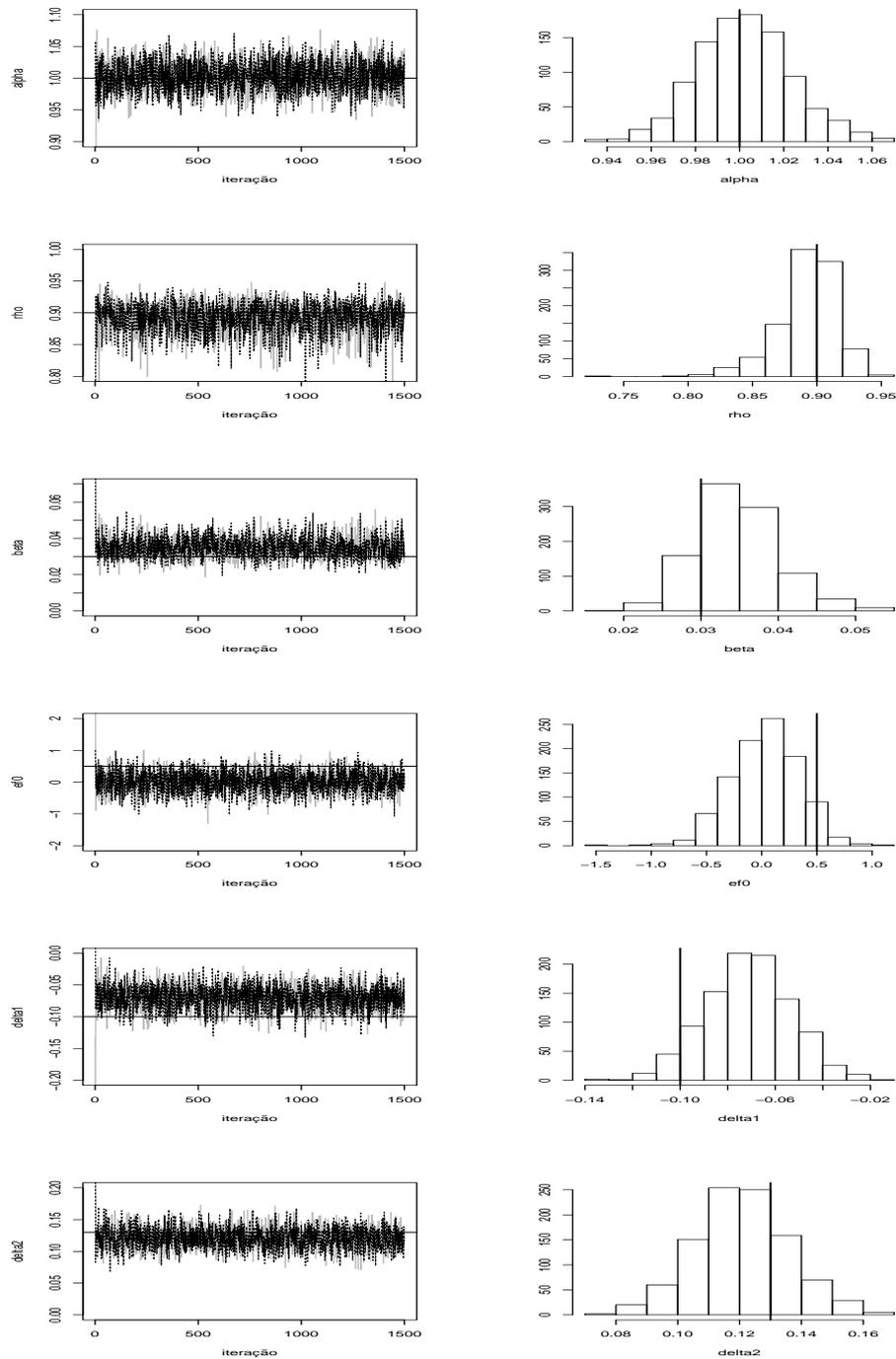


Figura 4.2 Resposta Poisson: Traços das cadeias geradas e histogramas da amostra da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).

os parâmetros, destacando-se a correlação entre β e ρ .

Tabela 4.1 Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem ($r = 1, s = 0$), para dados simulados com $\rho = 0,9$.

	α	ρ	β	E_0	δ_1	δ_2
Teórico	1,0	0,9	0,03	0,5	-0,1	0,13
Mín	0,93360	0,73170	0,01801	-1,47700	-0,13280	0,07496
Q1	0,98740	0,88090	0,03111	-0,17490	-0,08382	0,11110
Mediana	1,00200	0,89530	0,03450	0,03747	-0,07128	0,12050
Média	1,00200	0,89250	0,03473	0,03085	-0,07143	0,12080
Q3	1,01500	0,90840	0,03806	0,25020	-0,05963	0,13040
Máx.	1,06800	0,94820	0,05494	1,05800	-0,01336	0,16380
e.p.	0,00031	0,00051	0,00009	0,00596	0,00014	0,00010

Tabela 4.2 Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r = 1, s = 0$), para dados simulados com $\rho = 0,9$.

	α	ρ	β	E_0	δ_1	δ_2
α	1,00	-0,26	-0,11	-0,22	-0,08	-0,37
ρ	-0,26	1,00	-0,75	0,11	0,35	0,38
β	-0,11	-0,75	1,00	-0,02	-0,02	-0,03
E_0	-0,22	0,11	-0,02	1,00	-0,10	0,08
δ_1	-0,08	0,35	-0,02	-0,10	1,00	0,12
δ_2	-0,37	0,38	-0,03	0,08	0,12	1,00

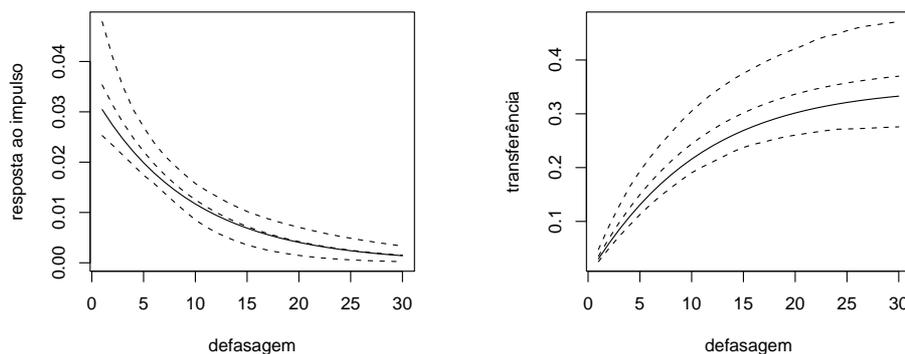


Figura 4.3 Resposta Poisson, modelo de ordem ($r=1, s=0$): Função de resposta ao impulso e de transferência - teóricas (linha sólida), médias e limites de credibilidade estimados a 95% (linhas tracejadas).

A figura 4.3 exibe as funções de resposta ao impulso e de transferência teóricas (linhas sólidas) e suas estimativas a posteriori, junto a intervalos de credibilidade a 95%. Mais especificamente, apresentam-se riscos relativos associados à exposição a uma unidade da regressora, em comparação ao nível nulo, uma vez que tais funções já se encontram exponenciadas e descontadas de $e^0 = 1$. As estimativas pontuais (linhas tracejadas) são dadas pela média de cada uma das funções, ao longo das 1000 iterações que compõem a amostra final do algoritmo MCMC. Embora as funções teóricas sejam ligeiramente superestimadas pelas estimativas pontuais, seus valores são capturados pelos limites de credibilidade a 95%. Na figura 4.4 pode-se comparar a função de resposta média (λ_t) teórica à sua estimativa (dada pela média a posteriori). Como se vê no diagrama de dispersão, em alguns poucos pontos a média a posteriori subestima a função de resposta média; entretanto, ainda assim, todos os valores da função de resposta média teórica são capturados pelo intervalo de credibilidade a 95%.

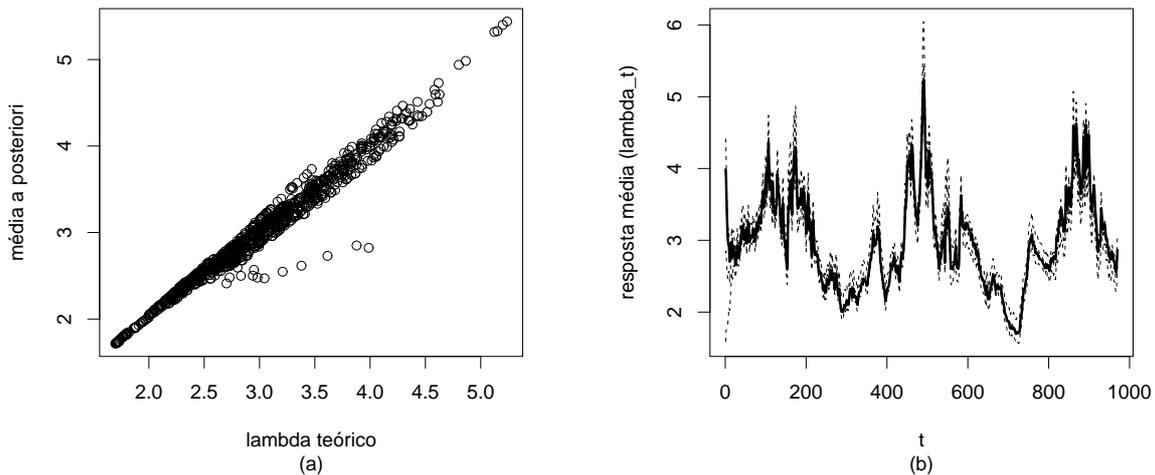


Figura 4.4 Resposta Poisson, modelo de ordem ($r=1, s=0$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

Os erros quadráticos médios associados à média, à mediana e à moda das amostras das distribuições preditivas, para horizontes variando de 1 a 30, foram, respectivamente: 1,69, 1,57 e 1,83, indicando uma ligeira vantagem da mediana em relação às demais estimativas pontuais. A forma das distribuições preditivas pode ser vista na figura 4.5, que exibe histogramas das amostras das mesmas, nos quais a linha vertical corresponde ao valor real observado para a resposta a cada instante, capturado pela distribuição preditiva (ao levarmos em conta intervalos de credibilidade a 95%) em todos os 30 instantes no horizonte de previsão.

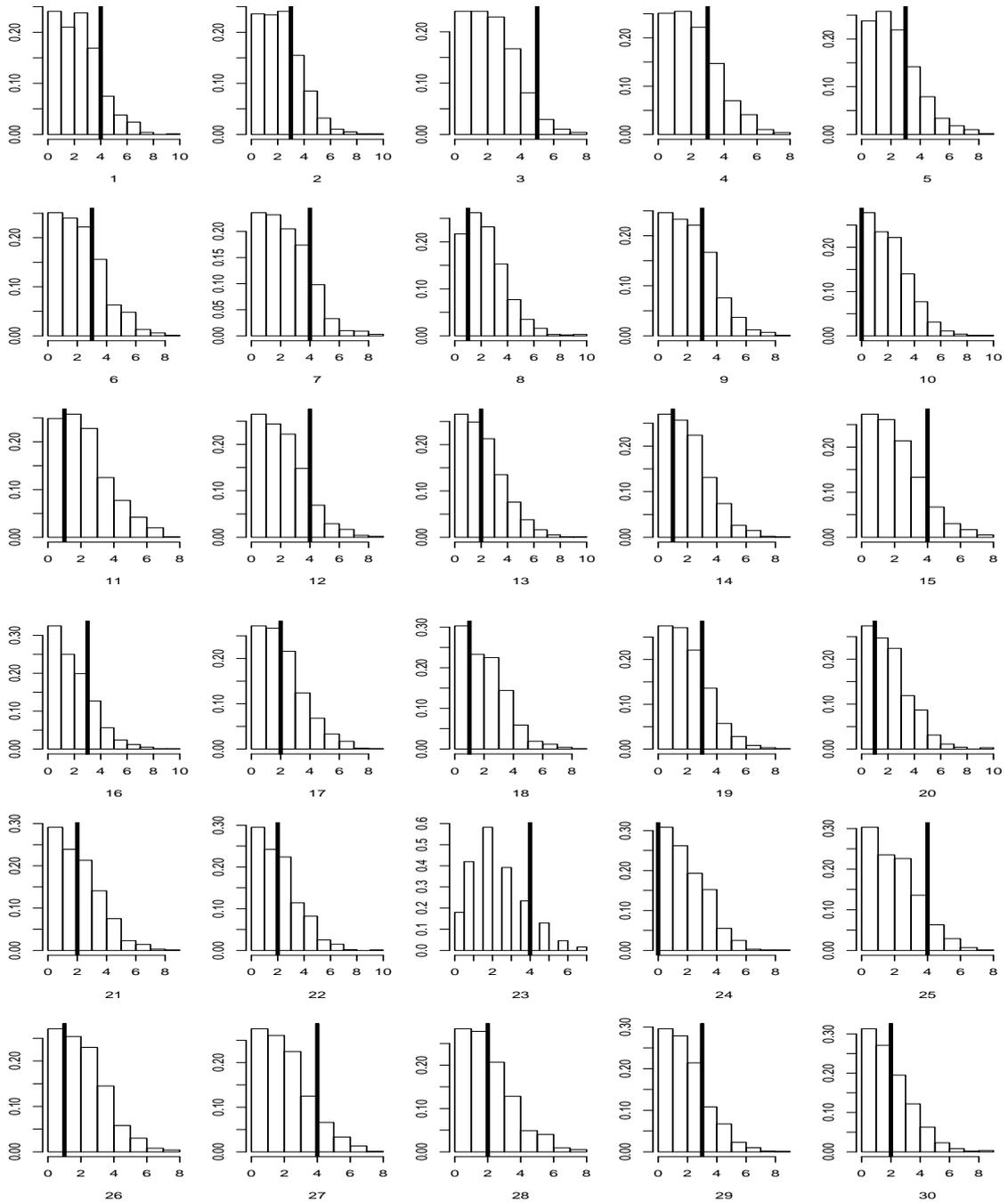


Figura 4.5 Resposta Poisson, modelo de ordem $(r=1,s=0)$: Histogramas das amostras das distribuições preditivas para horizontes variando de 1 a 30.

Uma questão relevante avaliada neste exercício foi a quantidade de informação contida nos dados necessária à estimação dos parâmetros do modelo. O modelo autoregressivo foi ajustado com o tamanho da série assumindo valores: 100, 300, 500, 700 e 1000 (os resultados deste último foram apresentados acima). A figura 4.6 exhibe, para os diferentes tamanhos da série, os histogramas da amostra a posteriori para o parâmetro autoregressivo (seu valor teórico é 0,90). Percebe-se que o verdadeiro valor do parâmetro passa a ser bem estimado com algo em torno de 700 observações.

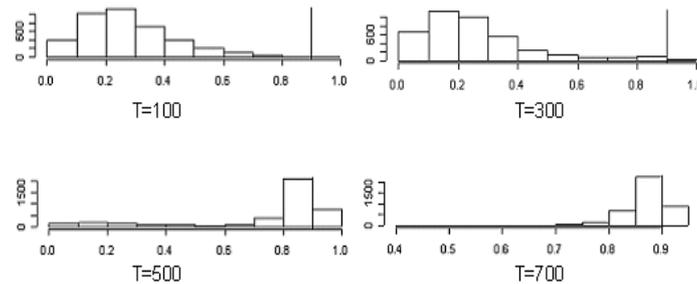


Figura 4.6 Resposta Poisson: Histogramas das amostras das posteriores de ρ , para dados simulados segundo o modelo de ordem $(r=1, s=0)$, ajustado para $T=100, 300, 500$ e 700 .

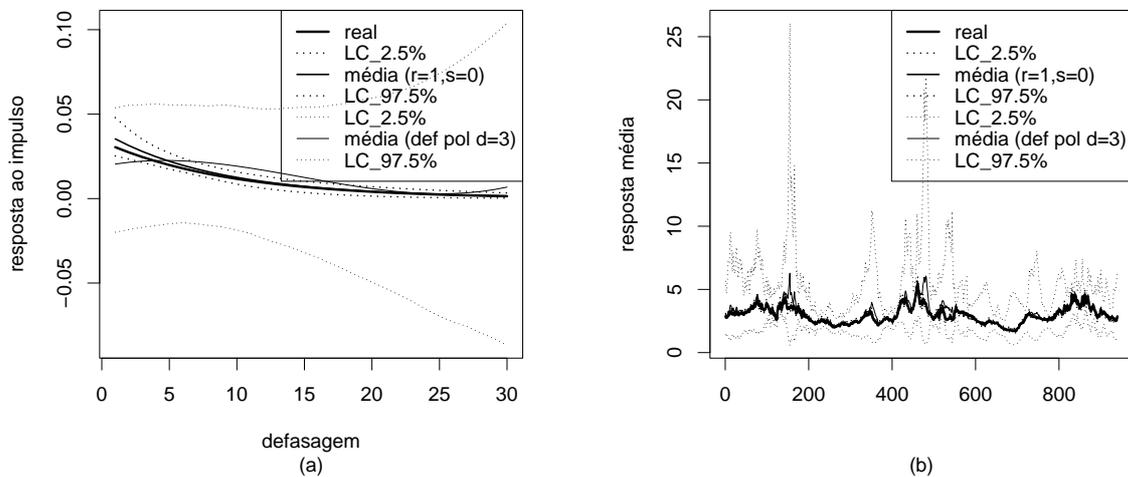


Figura 4.7 Resposta Poisson, modelo de ordem $(r=1, s=0)$ com $\rho = 0,9$ e $\beta = 0,03$: (a) Função de resposta ao impulso; (b) Função de resposta média. Funções teóricas (linha em negrito), estimativas das médias a posteriori e de limites de credibilidade obtidos do ajuste dos modelos de ordem $(r=1, s=0)$ e polinomial de grau 3.

Ainda, modelos polinomiais de graus 2 e 3 foram ajustados, considerando-se o mesmo conjunto de dados, simulado de acordo com o modelo de ordem $(r = 1, s = 0)$ com parâmetro

autoregressivo $\rho = 0,9$. Para o modelo de grau 2, a convergência das cadeias geradas não foi atingida. Já para o modelo de grau 3, embora o procedimento MCMC tenha convergido, a função de resposta ao impulso estimada - como se pode ver na figura 4.7 - tem grande incerteza associada, apresentando-se não significativa, apesar de seus valores médios a posteriori estarem razoavelmente próximos aos efeitos teóricos. A função de resposta média estimada com o modelo polinomial apresenta comportamento semelhante, com valores médios acompanhando os teóricos, mas sujeita a grande incerteza. Entretanto, como os dados foram gerados a partir do modelo autoregressivo, é natural que o ajuste segundo tal estrutura exiba menor incerteza.

Um segundo conjunto de dados foi simulado, repetindo-se o exercício anterior, de forma a avaliar o comportamento do procedimento de estimação para valores do parâmetro autoregressivo próximos à fronteira de estabilidade. Para isso, fez-se $\rho = 0,99$. Embora a convergência tenha sido um pouco mais lenta (após, aproximadamente, 5000 iterações), não houve problemas para estimação do parâmetro autoregressivo. As tabelas 4.3 e 4.4 exibem, respectivamente, estatísticas sumárias e correlações das amostras da posteriori e, com exceção dos parâmetros E_0 e δ_1 , média e mediana a posteriori apresentam-se como estimativas bastante acuradas.

Tabela 4.3 Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem ($r = 1, s = 0$), para dados simulados com $\rho = 0,99$.

	α	ρ	β	E_0	δ_1	δ_2
Teórico	1,0	0,99	0,03	0,5	-0,10	0,13
Mín.	0,89510	0,98700	0,02691	0,06480	-0,15060	0,10410
Q1	1,02400	0,98900	0,03013	0,33290	-0,07437	0,13140
Mediana	1,04700	0,98940	0,03107	0,39900	-0,05624	0,13810
Média	1,04700	0,98940	0,03107	0,39480	-0,05671	0,13820
Q3	1,07200	0,98990	0,03197	0,45610	-0,03783	0,14500
Máx.	1,15000	0,99170	0,03634	0,70950	0,04039	0,16790
e.p.	0,00175	0,00005	0,00007	0,00624	0,00177	0,00033

Tabela 4.4 Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r = 1, s = 0$), para dados simulados com $\rho = 0,99$.

	α	ρ	β	E_0	δ_1	δ_2
α	1,00	-0,73	0,43	-0,68	0,66	0,00
ρ	-0,73	1,00	-0,91	0,65	-0,91	-0,41
β	0,43	-0,91	1,00	-0,54	0,88	0,63
E_0	-0,68	0,65	-0,54	1,00	-0,63	-0,21
δ_1	0,66	-0,91	0,88	-0,63	1,00	0,37
δ_2	0,00	-0,41	0,63	-0,21	0,37	1,00

4.1.2 Função de transferência de ordem ($r = 1, s > 0$)

Suponha-se, agora, a seguinte forma para o preditor η_t em (4.1):

$$\begin{aligned}\eta_t &= \alpha + E_t + \delta_1 Z_{1t} + \delta_2 Z_{2t} \\ E_t &= \rho E_{t-1} + \beta_0 X_t + \beta_1 X_{t-1} + \cdots + \beta_s X_{t-s}.\end{aligned}\tag{4.3}$$

Inicialmente, simulou-se uma série com 1000 observações, de acordo com a estrutura acima, assumindo-se $s = 1$. Adotaram-se os seguintes valores para o vetor de parâmetros: $\alpha = 1,0$, $\rho = 0,9$, $\beta_0 = 0,03$, $\beta_1 = 0,07$, $E_0 = 0,5$, $\delta_1 = -0,1$, $\delta_2 = 0,13$. As distribuições a priori de todos os parâmetros foram não-informativas e a geração de amostras da posteriori foi feita usando-se o algoritmo de Metropolis-Hastings. Novamente, as densidades propostas utilizadas foram: passeio aleatório para $\log(\rho)$ e a proposta baseada em aproximação em torno da média para os demais parâmetros. μ , ρ e E_0 foram amostrados individualmente; β_0 e β_1 , bem como δ_1 e δ_2 foram gerados em conjunto.

A figura 4.8 exibe os traços, a cada 20 iterações (de um total de 30000), das cadeias geradas e histogramas das amostras após suposta convergência. Na tabela 4.5, encontram-se estatísticas sumárias da amostra da posteriori.

O parâmetro β_0 , que mede o efeito imediato da regressora, é ligeiramente subestimado, com intervalo de credibilidade a 95% estimado por (0,0165, 0,0297). A taxa de aceitação para o bloco (β_0, β_1) é bastante baixa (em torno de 10%), sugerindo a possível necessidade de uma cadeia longa para se atingir convergência. O procedimento foi repetido, então, com 100000 iterações, com resultados bastante semelhantes àqueles obtidos com 30000 iterações e, portanto, levando à conclusão - já indicada pelos traços da figura 4.8 - de convergência da cadeia. Na tabela 4.6 são apresentadas as estimativas das correlações a posteriori. Comparando-se os resultados àqueles obtidos para o modelo de ordem ($r = 1, s = 0$) observa-se que, naquele, o parâmetro β (de efeito imediato da regressora) apresentava-se bastante correlacionado ao parâmetro autoregressivo. Ao se adotar uma defasagem, esta correlação diminui consideravelmente. Em contrapartida, a correlação entre ρ e β_1 é bastante elevada, uma vez que estes dois parâmetros retratam efeitos passados da regressora sobre a resposta média.

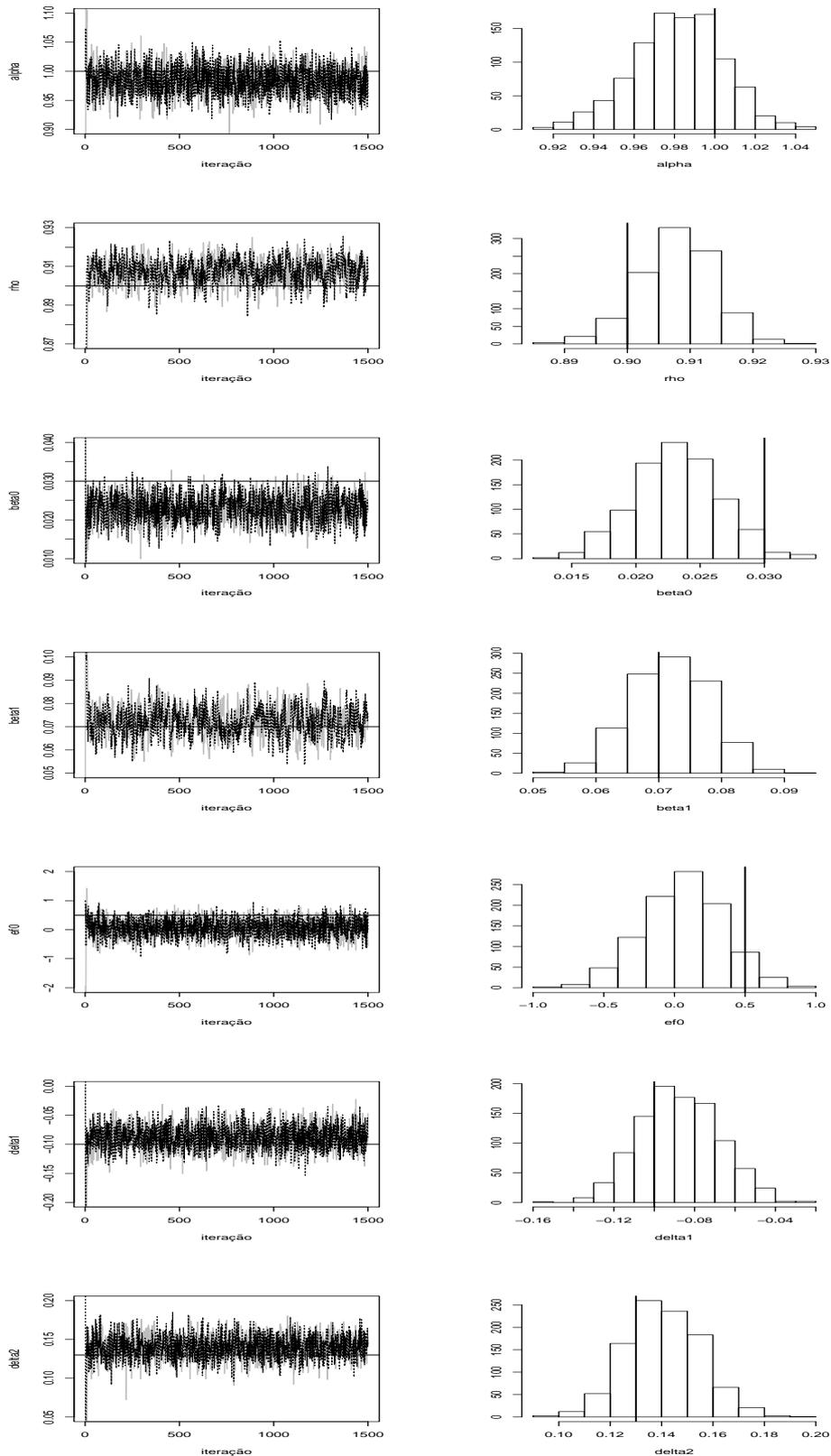


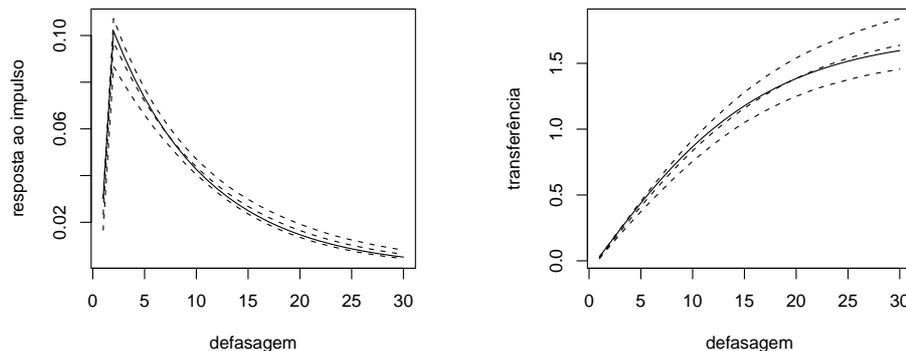
Figura 4.8 Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de função de transferência de ordem ($r = 1, s = 1$).

Tabela 4.5 Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $r=1, s=1$.

	α	ρ	β_0	β_1	E_0	δ_1	δ_2
Teórico	1,0	0,9	0,03	0,07	0,5	-0,10	0,13
Mín.	0,91810	0,88870	0,01267	0,05453	-0,99740	-0,15330	0,09333
Q1	0,96650	0,90410	0,02097	0,06765	-0,12240	-0,10130	0,13100
Mediana	0,98250	0,90840	0,02328	0,07174	0,06810	-0,08840	0,14040
Média	0,98180	0,90790	0,02317	0,07177	0,06255	-0,08732	0,14090
Q3	0,99650	0,91200	0,02538	0,07609	0,25080	-0,07454	0,15110
Máx.	1,04900	0,92550	0,03386	0,09014	0,90230	-0,02711	0,19260
e.p.	0,00038	0,00020	0,00004	0,00009	0,00999	0,00035	0,00024

Tabela 4.6 Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $(r=1, s=1)$.

	α	ρ	β_0	β_1	E_0	δ_1	δ_2
α	1,00	-0,28	-0,26	0,01	-0,07	-0,02	-0,39
ρ	-0,28	1,00	0,38	-0,79	0,01	0,25	0,46
β_0	-0,26	0,38	1,00	-0,61	0,03	0,27	0,40
β_1	0,01	-0,79	-0,61	1,00	-0,01	-0,01	-0,19
E_0	-0,07	0,01	0,03	-0,01	1,00	-0,05	0,01
δ_1	-0,02	0,25	0,27	-0,01	-0,05	1,00	0,27
δ_2	-0,39	0,46	0,40	-0,19	0,01	0,27	1,00

**Figura 4.9** Resposta Poisson, modelo de ordem $(r=1, s=1)$: Função de resposta ao impulso e de transferência - teóricas (linha sólida), médias e limites de credibilidade estimados a 95% (linhas tracejadas).

A figura 4.9 exibe as funções de resposta ao impulso e de transferência teóricas (linhas sólidas), suas médias a posteriori e limites de credibilidade estimados. Apesar da ligeira subestimativa de β_0 , o processo de estimação exibe excelentes resultados para tais funções.

Na figura 4.10 pode-se comparar a função de resposta média (λ_t) teórica à sua estimativa (média a posteriori). Os limites estimados do intervalo de credibilidade a 95% compreendem todos os valores da função de resposta média teórica.

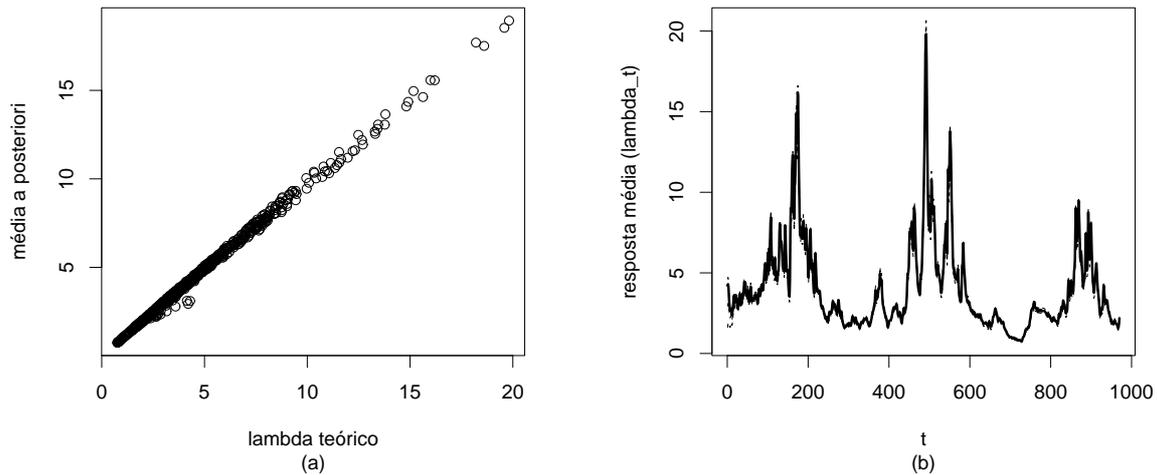


Figura 4.10 Resposta Poisson, modelo de ordem ($r=1, s=1$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

A capacidade preditiva do modelo foi avaliada para horizontes variando de 1 a 30. A média amostral das distribuições preditivas apresentou erro quadrático médio ligeiramente menor que mediana e moda: 1,76, contra 1,93 (erro associado à mediana) e 1,97 (à moda). Histogramas das amostras das distribuições preditivas preditivas são dispostos na figura 4.11, e nota-se que as distribuições preditivas capturam bastante bem os valores reais observados.

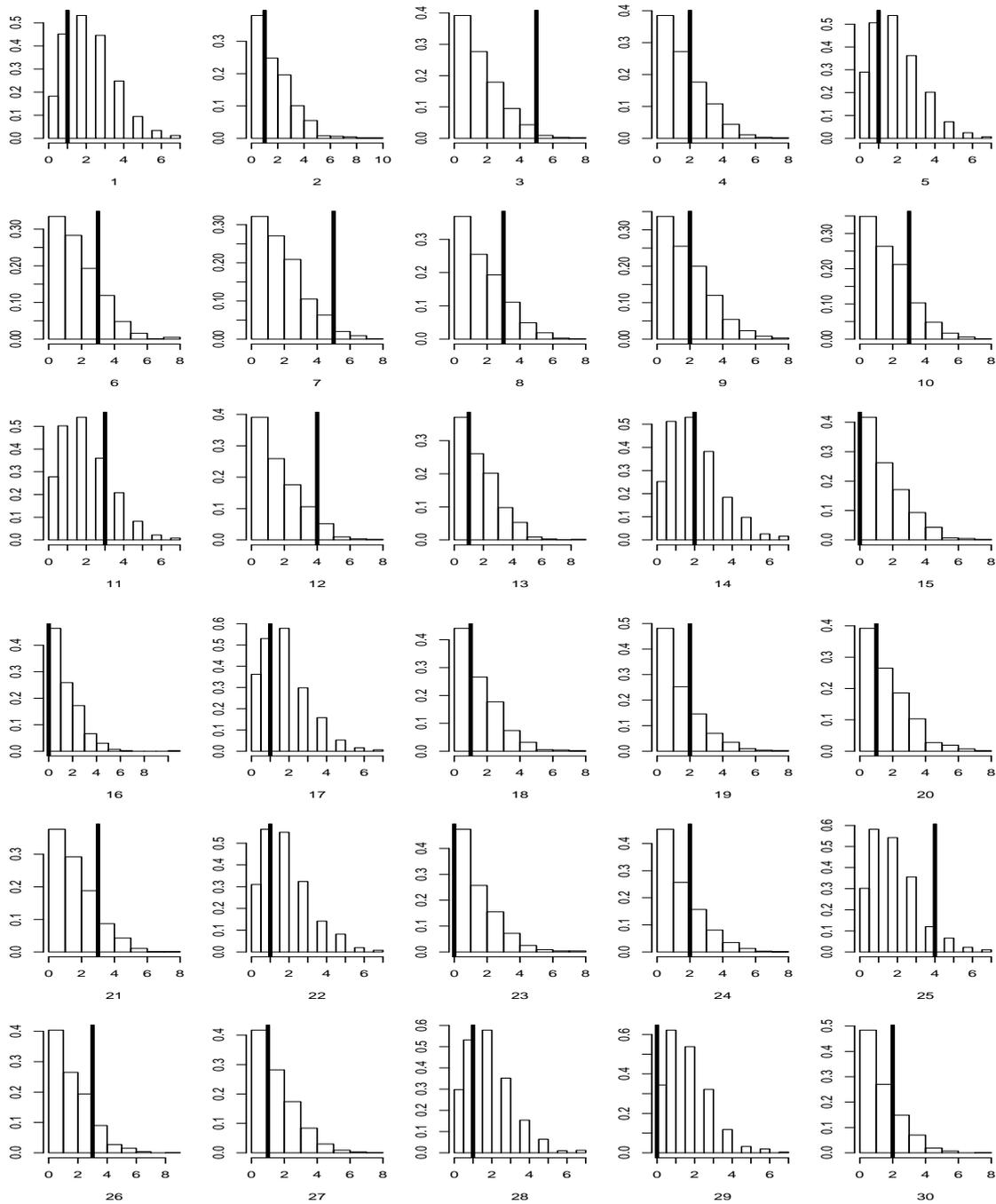


Figura 4.11 Resposta Poisson, modelo de ordem ($r=1,s=1$): Amostra das distribuições preditivas para horizontes variando de 1 a 30.

As dificuldades na estimação aumentam à medida em que a ordem das defasagens na regressora, s , cresce. A figura 4.12 e as tabelas 4.7 e 4.8 exibem o sumário das amostras da posteriori para o modelo de ordem ($r = 1, s = 2$), tendo como resposta um conjunto de dados simulados a partir dos seguintes valores para o vetor de parâmetros: $\alpha = 1,0, \rho = 0,9, \beta_0 = 0,1, \beta_1 = 0,06, \beta_2 = 0,03, E_0 = 0,5, \delta_1 = -0,1, \delta_2 = 0,13$.

Como se pode ver na tabela 4.7, os parâmetros β_1 e β_2 são subestimados. Para avaliar se a verossimilhança contém informação suficiente para a estimação do parâmetro β_2 , que foi particularmente mal estimado, construiu-se uma grade cruzada para os parâmetros β_2 e ρ (observe-se, na tabela 4.8, a elevada correlação a posteriori entre estes), condicionando-se aos verdadeiros valores dos demais parâmetros. A figura 4.13 exhibe as verossimilhanças perfiladas para β_2 e ρ . Pode-se observar que a verossimilhança concentra-se abaixo do valor teórico de β_2 .

A estimação espúria dos parâmetros de impacto defasado da regressora obviamente é refletida na estimação das funções de reposta ao impulso, subestimada, como se pode ver na figura

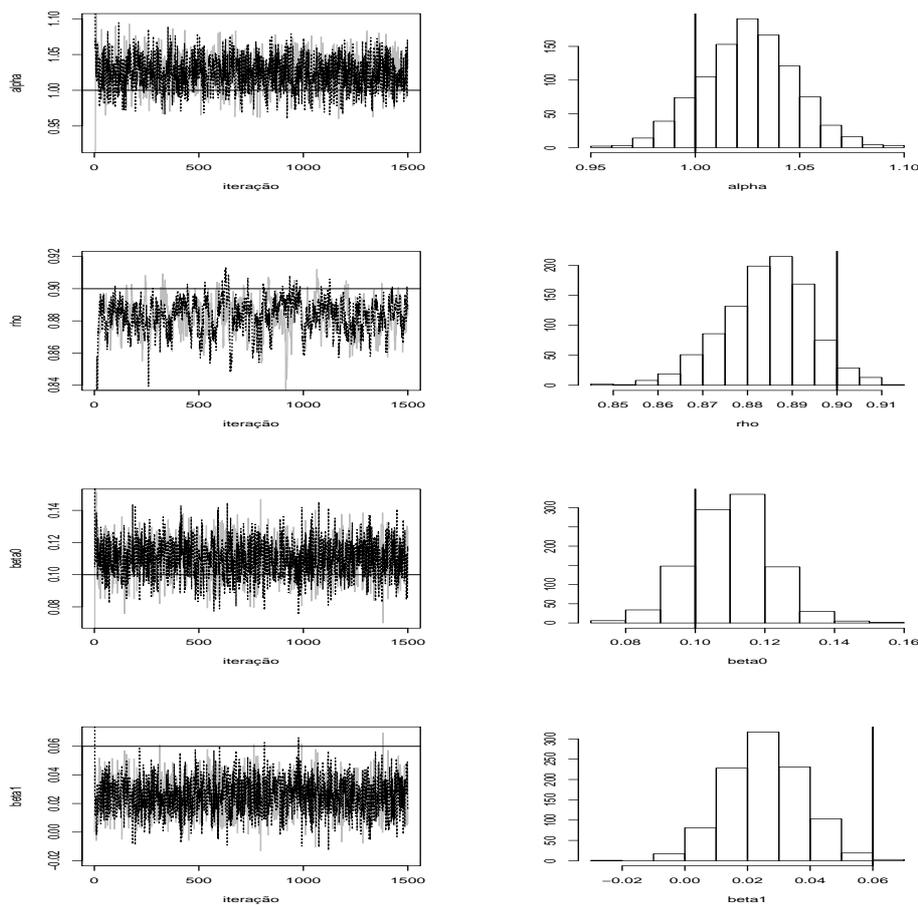


Figura 4.12 Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de ordem ($r=1, s=2$).

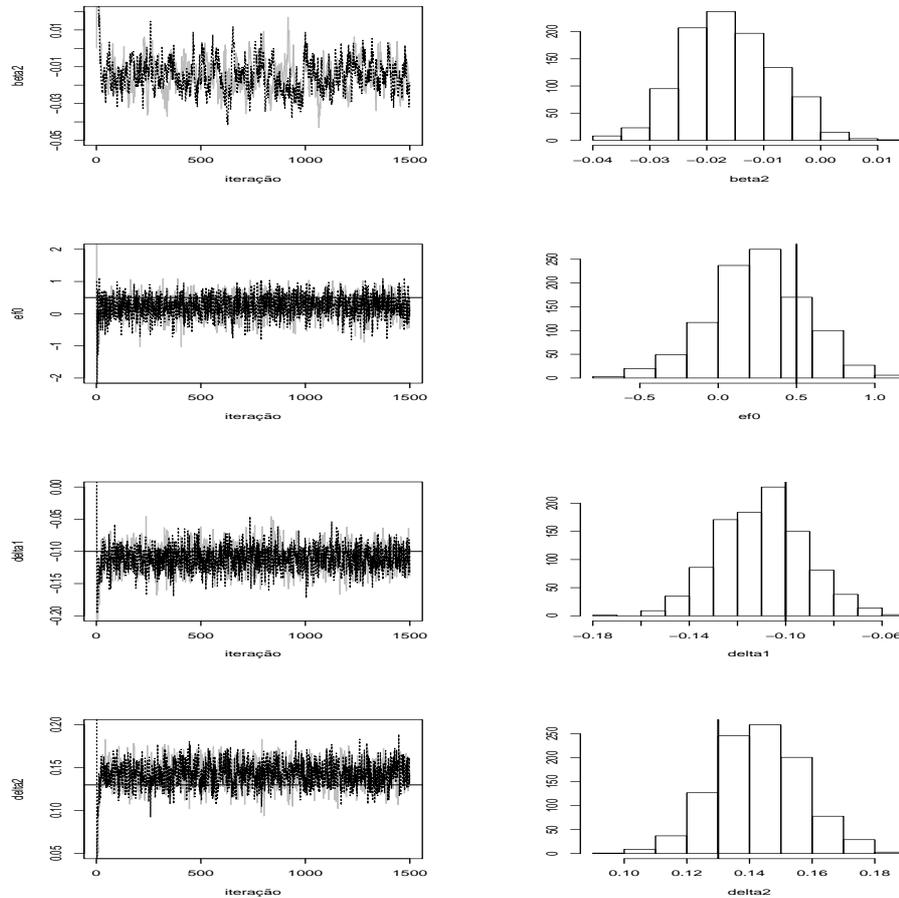


Figura 4.12 Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de ordem $(r=1, s=2)$ (continuação).

Tabela 4.7 Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $(r=1, s=2)$.

	α	ρ	β_0	β_1	β_2	E_0	δ_1	δ_2
Teórico	1,0	0,9	0,1	0,06	0,03	0,5	-0,10	0,13
Mín.	0,9550	0,8493	0,0765	-0,0221	-0,0380	-0,7801	-0,1763	0,0990
Q1	1,0110	0,8785	0,1022	0,0174	-0,0214	0,0572	-0,1226	0,1338
Mediana	1,0270	0,8851	0,1107	0,0250	-0,0165	0,2639	-0,1093	0,1426
Média	1,0250	0,8844	0,1101	0,0256	-0,0161	0,2553	-0,1099	0,1431
Q3	1,0400	0,8909	0,1175	0,0338	-0,0106	0,4466	-0,0982	0,1527
Máx.	1,0990	0,9112	0,1561	0,0626	0,0101	1,0670	-0,0575	0,1875
e.p.	0,0005	0,0005	0,0002	0,0001	0,0001	0,0150	0,0002	0,0005

4.14. Entretanto, a função de resposta média, exibida na mesma figura, é bastante bem estimada, configurando-se um problema de identificação no modelo, acarretado, provavelmente,

Tabela 4.8 Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r=1, s=2$).

	α	ρ	β_0	β_1	β_2	E_0	δ_1	δ_2
α	1,00	-0,37	-0,12	0,02	0,20	-0,14	-0,07	-0,42
ρ	-0,37	1,00	-0,05	0,11	-0,88	0,06	0,30	0,53
β_0	-0,12	-0,05	1,00	-0,96	0,20	0,00	0,09	0,10
β_1	0,02	0,11	-0,96	1,00	-0,30	-0,01	0,00	0,01
β_2	0,20	-0,88	0,20	-0,30	1,00	-0,05	-0,12	-0,34
E_0	-0,14	0,06	0,00	-0,01	-0,05	1,00	-0,14	0,03
δ_1	-0,07	0,30	0,09	0,00	-0,12	-0,14	1,00	0,27
δ_2	-0,42	0,53	0,10	0,01	-0,34	0,03	0,27	1,00

pela forte correlação entre X_t , X_{t-1} e X_{t-2} , associada, ainda, à adoção de um parâmetro autoregressivo, que "compete" com as variáveis defasadas pela representação da memória do processo. Como se pode ver na mesma figura, o modelo de defasagem polinomial de grau 3, aplicado ao mesmo conjunto de dados, apresenta função de transferência com magnitude bastante semelhante ao modelo autoregressivo, embora não seja capaz de captar o efeito crescente nas primeiras defasagens. Naturalmente, o modelo polinomial apresenta maior incerteza, uma vez que os dados foram gerados da estrutura autoregressiva de ordem $r = 1, s = 2$ e também não foi capaz de recuperar a função de resposta ao impulso teórica.

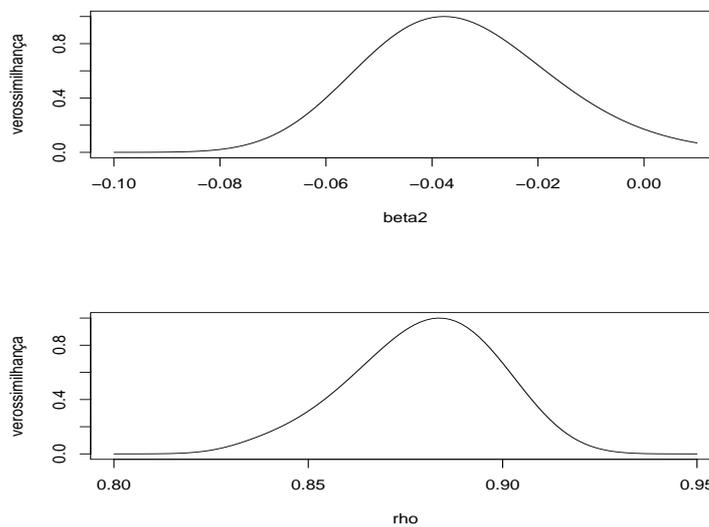


Figura 4.13 Resposta Poisson: verossimilhança para ρ e β_2 , para dados simulados de acordo com o modelo de ordem ($r=1, s=2$).

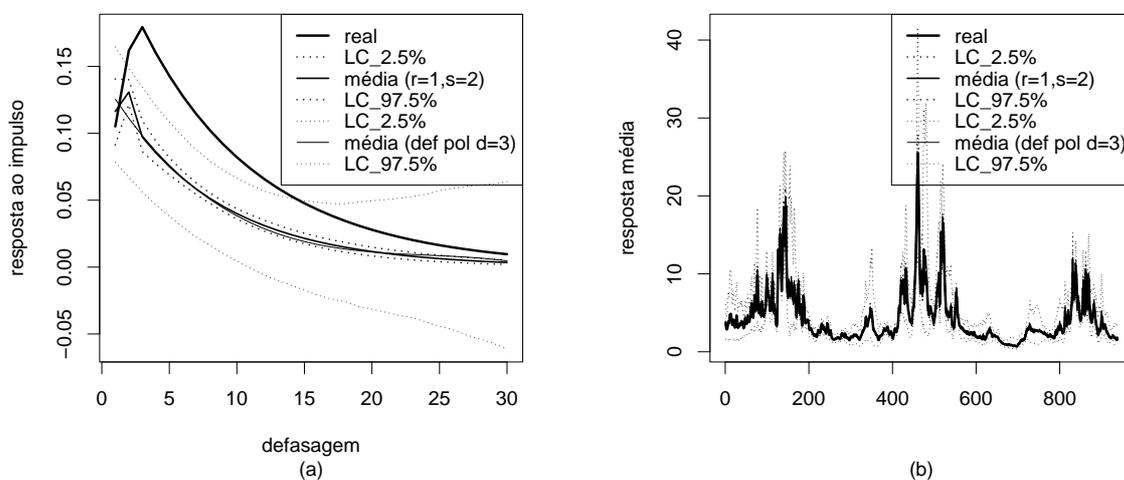


Figura 4.14 Resposta Poisson, modelo de ordem ($r=1, s=2$) com $\rho = 0,9$, $\beta_0 = 1,0$, $\beta_1 = 0,06$ e $\beta_2 = 0,03$: (a) Função de resposta ao impulso; (b) Função de resposta média. Funções teóricas (linha em negrito), estimativas das médias a posteriori e de limites de credibilidade obtidos do ajuste dos modelos de ordem ($r=1, s=2$) e polinomial de grau 3.

4.1.3 Função de transferência de ordem ($r = 1, s = 0$) com erro aleatório

Assume-se, agora, a estrutura (3.16) para o efeito acumulado de X_t , em que um termo aleatório é acrescido a E_t . Um conjunto de 1000 observações foi simulado assumindo-se os seguintes valores para as componentes do vetor paramétrico: $\alpha = 1,0$, $\rho = 0,9$, $\beta = 0,03$, $E_0 = 0,5$, $Q_\varepsilon = 0,07$, $\delta_1 = -0,10$, $\delta_2 = 0,13$.

A figura 4.15 exibe os traços das cadeias geradas no procedimento de estimação, com saltos de 20 iterações, de forma a reduzir a autocorrelação serial, junto a histogramas das amostras da posteriori.

Como se pode ver na tabela 4.9, o parâmetro autoregressivo é subestimado pela média e pela mediana da amostra da posteriori obtida via MCMC; entretanto o intervalo de credibilidade estimado a 95%, dado por $(0,703, 0,915)$ captura seu verdadeiro valor. O mesmo ocorre com a variância dos erros, cujo intervalo de credibilidade a 95% é estimado por $(0,019, 0,085)$. A tabela 4.10 reporta correlações a posteriori estimadas, destacando-se a correlação entre ρ e β , assim como no caso estático de ordem ($r = 1, s = 0$).

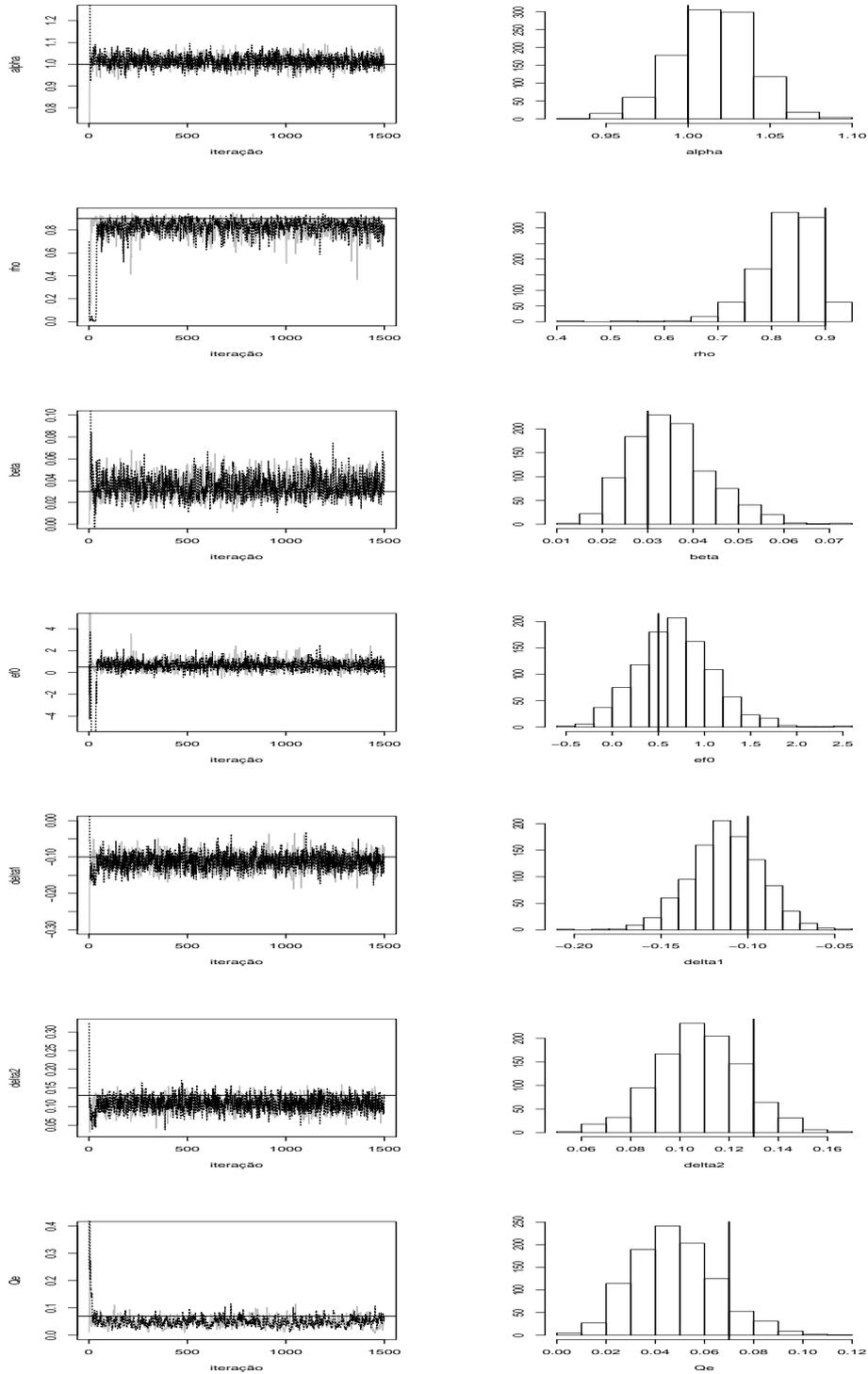


Figura 4.15 Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem $(r=1, s=0)$ com erro aleatório.

Tabela 4.9 Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem ($r=1, s=0$) com erro aleatório.

	α	ρ	β	E_0	δ_1	δ_2	Q_ε
Teórico	1,0	0,9	0,03	0,5	-0,10	0,13	0,07
Mín	0,93700	0,42070	0,0143	-0,44980	-0,20120	0,05471	0,00611
Q1	0,99980	0,79840	0,0287	0,41860	-0,12570	0,09708	0,03572
Mediana	1,01700	0,83730	0,0341	0,66760	-0,11290	0,10780	0,04696
Média	1,01500	0,82870	0,0351	0,69180	-0,11280	0,10810	0,04778
Q3	1,03100	0,86770	0,0402	0,94770	-0,09895	0,11980	0,05853
Máx.	1,09400	0,94550	0,0744	2,54300	-0,04853	0,16360	0,11230
e.p.	0,00065	0,00131	0,0001	0,00789	0,00033	0,00023	0,00099

Tabela 4.10 Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r=1, s=0$) com erro aleatório.

	α	ρ	β	E_0	δ_1	δ_2	Q_ε
α	1,00	-0,22	-0,12	-0,01	-0,10	-0,27	-0,31
ρ	-0,22	1,00	-0,64	-0,41	0,31	0,37	-0,13
β	-0,12	-0,64	1,00	0,23	0,13	0,01	0,13
E_0	-0,01	-0,41	0,23	1,00	-0,29	-0,14	0,05
δ_1	-0,10	0,31	0,13	-0,29	1,00	0,10	-0,05
δ_2	-0,27	0,37	0,01	-0,14	0,10	1,00	-0,02
Q_ε	-0,31	-0,13	0,13	0,05	-0,05	-0,02	1,00

4.1.4 Função de transferência de ordem ($r = 1, s = 0$) e nível dinâmico

Embora a evolução dinâmica do nível α_t - como postulada em (3.1) - não implique caráter dinâmico para as funções de transferência e de resposta ao impulso, pode-se conseguir, ao se permitir tal evolução, modelos mais adaptativos que os estáticos. Para avaliar a identificabilidade do modelo sob tais condições, simulou-se um conjunto de dados com função de transferência de ordem ($r = 1, s = 0$), gerada a partir de $\beta = 0,03$, $\rho = 0,9$. O efeito inicial da regressora X_t foi fixado em $E_0 = 0,1$. O preditor utilizado na simulação contou com duas covariáveis, cujos parâmetros são $\delta_1 = -0,10$ e $\delta_2 = 0,13$. O nível α_t evolui segundo um passeio aleatório com variância dos erros dada por $W = 0,005$.

As cadeias geradas pelo algoritmo MCMC apresentam comportamento convergente com menos de 5.000 iterações. Ainda assim, foram realizadas, ao todo, 30.000 iterações do algoritmo. A figura 4.16 exhibe os traços das cadeias paralelas construídas no procedimento de estimação, com saltos de 20 iterações, junto aos histogramas das amostras finais para cada parâmetro.

Na tabela 4.11, encontram-se estatísticas sumárias associadas às amostras da posteriori. Como se pode verificar, as distribuições a posteriori foram capazes de recuperar bastante bem os valores teóricos dos parâmetros estáticos, à exceção do parâmetro autoregressivo, um tanto subestimado, com intervalo de credibilidade a 95% estimado por $(0,0598, 0,8912)$. A inserção de um nível governado por passeio aleatório introduz no preditor mais um bloco estrutural para avaliação da memória do processo sob análise, além do bloco de função de transferência governado pelo parâmetro autoregressivo. Daí decorre, possivelmente, a dificuldade na identificação deste último. Média e mediana apresentam-se como estimativas pontuais acuradas para os parâmetros β e W e a média, em particular, apresenta-se bastante precisa (como se pode ver pelo erro-padrão reportado). Na tabela 4.12 encontram-se estimativas das correlações a posteriori. A correlação entre os parâmetros ρ e β , que nas aplicações com nível estático foram bastante significativas, agora é praticamente nula.

Tabela 4.11 Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem ($r=1, s=0$) com nível dinâmico.

	ρ	β	E_0	δ_1	δ_2	W
Teórico	0,9	0,03	0,1	-0,10	0,13	0,005
Mín	0,00337	-0,0197	-7,68400	-0,49260	0,03306	0,00222
Q1	0,46410	0,0292	-0,51800	-0,13200	0,14090	0,00448
Mediana	0,64770	0,0389	0,13240	-0,04564	0,18630	0,00544
Média	0,59220	0,0388	0,09489	-0,05023	0,19210	0,00576
Q3	0,75360	0,0493	0,73040	0,03882	0,23870	0,00666
Máx.	0,93980	0,0866	8,84200	0,35140	0,43620	0,01289
e.p.	0,00624	0,0003	0,02193	0,00781	0,00240	0,00014

As figuras 4.17 e 4.18 exibem a comparação entre valores teóricos e ajustados, respectivamente, para o nível α_t e para as funções de resposta ao impulso e de transferência (exponenciais).

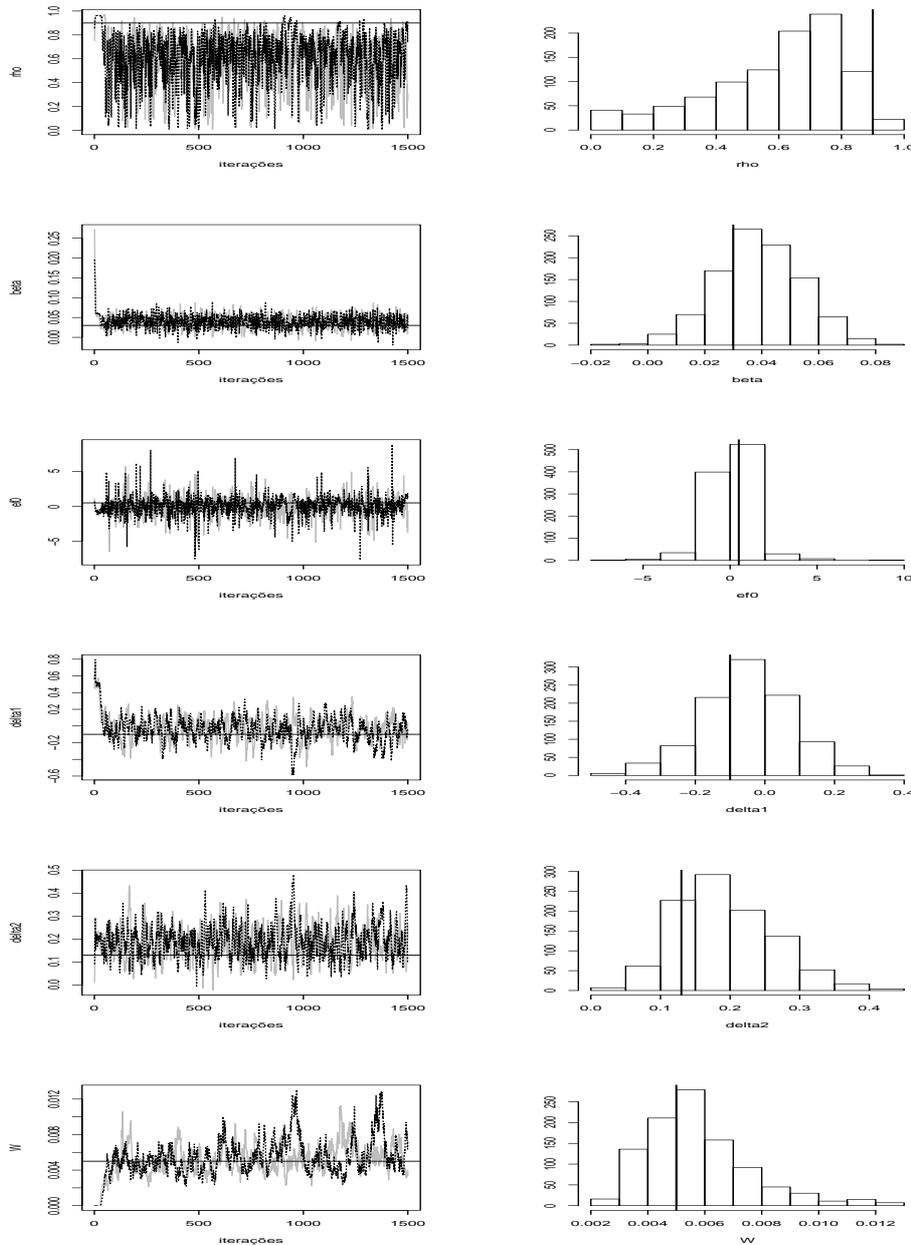


Figura 4.16 Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem ($r=1, s=0$) com nível dinâmico .

O nível foi bem recuperado pelo processo de estimação, mas, claramente, as funções de resposta ao impulso e de transferência são afetadas pela dificuldade em identificar o parâmetro autoregressivo ρ . A função de resposta média, λ_t , parece bem estimada, como mostra a figura 4.19. Na figura 4.20 encontram-se histogramas das amostras das distribuições preditivas condicionais a X_{t+h} , para horizontes $h = 1$ a 30.

Tabela 4.12 Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r=1, s=0$) com nível dinâmico.

	ρ	β	E_0	δ_1	δ_2	W
ρ	1,00	0,03	0,06	0,18	-0,04	-0,16
β	0,03	1,00	-0,03	0,11	-0,09	-0,12
E_0	0,06	-0,03	1,00	0,02	0,03	-0,01
δ_1	0,18	0,11	0,02	1,00	-0,35	-0,24
δ_2	-0,04	-0,09	0,03	-0,35	1,00	0,35
W	-0,16	-0,12	-0,01	-0,24	0,35	1,00

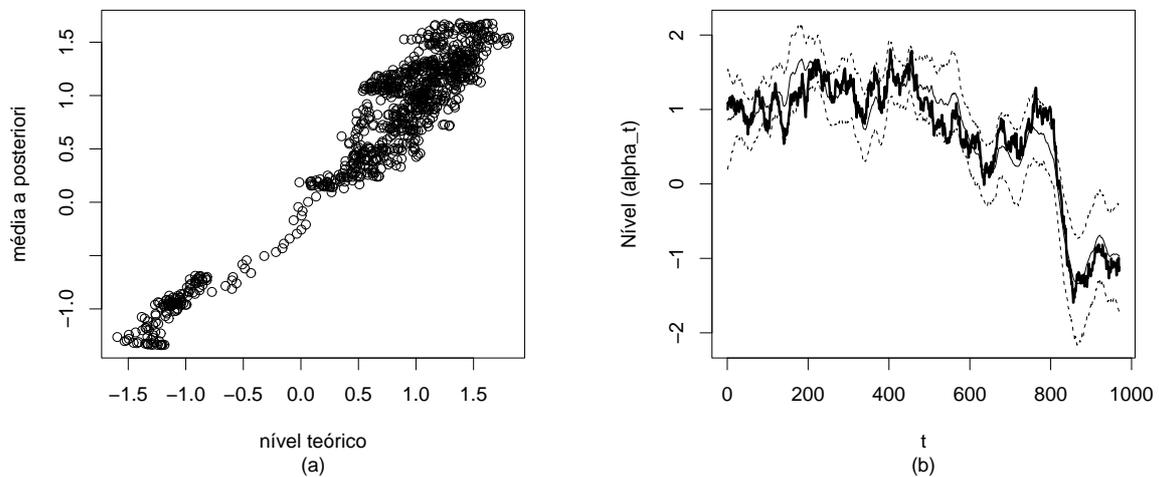


Figura 4.17 Resposta Poisson, modelo de ordem ($r=1, s=0$) com nível dinâmico. Nível α_t : (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços de α_t teórico (linha em negro), estimativas da média a posteriori e de limites de credibilidade a 95 %.

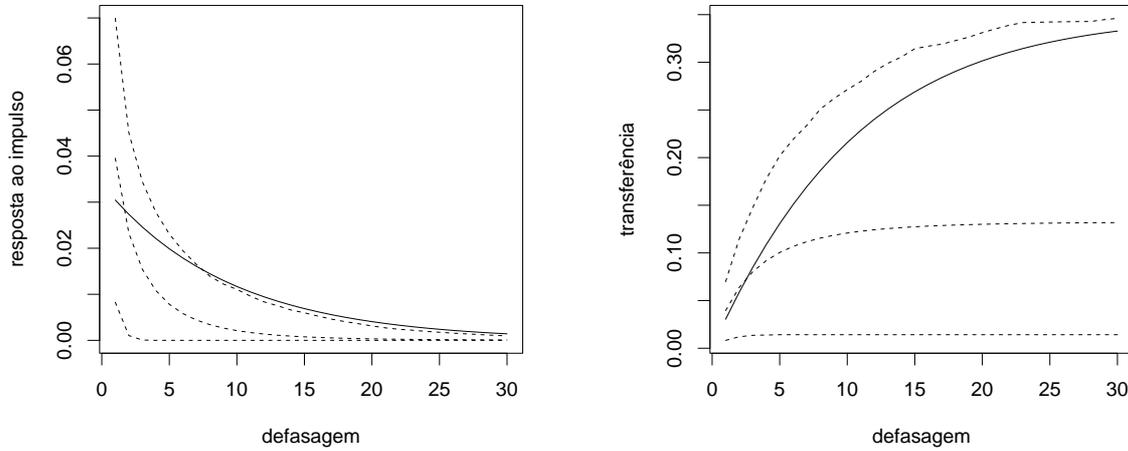


Figura 4.18 Resposta Poisson, modelo de ordem $(r=1, s=0)$ com nível dinâmico: Função de resposta ao impulso e de transferência - teóricas (linha sólida), médias e limites de credibilidade estimados a 95% (linhas tracejadas).

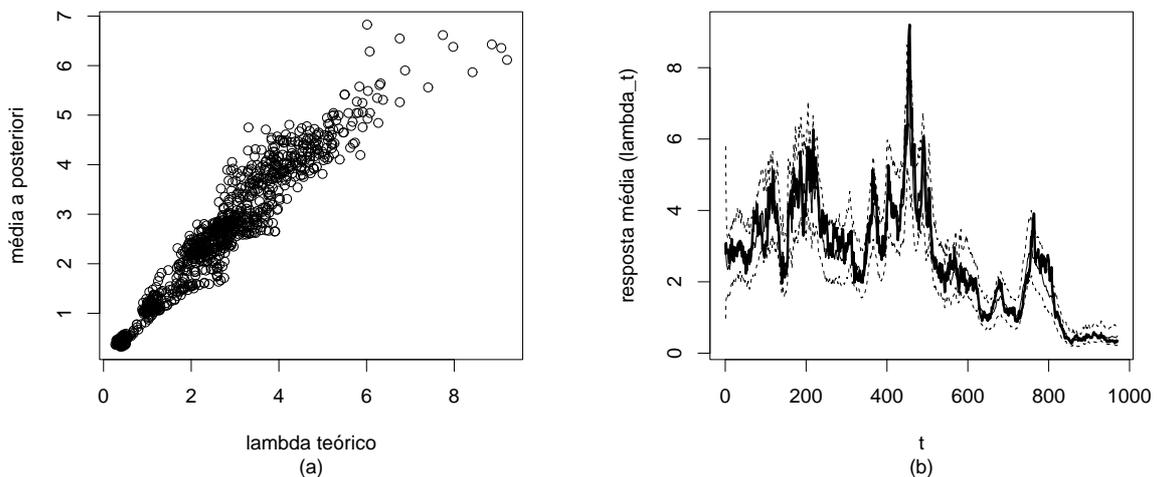


Figura 4.19 Resposta Poisson, modelo de ordem $(r=1, s=0)$ com nível dinâmico. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

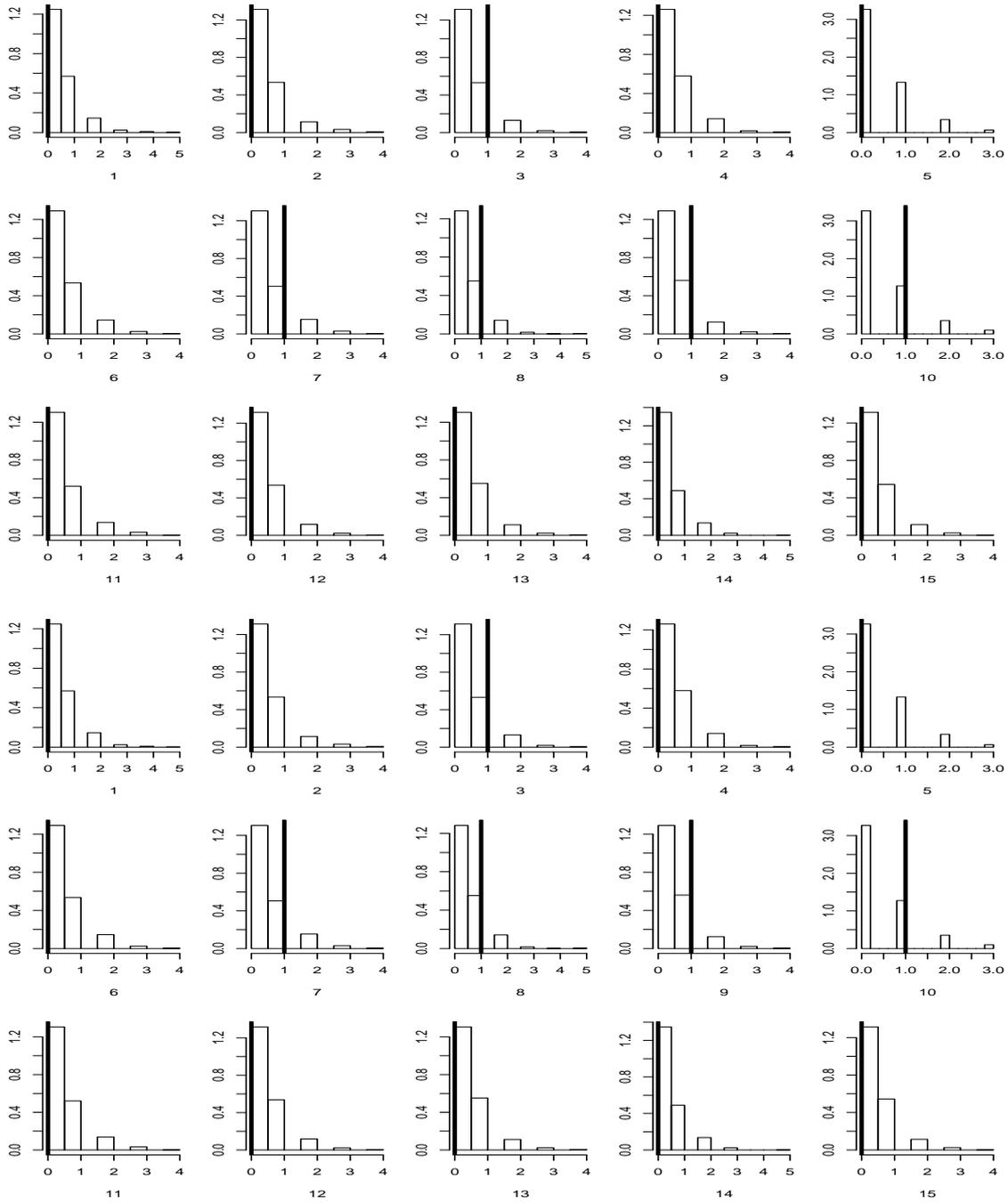


Figura 4.20 Resposta Poisson, modelo de ordem ($r=1,s=0$) com nível dinâmico : Histogramas das amostras das distribuições preditivas para horizontes variando de 1 a 30.

4.1.5 Função de transferência de ordem ($r = 1, s = 0$) com fator de ganho dinâmico

Pressupondo-se que o choque imediato da regressora X_t sobre a resposta esperada possa variar no tempo e assumindo-se, com tal propósito, a estrutura ditada pelas equações (3.14) e (3.15), obtém-se, como mencionado na subseção 3.3.4, funções de resposta ao impulso que a cada instante t podem ter formatos e magnitudes diferentes, de acordo com o valor de β_t . Seguindo tais pressupostos, simulou-se um conjunto de dados artificiais, adotando-se os seguintes valores para os parâmetros de interesse: $\alpha = 1,0$, $\rho = 0,7$, $E_0 = 0,1$, $\delta_1 = -0,05$, $\delta_2 = 0,05$ e passeio aleatório para os fatores de ganho governados pela variância $Q = 0,0005$.

No procedimento MCMC, como já visto anteriormente, a condicional completa de Q é uma Gama inversa. Para os parâmetros restantes, que não têm condicional completa conhecida, foi necessário inserir passos Metropolis-Hastings no amostrador de Gibbs. Os erros de evolução da equação (3.15) foram gerados individualmente, a partir das propostas baseadas em (3.17) e (3.18). A mesma proposta foi utilizada para os demais parâmetros, gerados individualmente, à exceção de δ_1 e δ_2 , gerados em bloco. A figura 4.21 exhibe os traços das cadeias geradas (a saltos de 20 iterações) e histogramas das amostras após suposta a convergência.

A tabela 4.13 exhibe estatísticas sumárias das amostras da distribuição a posteriori. Os parâmetros α e ρ são particularmente bem estimados pela média. Os parâmetros δ_1 e δ_2 são superestimados, com intervalos de credibilidade estimados a 95%, respectivamente: $(-0,046; 0,042)$ e $(0,047; 0,143)$. A cadeia gerada para Q exhibe forte autocorrelação, como se pode ver na figura 4.21, mas o erro-padrão associado à média estimada a posteriori é bastante baixo, indicando que o tamanho da amostra final, mesmo sob presença de autocorrelação, foi suficiente para a estimação deste parâmetro. Na tabela 4.14, destaca-se a correlação a posteriori entre os parâmetros ρ e Q .

Tabela 4.13 Resposta Poisson: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem ($r=1, s=0$) com fator de ganho dinâmico.

	α	ρ	Q	E_0	δ_1	δ_2
Teórico	1,0	0,7	0,0005	0,1	-0,05	0,05
Mín	0,91710	0,48190	0,00006	-3,08600	-0,07227	-0,00612
Q1	0,99530	0,68190	0,00020	-0,32860	-0,01857	0,07900
Mediana	1,01300	0,72380	0,00029	0,07496	-0,00292	0,09499
Média	1,01300	0,71920	0,00031	0,04271	-0,00266	0,09511
Q3	1,03000	0,76040	0,00039	0,44850	0,01220	0,11220
Máx.	1,11100	0,87410	0,00085	1,92200	0,08234	0,17680
e.p.	0,00038	0,00292	0,00001	0,00840	0,00024	0,00033

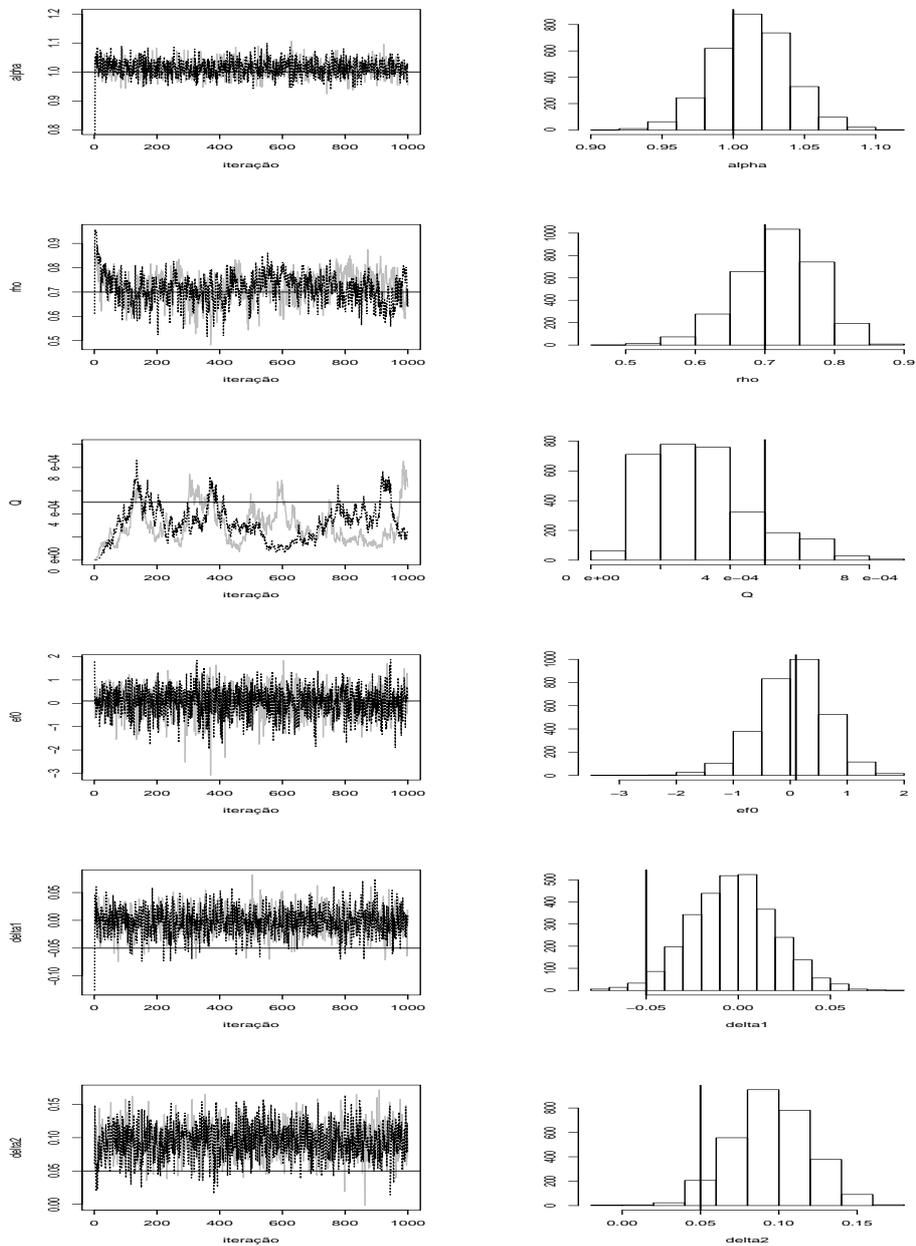


Figura 4.21 Resposta Poisson: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem ($r=1, s=0$) com fator de ganho dinâmico.

Tabela 4.14 Resposta Poisson: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r=1, s=0$) com fator de ganho dinâmico.

	α	ρ	Q	E_0	δ_1	δ_2
α	1,00	0,00	-0,07	-0,10	-0,03	-0,23
ρ	0,00	1,00	-0,48	0,02	0,08	0,06
Q	-0,07	-0,48	1,00	0,02	-0,03	-0,04
E_0	-0,10	0,02	0,02	1,00	-0,13	0,00
δ_1	-0,03	0,08	-0,03	-0,13	1,00	-0,19
δ_2	-0,23	0,06	-0,04	0,00	-0,19	1,00

Pode-se verificar, na figura 4.21, que a distribuição a posteriori de Q é assimétrica positiva, colocando maior massa na região abaixo do valor teórico postulado para a criação dos dados. Em decorrência da subestimativa da variância de evolução, as estimativas para β_t resultam mais suaves que os valores teóricos, como ilustra a figura 4.22. Ainda assim, exceto por alguns poucos pontos, os valores teóricos encontram-se dentro dos limites de credibilidade estimados a 95%.

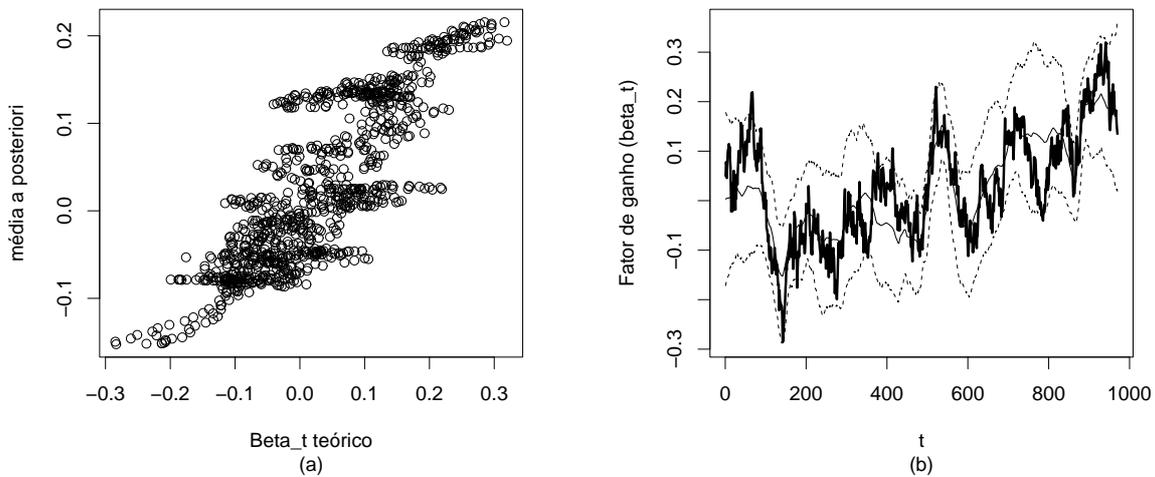


Figura 4.22 Resposta Poisson, modelo de ordem ($r=1, s=0$) com nível dinâmico. Fator de ganho β_t : (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços de β_t teórico (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

A figura 4.23 exibe a função de resposta média, λ_t , cuja estimação apresenta-se bastante satisfatória. Na figura 4.24, exibem-se as funções de resposta ao impulso para três instantes, sendo seus valores teóricos muito bem estimados nos três casos. Na figura 4.25, estão dispostos os histogramas das amostras das distribuições preditivas condicionais a X_{970+h} , $h = 1, 2, \dots, 30$. Os intervalos de credibilidade a 95% para as previsões capturam os valores reais de y_{971}, \dots, y_{1000} . Com exceção dos instantes 21 e 22 do horizonte de previsão, nos quais os

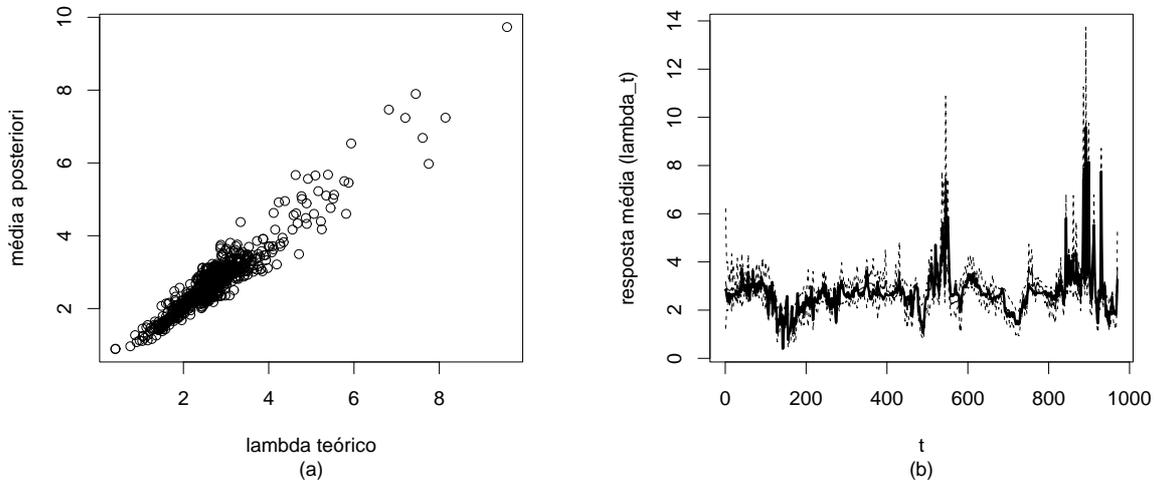


Figura 4.23 Resposta Poisson, modelo de ordem ($r=1, s=0$) com fator de ganho dinâmico. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

valores reais da resposta (y_{991} e y_{992}) são atipicamente elevados, se comparados ao restante do horizonte.

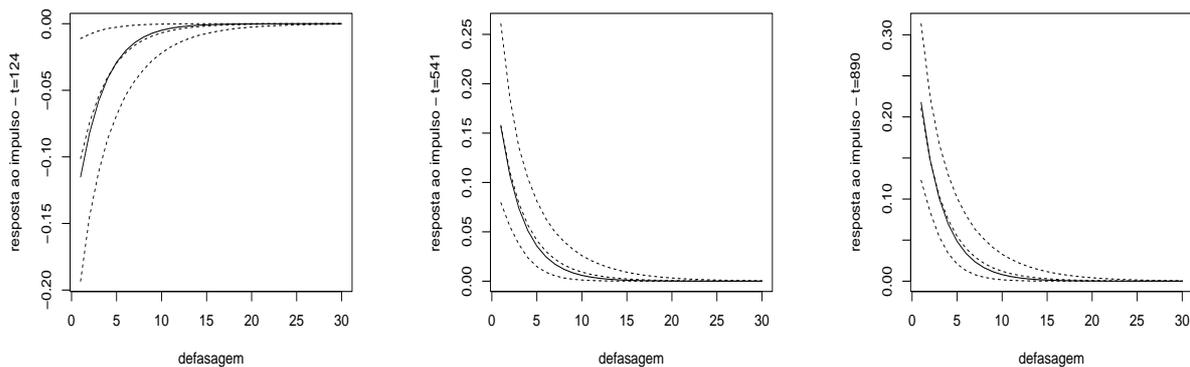


Figura 4.24 Resposta Poisson, modelo de ordem ($r=1, s=0$) com fator de ganho dinâmico, função de resposta ao impulso nos instantes $t = 124$, $t = 541$ e $t = 890$: função teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.

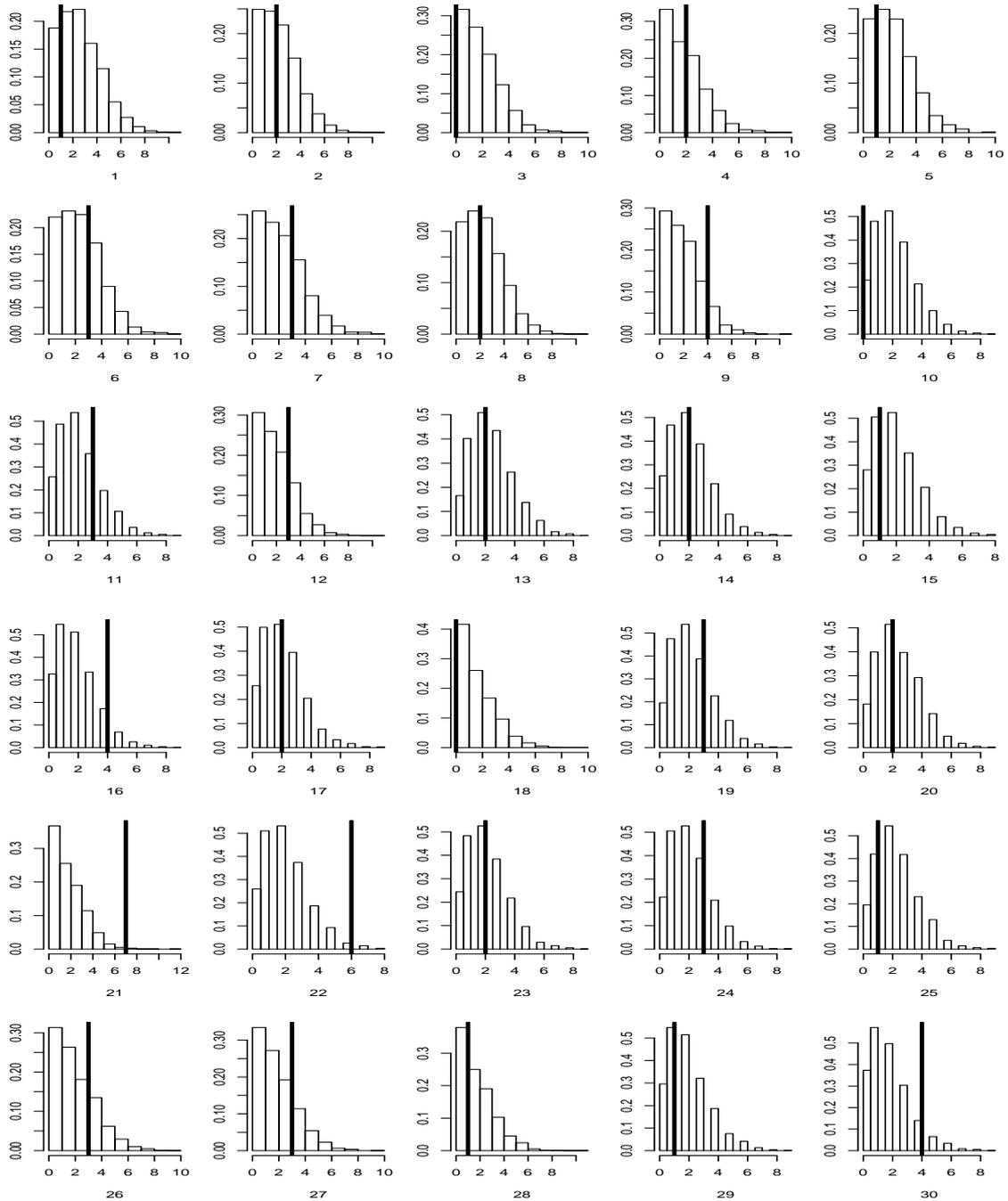


Figura 4.25 Resposta Poisson, modelo de ordem ($r=1, s=0$) com fator de ganho dinâmico: Histogramas das amostras das distribuições preditivas para horizontes variando de 1 a 30.

4.2 Resposta Binomial

Suponham-se modelos da seguinte forma:

$$y_t \sim \text{Binomial}(n, p_t)$$

$$\log\left(\frac{p_t}{1-p_t}\right) = \eta_t,$$

4.2.1 Função de transferência de ordem ($r = 1, s = 0$)

Suponha-se a seguinte forma para o preditor:

$$\eta_t = \alpha + E_t + \delta_1 Z_{1t} + \delta_2 Z_{2t}$$

$$E_t = \rho E_{t-1} + \beta X_t.$$

Gerou-se um conjunto de dados artificiais a partir da formulação acima, fixando-se $n = 5$ e atribuindo-se os seguintes valores aos parâmetros de interesse: $\alpha = 1,0$, $\beta = 0,05$, $\rho = 0,9$, $E_0 = 0,5$, $\delta_1 = -0,1$, $\delta_2 = 0,1$.

Adotaram-se prioris vagas para todos os parâmetros e, no algoritmo Metropolis em Gibbs, aplicou-se a proposta dada pela condicional completa de cada parâmetro, no modelo de trabalho (2.9) baseado em (3.34) e (3.35), para $\alpha, \beta, E_0, \delta_1$ e δ_2 , sendo estes dois últimos gerados conjuntamente. O logito de ρ foi gerado de uma proposta passeio-aleatório.

Os traços das cadeias geradas e histogramas das amostras finais da posteriori, após convergência, são exibidos na figura 4.26. Os parâmetros ρ e β apresentam forte correlação a posteriori, como indica a tabela 4.16, sendo ρ subestimado e β superestimado por suas estimativas pontuais - exibidas na tabela 4.15. Ainda assim, os valores teóricos de ambos os parâmetros encontram-se dentro dos limites de credibilidade estimados a 95%, dados, respectivamente, por : (0,683; 0,929) e (0,044; 0,130), para ρ e β . Conseqüentemente, as estimativas pontuais das funções de resposta ao impulso e de transferência - dadas por suas médias a posteriori estimadas - não são acuradas. Entretanto, também estas funções têm seus valores teóricos dentro dos limites de credibilidade estimados a 95%, como mostrado na figura 4.27.

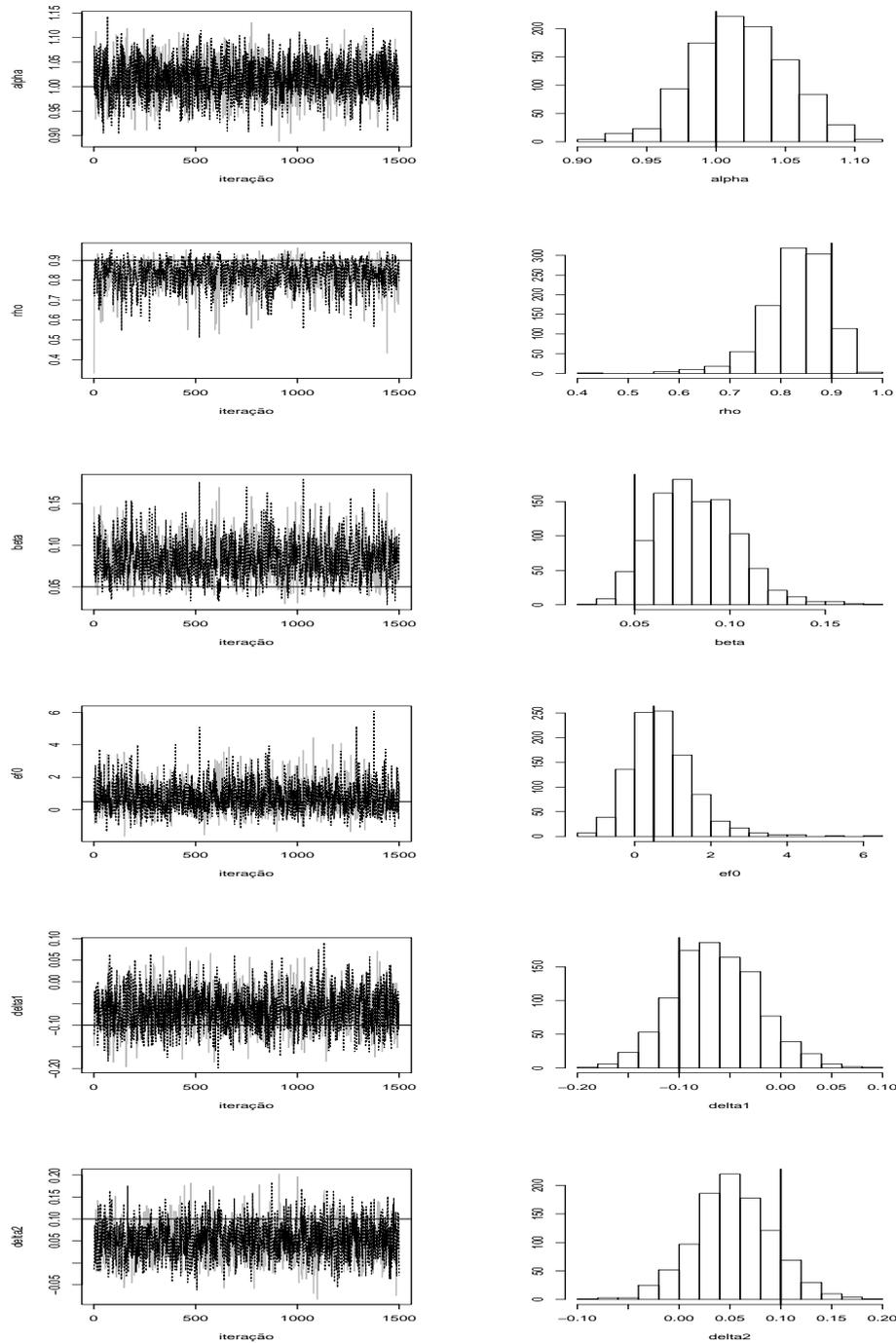


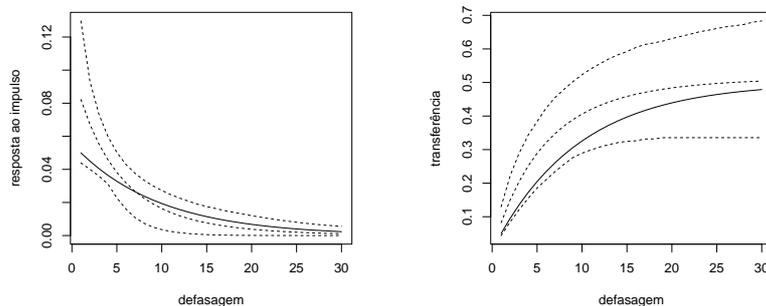
Figura 4.26 Resposta Binomial: Traços das cadeias geradas e histogramas da amostra da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).

Tabela 4.15 Resposta Binomial: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem ($r = 1, s = 0$).

	α	ρ	β	E_0	δ_1	δ_2
Teórico	1,0	0,9	0,05	0,5	-0,1	0,1
Mín	0,90860	0,43200	0,02830	-1,36000	-0,18400	-0,08332
Q1	0,99310	0,79820	0,06660	0,14190	-0,09175	0,02865
Mediana	1,01600	0,83920	0,08019	0,62270	-0,06534	0,05252
Média	1,01700	0,83210	0,08222	0,71010	-0,06357	0,05327
Q3	1,04100	0,87690	0,09682	1,16500	-0,03545	0,07714
Máx.	1,11900	0,96290	0,17870	6,08100	0,09082	0,19620
e.p.	0,00072	0,00099	0,00021	0,01192	0,00040	0,00074

Tabela 4.16 Resposta Binomial: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r = 1, s = 0$).

	α	ρ	β	E_0	δ_1	δ_2
α	1,00	-0,32	0,16	0,09	-0,27	-0,26
ρ	-0,32	1,00	-0,84	-0,38	0,42	0,42
β	0,16	-0,84	1,00	0,31	-0,17	-0,21
E_0	0,09	-0,38	0,31	1,00	-0,29	-0,17
δ_1	-0,27	0,42	-0,17	-0,29	1,00	0,10
δ_2	-0,26	0,42	-0,21	-0,17	0,10	1,00

**Figura 4.27** Resposta Binomial, modelo de ordem ($r=1, s=0$) função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.

4.2.2 Função de transferência de ordem ($r = 1, s = 1$)

Supondo-se

$$E_t = \rho E_{t-1} + \beta_0 X_t + \beta_1 X_{t-1},$$

gerou-se um conjunto de 1000 observações artificiais, fixando-se $n = 5$ e atribuindo-se aos parâmetros do modelo os seguintes valores: $\alpha = 1,0$, $\beta_0 = 0,05$, $\beta_1 = 0,10$, $\rho = 0,9$, $E_0 = 0,5$, $\delta_1 = -0,1$ e $\delta_2 = 0,1$.

Adotaram-se prioris vagas para todos os parâmetros e procedeu-se ao algoritmo MCMC com geração individual de α , ρ e E_0 e geração dos blocos (β_0, β_1) e (δ_1, δ_2) . Todos os parâmetros foram gerados da condicional completa obtida do modelo de trabalho (2.9), exceto por ρ , que teve seu logito gerado de uma proposta passeio-aleatório. A figura 4.28 exibe os traços das cadeias geradas (a cada 20 iterações) e os histogramas das amostras finais.

Como se pode ver na tabela 4.17, as estimativas, de forma geral, são satisfatórias e a incerteza associada ao processo de estimação é bastante pequena, como indicam os erros-padrão associados às médias a posteriori estimadas. A estimação do parâmetro que representa o efeito acumulado da regressora X no instante inicial da análise - E_0 - exibe, como nas aplicações anteriores, maior incerteza que os demais, mas ainda assim seu erro-padrão não é grande se comparado à magnitude deste parâmetro.

A estrutura de correlação a posteriori, exibida na tabela 4.18, apresenta-se bastante semelhante àquela obtida no ajuste do mesmo tipo de função de transferência para respostas Poisson, estando o parâmetro β_1 correlacionado a β_0 e a ρ .

Tabela 4.17 Resposta Binomial: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem ($r=1, s=1$).

	α	ρ	β_0	β_1	E_0	δ_1	δ_2
Teórico	1,0	0,9	0,05	0,1	0,5	-0,1	0,1
Mín.	0,90450	0,84370	-0,02625	0,06068	-1,74900	-0,22000	-0,05164
Q1	0,99380	0,88150	0,01478	0,11620	0,02178	-0,10380	0,07367
Mediana	1,01900	0,89320	0,02572	0,13340	0,36410	-0,07252	0,10490
Média	1,01900	0,89150	0,02632	0,13330	0,38220	-0,07385	0,10320
Q3	1,04700	0,90220	0,03820	0,14980	0,73270	-0,04530	0,13320
Máx.	1,13400	0,92450	0,07748	0,22590	2,73700	0,05650	0,22160
e.p.	0,00113	0,00112	0,00034	0,00045	0,03104	0,00075	0,00099

As funções de resposta ao impulso e de transferência estimadas, junto a seus valores teóricos, são apresentadas na figura 4.29, sendo ambas muito bem estimadas. O mesmo ocorre com a função de resposta média, np_t , cujos valores teóricos e estimados são exibidos na figura 4.30.

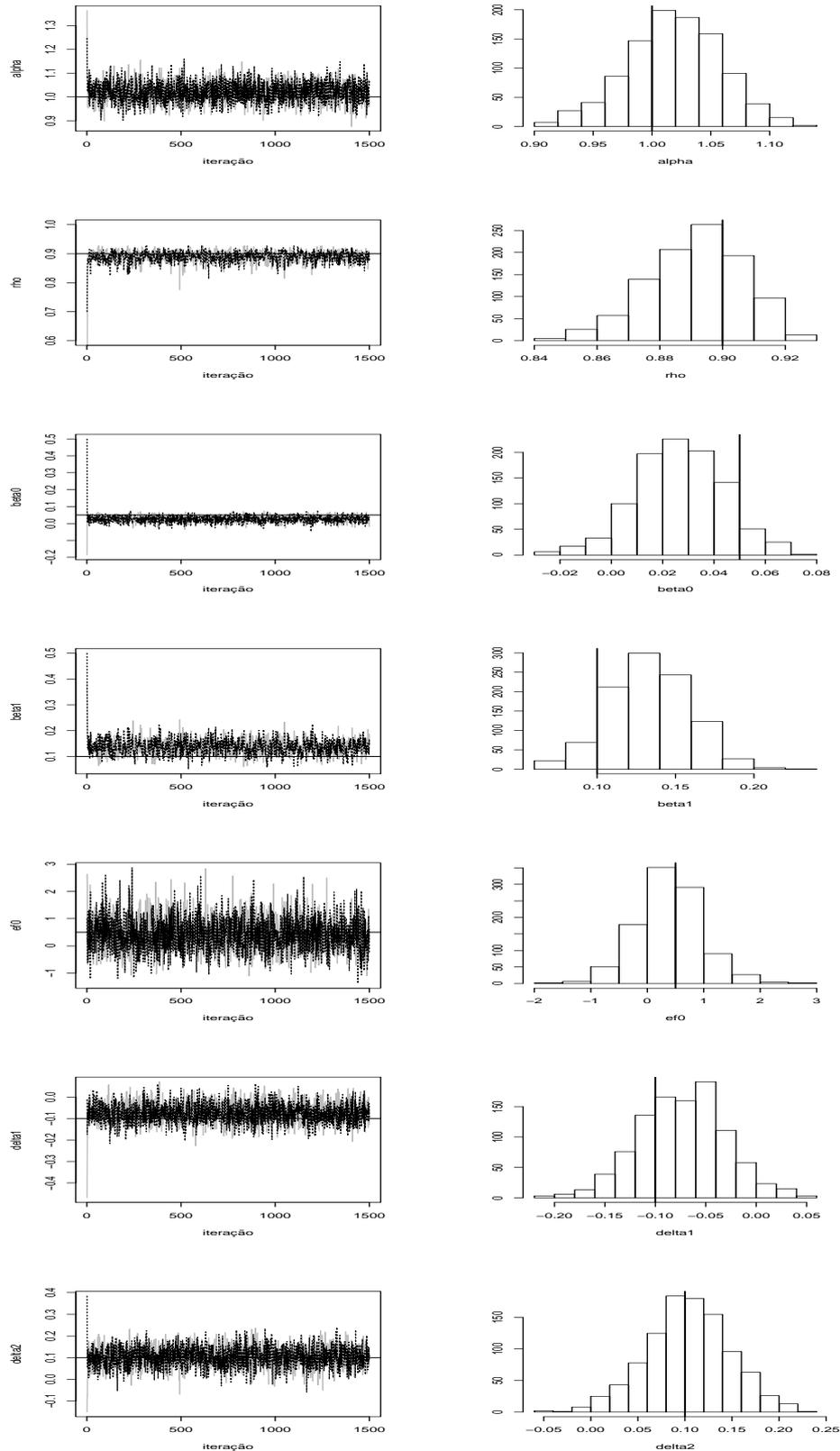


Figura 4.28 Resposta Binomial: Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de função de transferência de ordem $r = 1, s = 1$.

Tabela 4.18 Resposta Binomial: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r=1, s=1$).

	α	ρ	β_0	β_1	E_0	δ_1	δ_2
α	1,00	-0,36	-0,05	0,24	0,02	-0,24	-0,31
ρ	-0,36	1,00	0,12	-0,71	-0,15	0,35	0,42
β_0	-0,05	0,12	1,00	-0,69	0,03	0,14	0,12
β_1	0,24	-0,71	-0,69	1,00	0,04	-0,15	-0,25
E_0	0,02	-0,15	0,03	0,04	1,00	-0,23	-0,09
δ_1	-0,24	0,35	0,14	-0,15	-0,23	1,00	0,09
δ_2	-0,31	0,42	0,12	-0,25	-0,09	0,09	1,00

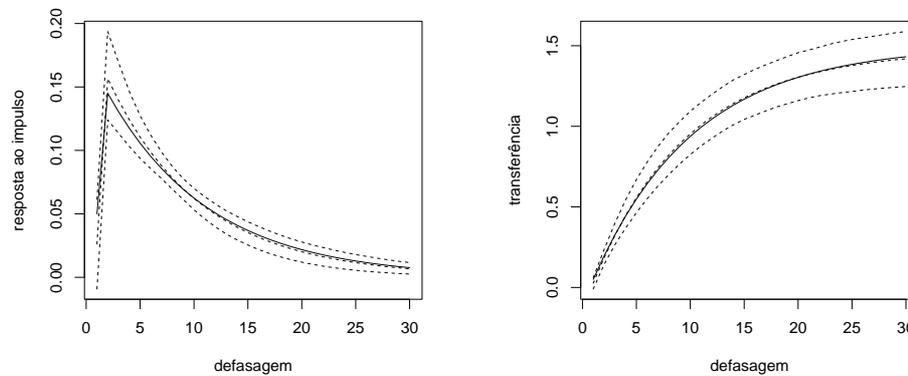


Figura 4.29 Resposta Binomial, modelo de ordem ($r=1, s=1$) função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.

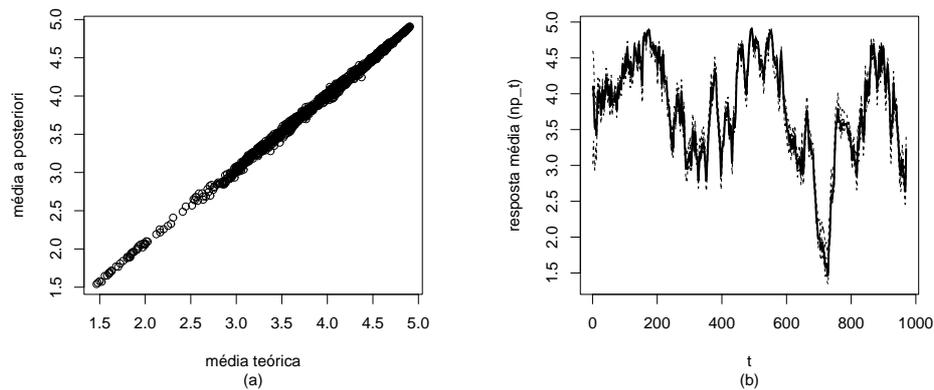


Figura 4.30 Resposta Binomial, modelo de ordem ($r=1, s=1$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

4.2.3 Função de transferência de ordem ($r = 1, s = 0$) com erro aleatório

Acrescentando-se um termo de erro à função de transferência de primeira ordem, tem-se:

$$E_t = \rho E_{t-1} + \beta X_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, Q_\varepsilon).$$

Assumindo-se a formulação acima, gerou-se um conjunto com 1000 observações, a partir da seguinte definição para o vetor paramétrico: $\alpha = 1,0$, $\beta = 0,05$, $\rho = 0,9$, $E_0 = 0,5$, $\delta_1 = -0,1$, $\delta_2 = 0,1$, $Q_\varepsilon = 0,07$.

Tabela 4.19 Resposta Binomial: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem ($r=1, s=0$) com erro aleatório.

	α	ρ	β	E_0	δ_1	δ_2	Q_ε
Teórico	1,0	0,9	0,05	0,5	-0,1	0,1	0,07
Mín	0,87480	0,49000	0,0246	-1,54500	-0,20510	-0,07063	0,00000
Q1	0,98030	0,77970	0,0681	0,43960	-0,10710	0,02492	0,00005
Mediana	1,00200	0,82390	0,0812	0,94210	-0,07938	0,05034	0,00018
Média	1,00300	0,81660	0,0826	1,04800	-0,07884	0,05037	0,00086
Q3	1,02700	0,85960	0,0954	1,57400	-0,05087	0,07507	0,00068
Máx.	1,10900	0,95860	0,1729	4,63400	0,03283	0,16780	0,01163
e.p.	0,00046	0,00289	0,0004	0,04146	0,00131	0,00102	0,00020

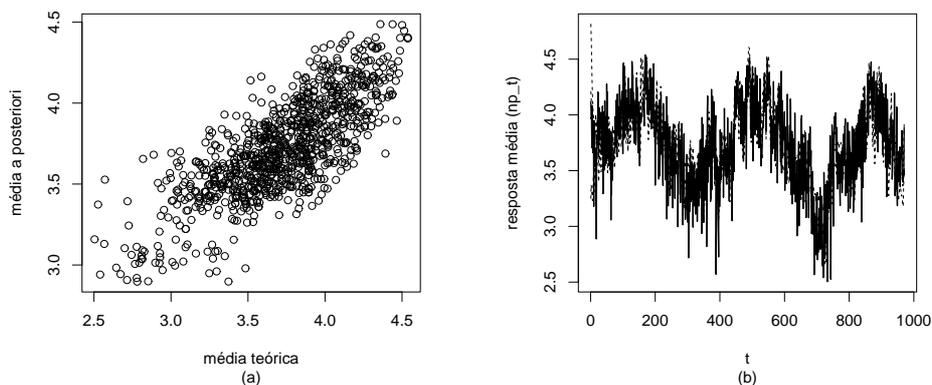


Figura 4.31 Resposta Binomial, modelo de ordem ($r=1, s=0$) com erro aleatório. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

Estatísticas sumárias das amostras obtidas via MCMC são apresentadas na tabela 4.19 e constata-se que as estimativas pontuais da variância dos erros ficam muito aquém do seu valor

teórico. De fato, a distribuição a posteriori não é capaz de captar o verdadeiro valor de Q_ε , sendo seu intervalo de credibilidade estimado a 95%: (0,00000945; 0,00570). A subestimativa de Q_ε compromete a obtenção da função de resposta média, como se vê na figura 4.31.

Repetiu-se o procedimento de estimação para dados gerados com os mesmos valores para o vetor paramétrico, porém assumindo-se $n = 50$. Nesse caso foi possível identificar o valor teórico da variância dos erros. Gráficos das trajetórias das cadeias geradas e histogramas das amostras da posteriori encontram-se na figura 4.32.

Os parâmetros ρ e β apresentam forte correlação a posteriori (-0,83) e o primeiro fica ligeiramente subestimado, com intervalo de credibilidade a 95% dado por (0,834; 0,898), enquanto β é superestimado, com intervalo de credibilidade a 95% (0,056; 0,083). Conseqüentemente, as funções de resposta ao impulso e de transferência são mal estimadas, como se pode ver na figura 4.33. As estimativas, exibidas na figura 4.34, para a função de resposta média são satisfatórias.

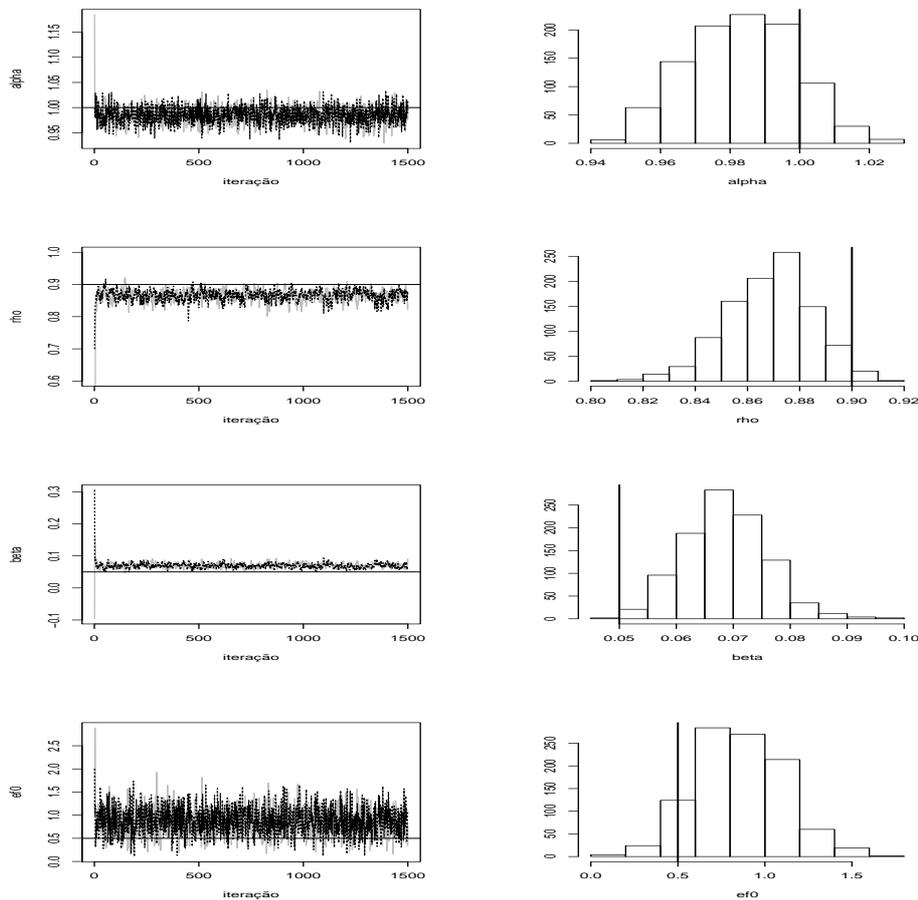


Figura 4.32 Resposta Binomial(50, p_t): Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de função de transferência de ordem ($r=1, s=0$) com erro aleatório.

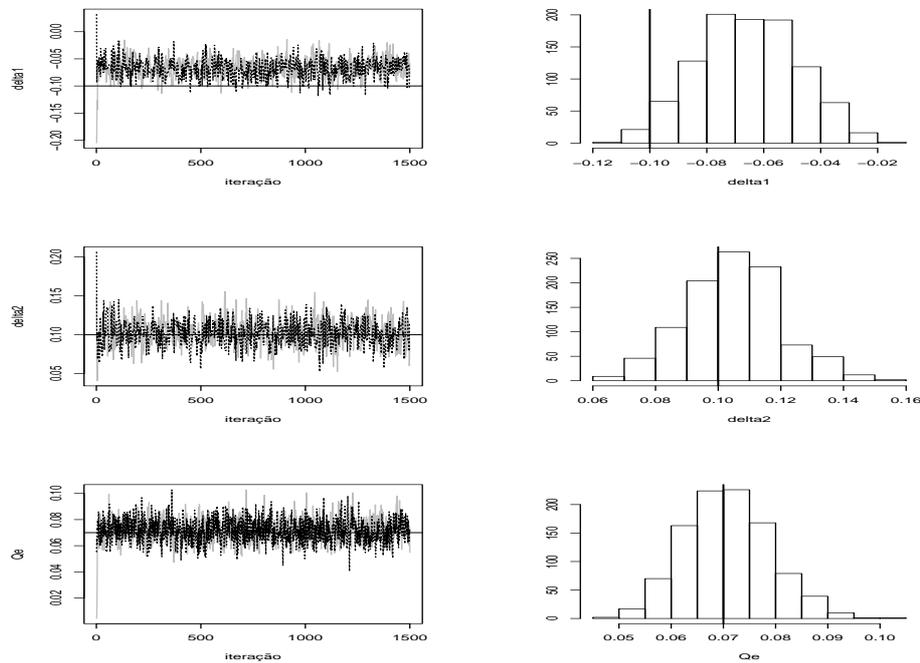


Figura 4.32 Resposta Binomial($50, p_t$): Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados de acordo com o modelo de função de transferência de ordem ($r=1, s=0$) com erro aleatório (continuação).

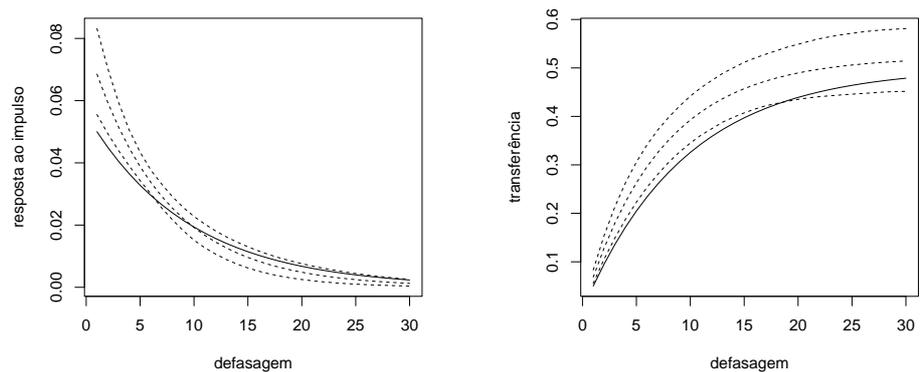


Figura 4.33 Resposta Binomial($50, p_t$), modelo de ordem ($r=1, s=0$) com erro aleatório: função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95 %.

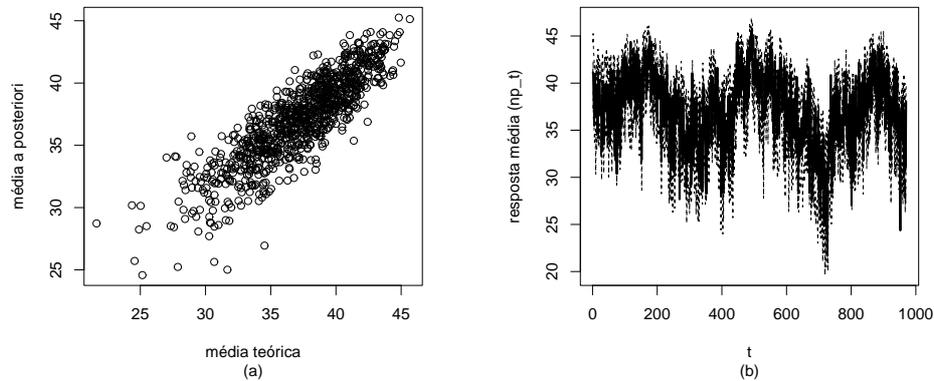


Figura 4.34 Resposta Binomial ($50, p_t$), modelo de ordem ($r=1, s=0$) com erros aleatórios. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

4.2.4 Função de transferência de ordem ($r = 1, s = 0$) e nível dinâmico

Pressupondo-se o preditor

$$\begin{aligned}\eta_t &= \alpha_t + E_t + \delta_1 Z_{1t} + \delta_2 Z_{2t} \\ \alpha_t &= \alpha_{t-1} + u_t, \quad u_t \sim N(0, W), \\ E_t &= \rho E_{t-1} + \beta X_t,\end{aligned}$$

simulou-se um conjunto de dados composto por 1000 pontos, a partir de $n = 5$ e dos seguintes valores para os parâmetros envolvidos: $\beta = 0,05$, $\rho = 0,9$, $E_0 = 0,5$, $\delta_1 = -0,1$, $\delta_2 = 0,1$, $W = 0,005$.

Como se vê na figura 4.35, diferentemente do que ocorreu ao se aplicar um nível dinâmico à resposta Poisson, no caso Binomial foi possível identificar o parâmetro autoregressivo, tendo-se, então, como mostra a figura 4.36, estimativas muito boas para as funções de resposta ao impulso e de transferência, as quais dependem, nesse caso, apenas de ρ e β . A variância da equação de sistema, entretanto, é subestimada, produzindo estimativas excessivamente suaves para o nível α_t , como visto na figura 4.37 e, conseqüentemente, obtêm-se estimativas suaves para a função de resposta média, mas, ainda assim, com intervalos de credibilidade englobando os valores teóricos em praticamente todos os pontos.

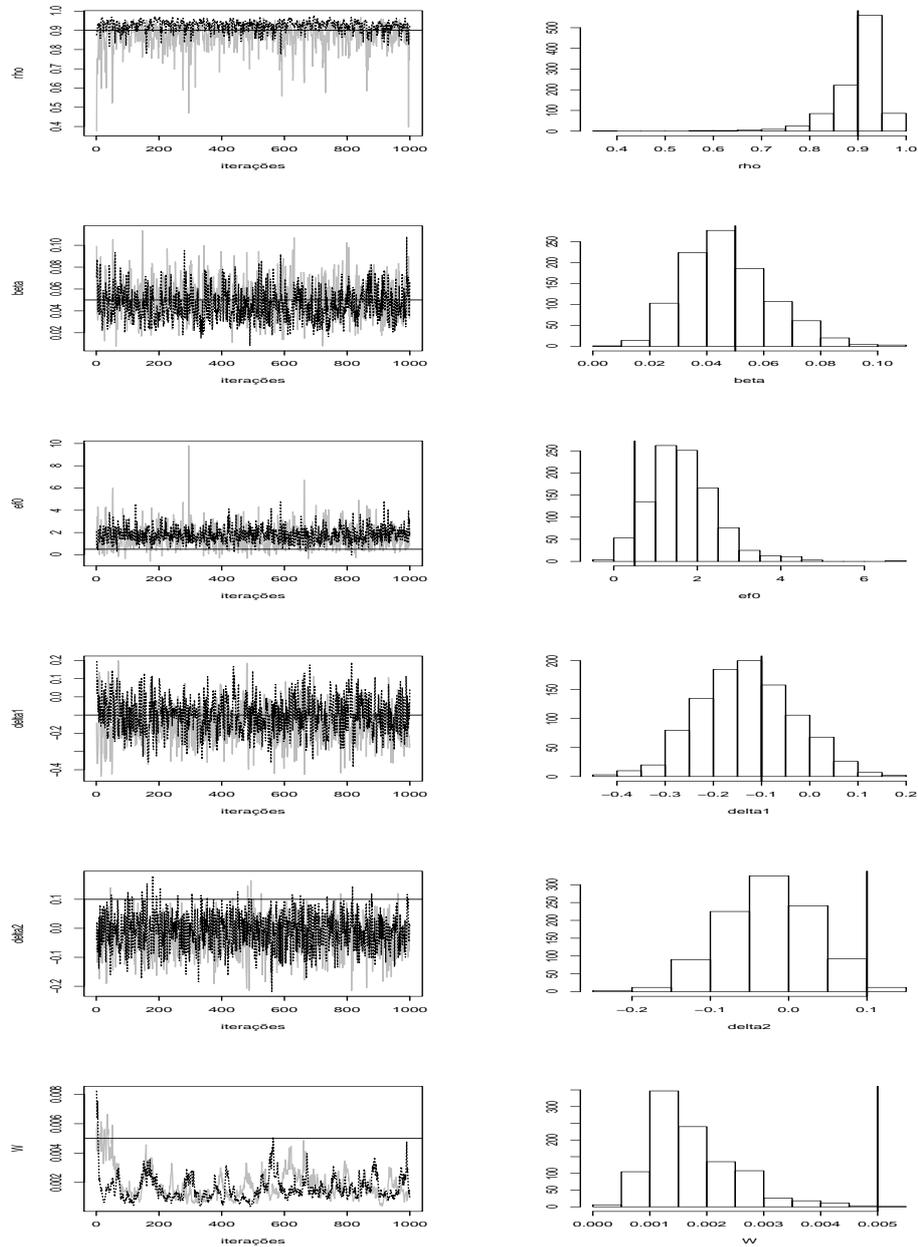


Figura 4.35 Resposta Binomial($5, p_1$): Traços das cadeias geradas e histogramas das amostras da posteriori, para dados simulados segundo o modelo autoregressivo de ordem ($r=1, s=0$) com nível dinâmico.

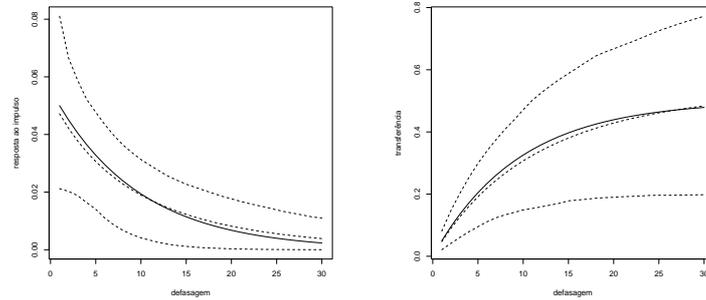


Figura 4.36 Resposta Binomial($5, p_t$), modelo de ordem ($r=1, s=0$) com nível dinâmico: função de resposta ao impulso teórica (linha sólida), estimativas da média a posteriori e de limites de credibilidade a 95%.

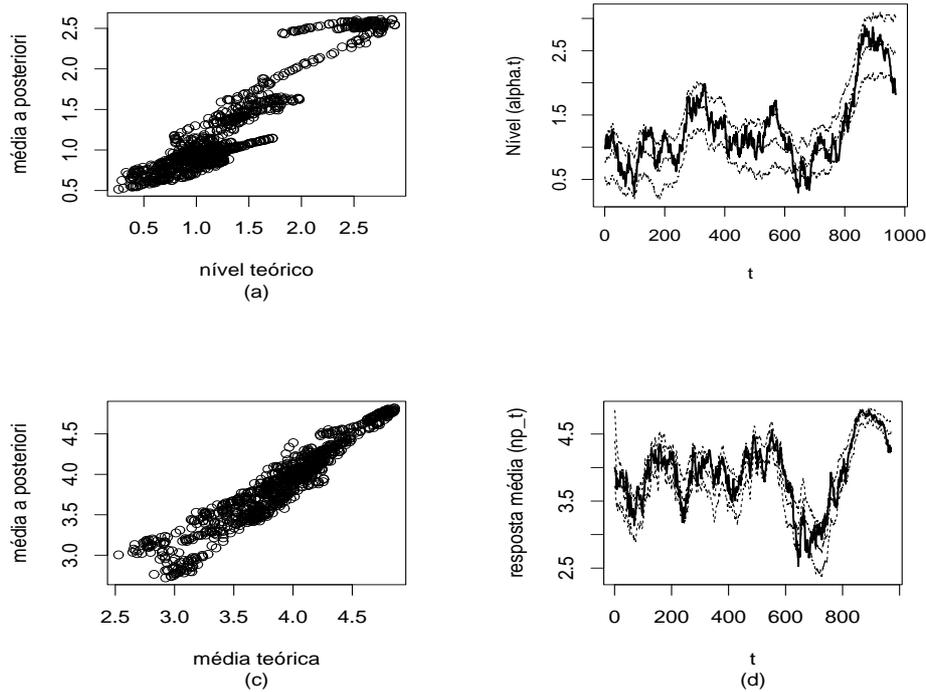


Figura 4.37 Resposta Binomial($5, p_t$), modelo de ordem ($r=1, s=0$) com nível dinâmico. Nível α_t : (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços de α_t teórico (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95%. Função de resposta média: (c) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (d) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95%.

4.3 Resposta Gama

Considere-se, nesta seção, a seguinte estrutura geral para os modelos propostos:

$$\begin{aligned} y_t &\sim \text{Gama}(\varphi, \lambda_t) \\ \log\left(\frac{\varphi}{\lambda_t}\right) &= \eta_t, \end{aligned} \quad (4.4)$$

4.3.1 Função de transferência de ordem ($r = 1, s = 0$)

Como visto anteriormente, o modelo estático de ordem ($r = 1, s = 0$) tem preditor dado pela equação (4.2). Utilizando esse preditor (com apenas uma variável de controle, Z_1) e a estrutura observacional (4.4), foram geradas 1000 observações artificiais, sendo 970 destas utilizadas para estimação do modelo e as 30 restantes, para avaliação da capacidade preditiva. A figura 4.38 exibe as trajetórias da regressora X_t e da covariável Z_t utilizadas ao longo das seções 4.3 e 4.4.

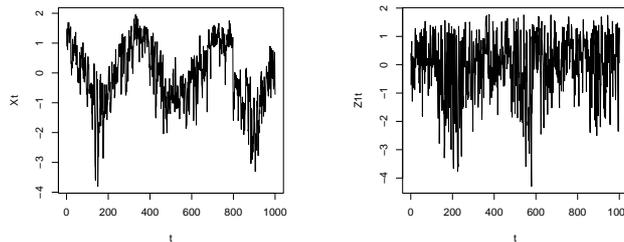


Figura 4.38 Regressora X_t e covariável Z_t , utilizadas nas simulações para respostas Gama e Gaussiana Inversa.

O conjunto de valores y_1, \dots, y_{1000} foi simulado a partir dos seguintes valores para o vetor paramétrico: $\varphi = 7,0$, $\alpha = 1,5$, $\rho = 0,9$, $\beta = -0,02$, $E_0 = 0,2$, $\delta = -0,12$. A estimação foi feita usando o algoritmo MCMC descrito no capítulo 3, com geração individual de cada parâmetro, a partir de propostas passeio aleatório para $\log(\varphi)$, α , $\log(\rho)$, E_0 e δ e proposta baseada na verossimilhança para β . As distribuições a priori para todos os parâmetros foram não-informativas: $\alpha, \beta, E_0, \delta_1, \delta_2 \stackrel{iid}{\sim} N(0, 10^5)$, $\rho \sim U(0, 1)$, $\varphi \sim G(10^{-5}, 10^{-5})$.

A figura 4.39 exibe as trajetórias de duas cadeias paralelas para cada um dos parâmetros, com saltos de 20 iterações e observa-se comportamento convergente após, aproximadamente, 10000 iterações. Na figura também estão dispostos os histogramas das amostras das distribuições a posteriori. Como se pode verificar, o procedimento de estimação foi capaz de identificar bastante bem os valores teóricos dos parâmetros.

As tabelas 4.20 e 4.21 exibem, respectivamente, estatísticas descritivas das amostras das posterioris marginais de cada parâmetro e suas correlações a posteriori estimadas.

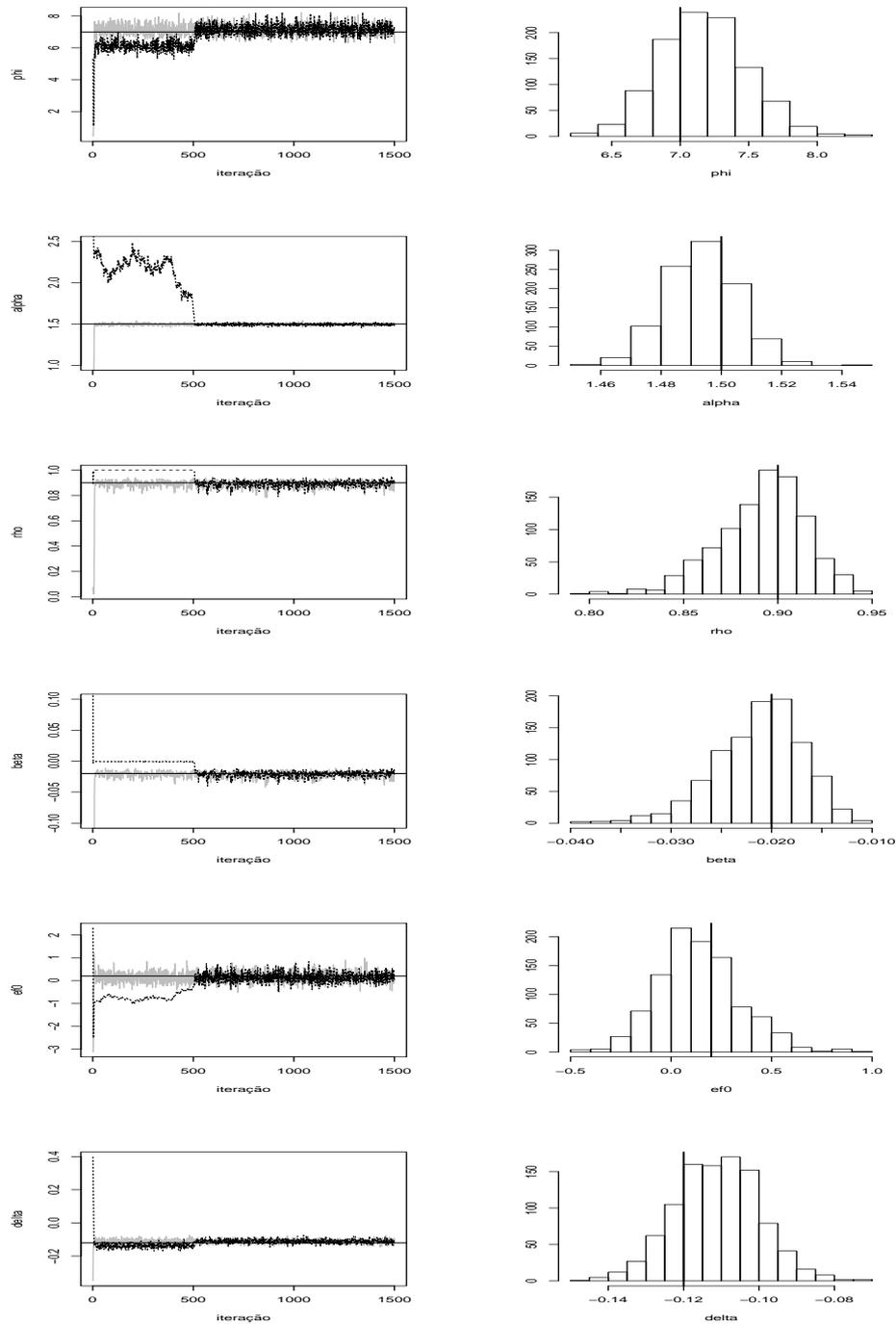


Figura 4.39 Resposta Gama: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).

Tabela 4.20 Resposta Gama: Sumário das amostras da posteriori obtidas da estimação do modelo autoregressivo de ordem ($r=1, s=0$) para dados simulados.

	φ	α	ρ	β	E_0	δ
Teórico	7,0	1,5	0,9	-0,02	0,2	-0,12
Mín	0,42940	1,45000	0,79960	-0,03871	-0,48100	-0,14740
Q1	6,95100	1,48600	0,87730	-0,02402	0,00504	-0,11870
Mediana	7,16300	1,49300	0,89500	-0,02078	0,11930	-0,11110
Média	7,14000	1,49300	0,89170	-0,02128	0,13700	-0,11110
Q3	7,37800	1,50100	0,90730	-0,01818	0,25460	-0,10350
Máx.	8,77000	1,54200	0,94570	-0,01064	0,97900	-0,07393
e.p.	0,06850	0,00016	0,00141	0,00010	0,01245	0,00012

Tabela 4.21 Resposta Gama: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem ($r=1, s=0$) para dados simulados.

	φ	α	ρ	β	E_0	δ
φ	1,00	0,01	0,01	0,01	0,00	0,00
α	0,01	1,00	-0,01	0,04	-0,15	-0,03
ρ	0,01	-0,01	1,00	0,95	-0,22	0,09
β	0,01	0,04	0,95	1,00	-0,24	0,03
E_0	0,00	-0,15	-0,22	-0,24	1,00	0,03
δ	0,00	-0,03	0,09	0,03	0,03	1,00

4.3.2 Função de transferência de ordem ($r = 1, s > 0$)

Suponha-se que o preditor η_t em (4.4) seja dado por (4.3). Assumindo-se $s = 1$, gerou-se um conjunto de dados composto por 1000 pontos, a partir de: $\varphi = 7,0$, $\alpha = 1,5$, $\rho = 0,9$, $\beta_0 = -0,03$, $\beta_1 = -0,01$, $E_0 = 0,2$, $\delta = -0,12$. Gráficos sumários das amostras geradas são exibidos na figura 4.40 e fica clara a dificuldade na identificação de β_0 e β_1 . A despeito disso, o impacto da regressora X_t sobre a resposta (excetuando-se seu impacto imediato) é estimado satisfatoriamente, como se vê na figura 4.41.

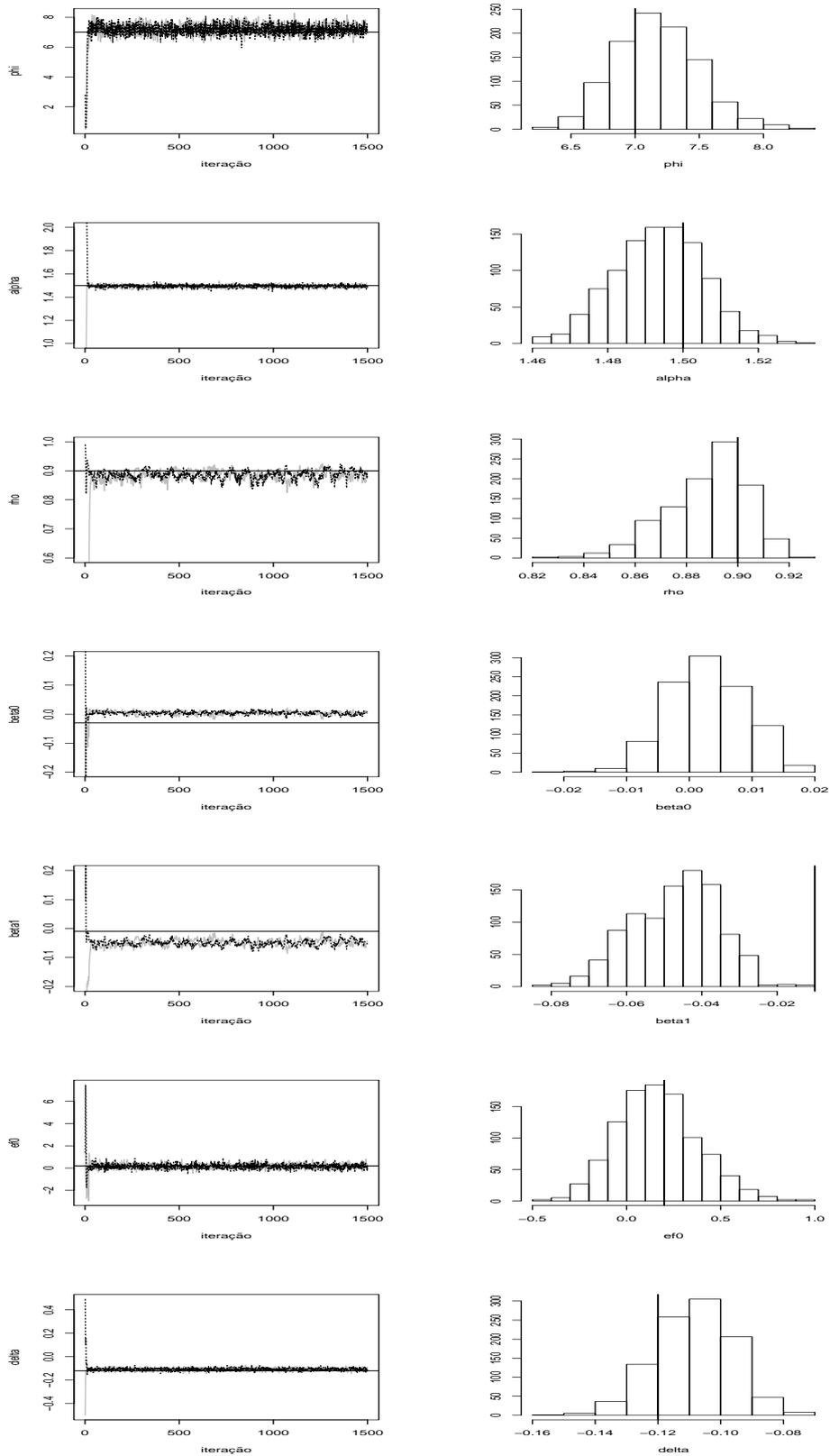


Figura 4.40 Resposta Gama: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=1$).

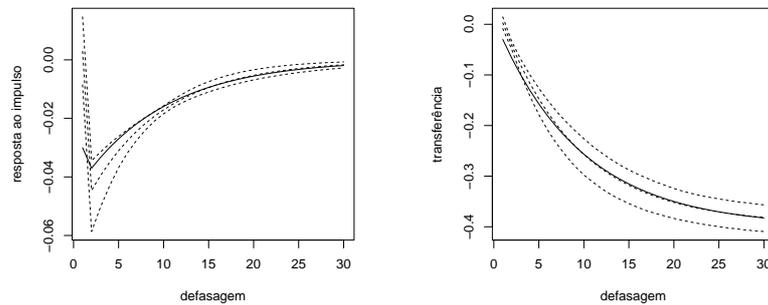


Figura 4.41 Resposta Gama, modelo de ordem $(r=1, s=1)$: Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).

Para $s = 2$ e trabalhando-se com dados gerados a partir de $\varphi = 7,0$, $\alpha = 1,5$, $\rho = 0,9$, $\beta_0 = -0,04$, $\beta_1 = -0,03$, $\beta_2 = -0,01$, $E_0 = 0,2$, $\delta = -0,12$, tem-se os resultados exibidos na figura 4.42. Estatísticas sumárias associadas às amostras das distribuições a posteriori de cada parâmetro são registradas nas tabelas 4.22 e 4.23.

Tabela 4.22 Resposta Gama: Sumário das amostras da posteriori obtidas da estimação do modelo de ordem $r=1, s=2$.

	φ	α	ρ	β_0	β_1	β_2	E_0	δ
Teórico	7,0	1,5	0,9	-0,04	-0,03	-0,01	0,2	-0,12
Mín.	6,2530	1,4580	0,8663	-0,0684	-0,1213	-0,0485	-0,5652	-0,1518
Q1	6,9560	1,4850	0,8915	-0,0196	-0,0835	-0,0130	0,0154	-0,1164
Mediana	7,1700	1,4940	0,8968	-0,0069	-0,0692	-0,0059	0,1363	-0,1083
Média	7,1700	1,4940	0,8966	-0,0082	-0,0688	-0,0054	0,1461	-0,1085
Q3	7,3780	1,5020	0,9023	0,0032	-0,0554	0,0031	0,2710	-0,1003
Máx.	8,2370	1,5300	0,9134	0,0380	0,0107	0,0224	0,9739	-0,0712
e.p.	0,0068	0,0002	0,0007	0,0008	0,0013	0,0011	0,0050	0,0004

Embora os intervalos de credibilidade estimados a 95% para β_0 , β_1 e β_2 - dados, respectivamente, por: $(-0,0397; 0,0221)$, $(-0,1044; -0,0303)$ e $(-0,0302; 0,0174)$ - contenham os valores teóricos de cada um destes parâmetros, não se obtêm estimativas acuradas para os mesmos, estando os valores teóricos para β_0 e β_1 (estes altamente correlacionados), localizados próximos às caldas de suas distribuições marginais a posteriori. Tem-se, então, subestimativas para as funções de resposta ao impulso e transferência, como se vê na figura 4.43, ainda que a função de resposta média, exibida na figura 4.44, tenha sido muito bem estimada, o que nos traz indícios de problemas de identicabilidade para esta especificação do modelo.

Tabela 4.23 Resposta Gama: Correlações a posteriori obtidas da estimação do modelo autoregressivo de ordem $r=1, s=2$.

	φ	α	ρ	β_0	β_1	β_2	E_0	δ
φ	1,00	-0,04	0,01	0,01	0,00	-0,01	-0,02	0,01
α	-0,04	1,00	0,01	0,01	0,00	0,01	-0,12	0,05
ρ	0,01	0,01	1,00	0,11	-0,35	0,90	-0,19	0,03
β_0	0,01	0,01	0,11	1,00	-0,94	0,22	-0,06	0,08
β_1	0,00	0,00	-0,35	-0,94	1,00	-0,52	0,10	-0,10
β_2	-0,01	0,01	0,90	0,22	-0,52	1,00	-0,20	0,03
E_0	-0,02	-0,12	-0,19	-0,06	0,10	-0,20	1,00	0,07
δ	0,01	0,05	0,03	0,08	-0,10	0,03	0,07	1,00

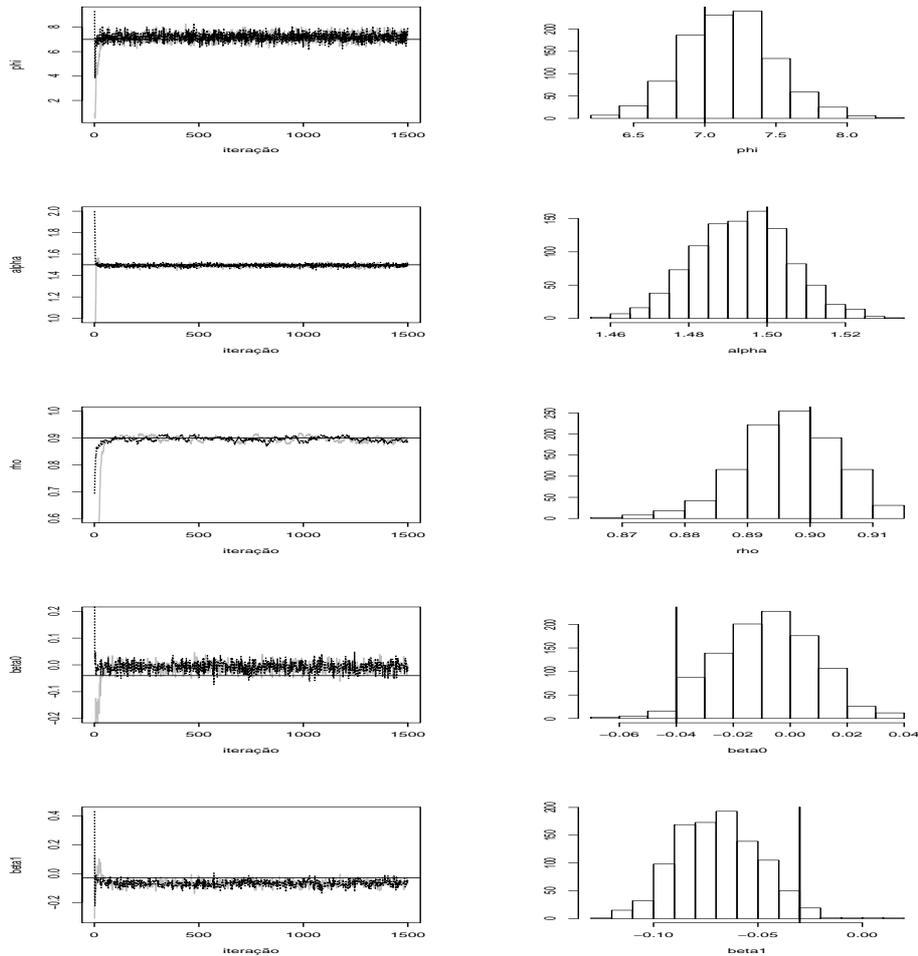


Figura 4.42 Resposta Gama: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=2$).

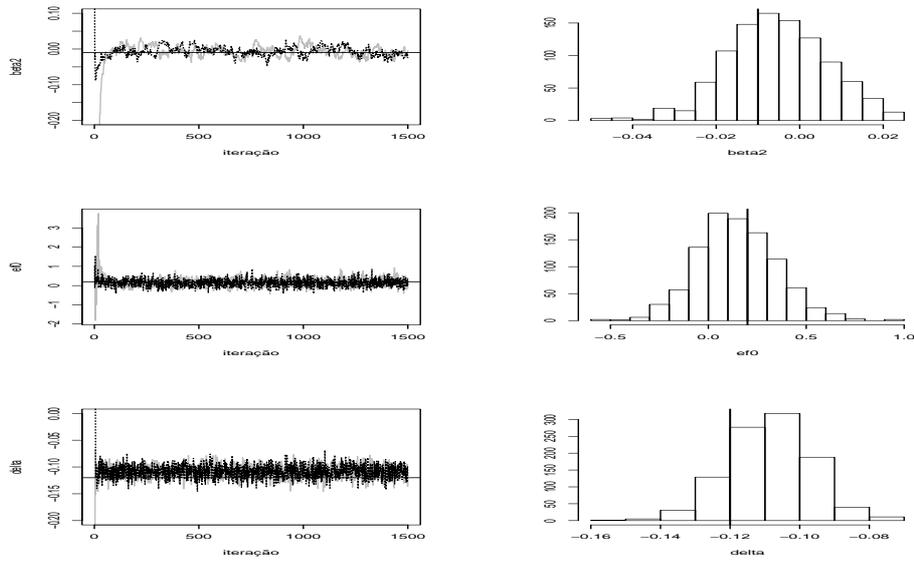


Figura 4.42 Resposta Gama: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=2$) (continuação).

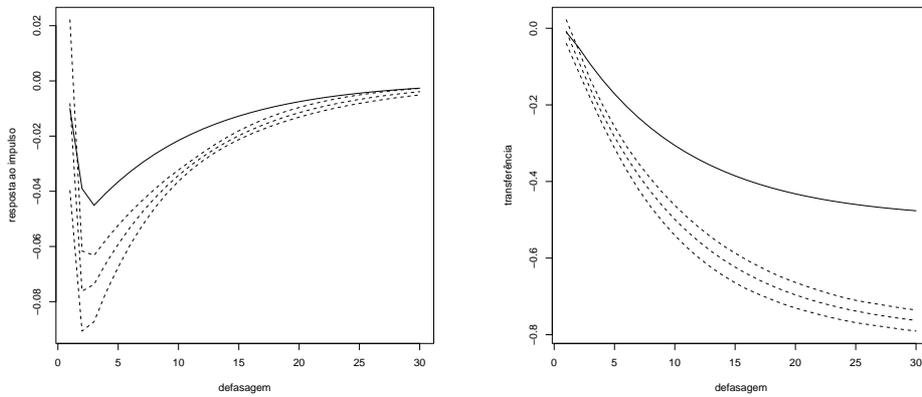


Figura 4.43 Resposta Gama, modelo de ordem ($r=1, s=2$): Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).

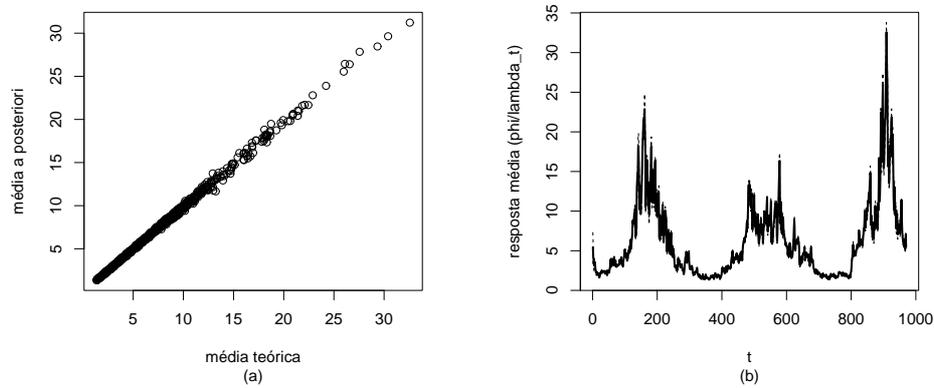


Figura 4.44 Resposta Gama, modelo de ordem ($r=1, s=2$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

4.4 Resposta Gaussiana Inversa

4.4.1 Função de transferência de ordem ($r = 1, s = 0$)

Considerando-se a formulação:

$$Y_t \sim \text{Gaussiana Inversa}(\mu_t, \sigma^2)$$

$$\log(\mu_t) = \eta_t = \alpha + E_t + \delta Z_t,$$

simulou-se, inicialmente, um conjunto de dados assumindo-se

$$E_t = \rho E_{t-1} + \beta X_t,$$

com $\alpha = 1,0$, $\rho = 0,9$, $\beta = 0,05$, $\varepsilon_0 = 0,5$, $\delta = -0,1$ e $\sigma^2 = 0,05$.

Para estimação do modelo foi executado o algoritmo MCMC descrito na seção 3.4, com movimentos individuais baseados em propostas passeio aleatório para $\log(\rho)$, α , β , ε_0 e δ , sendo σ^2 gerado a partir de sua densidade condicional completa, gama inversa. A figura 4.45 exhibe traços das cadeias geradas e histogramas das amostras da posteriori.

As funções estimadas de resposta ao impulso e de transferência obtidas a partir do procedimento MCMC são exibidas na figura 4.46. Na figura 4.50, observa-se a função de resposta média. Nota-se que houve uma clara dificuldade em identificar o valor do parâmetro σ^2 . Entretanto, como todos os parâmetros que compõem o preditor η_t foram identificados pelo procedimento de estimação e, ainda, como a estrutura preditiva é colocada na média μ_t , esta é muito bem estimada.

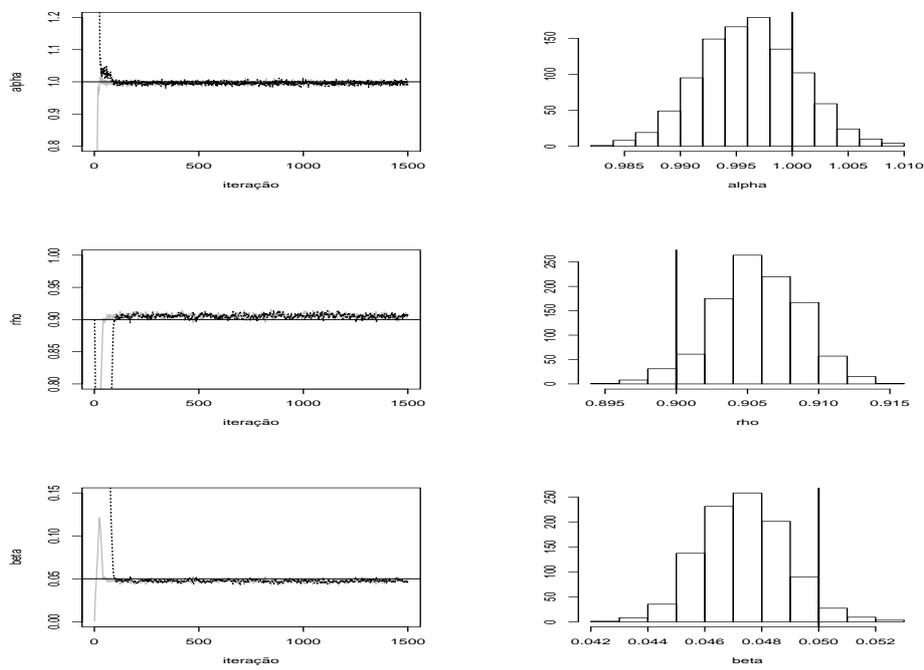


Figura 4.45 Resposta Gaussiana Inversa: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).

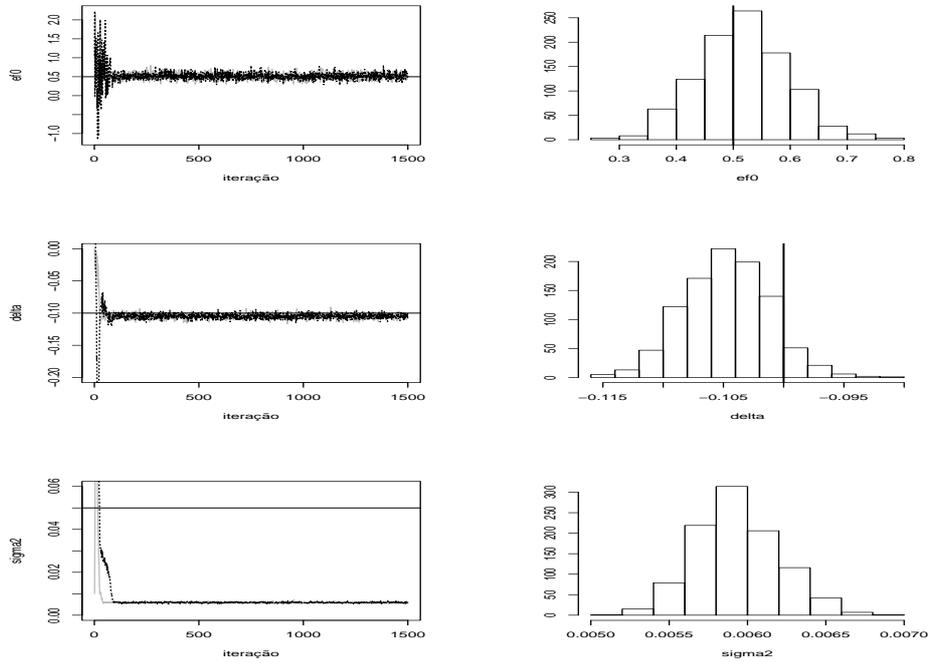


Figura 4.45 Resposta Gaussiana Inversa: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=0$).

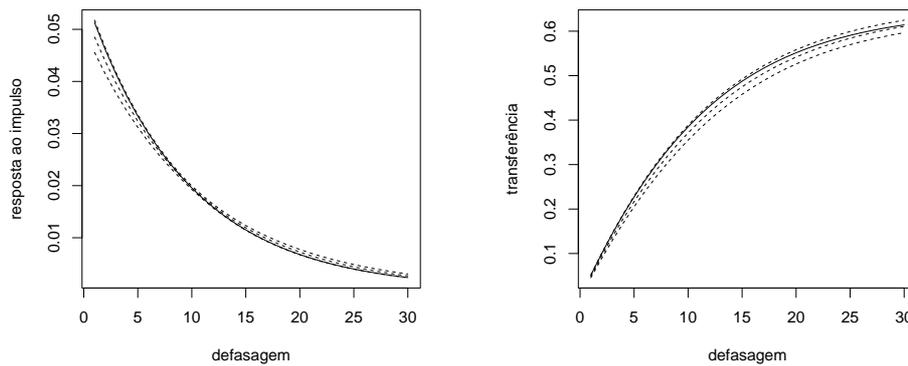


Figura 4.46 Resposta Gaussiana Inversa, modelo de ordem ($r=1, s=0$) : Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).

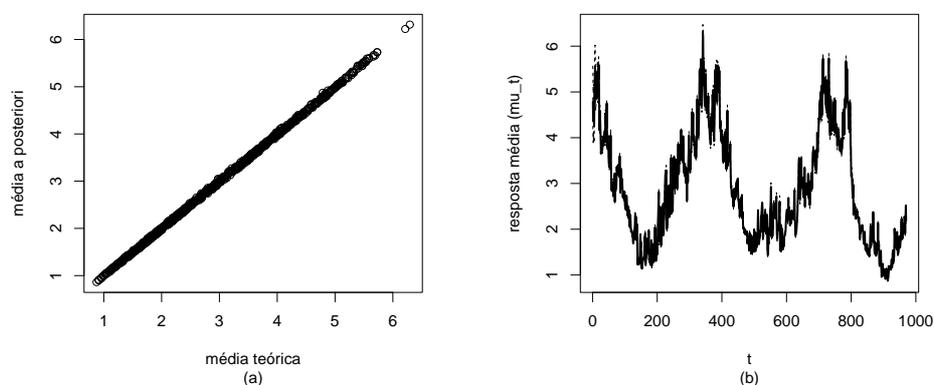


Figura 4.47 Resposta Gaussiana Inversa, modelo de ordem ($r=1, s=0$). Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

4.4.2 Função de transferência de ordem ($r = 1, s = 1$)

Em um segundo exercício, supõe-se

$$E_t = \rho E_{t-1} + \beta_0 X_t + \beta_1 X_{t-1},$$

com $\alpha = 1,0$, $\rho = 0,9$, $\beta_0 = 0,05$, $\beta_1 = 0,01$, $\varepsilon_0 = 0,5$, $\delta = -0,1$ e $\sigma^2 = 0,05$. O resumo do procedimento MCMC encontra-se na figura 4.48 e as figuras 4.49 e 4.50 exibem, respectivamente, as estimativas dos impactos defasados de X_t e da função de resposta média, μ_t .

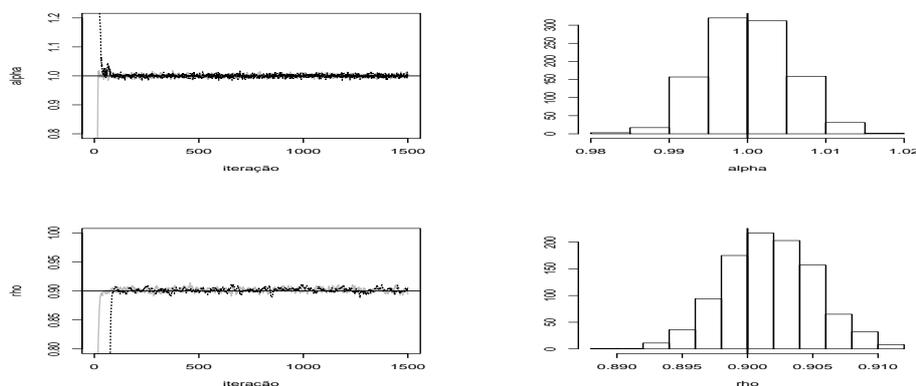


Figura 4.48 Resposta Gaussiana Inversa: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=1$).

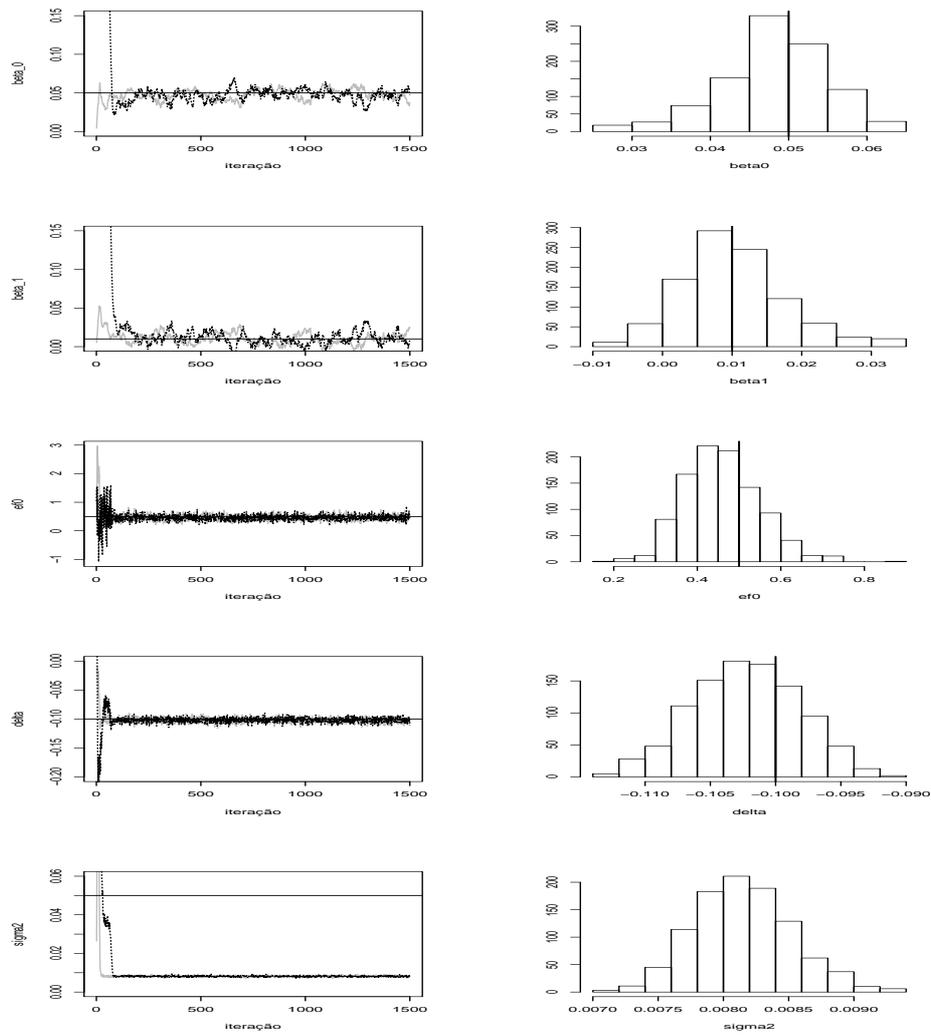


Figura 4.48 Resposta Gaussiana Inversa: Traços das cadeias geradas e histogramas da posteriori, para dados simulados de acordo com o modelo autoregressivo de ordem ($r=1, s=1$) - (continuação).

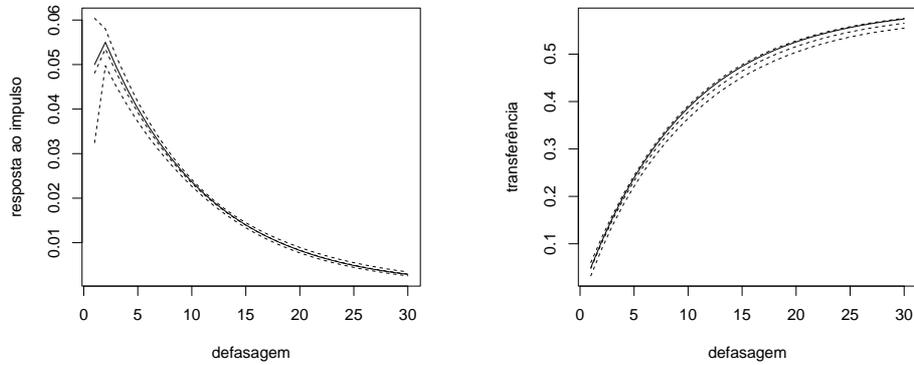


Figura 4.49 Resposta Gaussiana Inversa, modelo de ordem $(r=1, s=1)$: Função de resposta ao impulso e de transferência - real (linha sólida), média e limites de credibilidade estimados a 95% (linhas tracejadas).

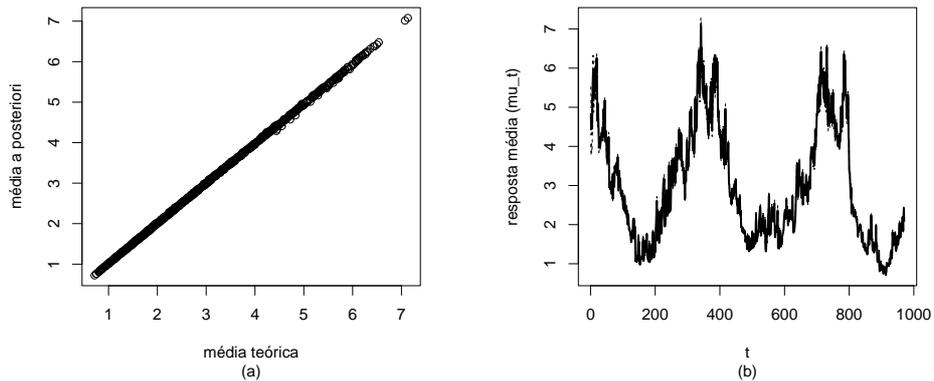


Figura 4.50 Resposta Gaussiana Inversa, modelo de ordem $(r=1, s=1)$. Função de resposta média: (a) diagrama de dispersão dos valores teóricos versus média a posteriori estimada; (b) traços da função teórica (linha em negrito), estimativas da média a posteriori e de limites de credibilidade a 95 %.

Mais uma vez, percebe-se que os parâmetros $\alpha, \rho, \beta, \varepsilon_0$ e δ são bem estimados. Toda a estrutura preditiva, dependente de tais parâmetros, é posta sobre a resposta média, μ_t . O parâmetro σ^2 , ortogonal a μ_t , é severamente subestimado nos dois casos tratados. Assim, tem-se uma redução na incerteza associada às estimativas induzida pela má estimação de σ^2 . Este fato não compromete as estimativas pontuais das funções de resposta ao impulso, de transferência e de resposta média, mas deve-se ter cautela ao se inferir sobre tais modelos, pois a credibilidade associada aos resultados obtidos é questionável.

4.5 Alguns Comentários sobre os Exercícios com Dados Simulados

Com exceção do caso Binomial, em que houve problemas de identificação do par ρ, β_0 , a estimação de funções de transferência em modelos estáticos de ordem $r = 1, s = 0$ foi muito bem sucedida, com identificação de todos os parâmetros envolvidos e bastante precisão associada às estimativas; quadro que não se alterou ao se adicionar, ao preditor, termos de erro independentes e identicamente distribuídos.

Para os modelos de ordem $r = 1, s = 1$, as estimativas pontuais de β_0 e β_1 não são muito acuradas, o que não chega a comprometer a estimação das funções de resposta ao impulso, de transferência e de resposta média, como se constata ao se levar em conta intervalos de credibilidade estimados para tais funções.

Já os modelos de ordem $r = 1, s = 2$ apresentaram problemas de identificação acarretados pela autocorrelação serial na regressora, tendo as estimativas espúrias de β_0, β_1 e β_2 comprometido a identificação da função de transferência teórica. As dificuldades encontradas já com $r = 2$ defasagens aplicadas à regressora sugerem que, para a modelagem de efeitos que atinjam seu ápice algum tempo após um pulso, é recomendável que se invista em abordagens livres dos problemas de autocorrelação a que estão sujeitos os modelos que dependem de defasagens na regressora.

No caso da distribuição Gaussiana Inversa, embora tenham sido obtidos resultados acurados na estimação dos parâmetros presentes no preditor (e, conseqüentemente, das funções de resposta ao impulso, de transferência e de resposta média, μ_t), mesmo para modelos bastante simples foi impossível identificar o parâmetro σ^2 , uma vez que a estrutura preditiva especificada sobre μ_t não fornece informação sobre este, o que sugere a adequação de um nível hierárquico adicional, no qual σ^2 seja modelado.

Quanto à adoção de parâmetros com evolução estocástica, embora tenha-se encontrado certo grau de dificuldade na identificação do parâmetro autoregressivo ρ ao se especificar nível dinâmico associado à resposta Poisson, quando a mesma estrutura foi aplicada a dados binomiais o modelo mostrou-se identificável. Também foi bem sucedida a estimação de modelos com funções de transferência dinâmicas de ordem $(r = 1, s = 0)$, determinadas pela evolução do parâmetro β , que mede o impacto imediato da regressora sobre a resposta esperada. A dinâmica da função de resposta ao impulso $(r = 1, s = 0)$ é determinada pelo amortecimento de β , nos instantes posteriores a um pulso, causado por potências crescentes em ρ . Portanto, ao se permitir a evolução de β , confere-se dinâmica não somente os impactos imediatos da

regressora, mas também a seus efeitos defasados. Embora as variâncias dos erros de evolução tenham sido subestimadas, produzindo trajetórias suavizadas para os parâmetros de estado, ainda assim obteve-se ótimos resultados na identificação desses modelos. É particularmente digna de nota a possibilidade de estimação de modelos tão flexíveis, a partir de dados gerados de distribuições que distam severamente da usual distribuição gaussiana, como foi o caso nos exemplos discutidos, em que se buscou produzir dados para os quais aproximações normais fossem inadequadas.

No próximo capítulo, os modelos avaliados são aplicados a conjuntos de dados reais.

Aplicações a Dados Reais

Em diversas grandes metrópoles ao redor de todo o mundo, especialistas vêm tentando identificar quais são os poluentes mais danosos à saúde, quantificando seu impacto sobre diferentes patologias. Uma questão importante - embora ignorada em vários desses estudos - é a consideração do impacto cumulativo de poluentes sobre desfechos epidemiológicos e do tempo de persistência dos efeitos de cada poluente, bem como sua forma de evolução, ou seja, se o efeito de elevações de determinado poluente, por exemplo, decresce ao longo do tempo ou cresce, atingindo impacto máximo dias após a ocorrência de elevação, para então decair.

Na maioria dos trabalhos que chegam a considerar impactos defasados, o horizonte de influência do poluente é fixado, adotando-se, usualmente, modelos com defasagens distribuídas ou médias móveis de até, no máximo, três dias. Para a aplicação de modelos com defasagens por vários dias, tem-se problemas de estimação devido à forte correlação entre os níveis de poluentes em dias subseqüentes. Tais dificuldades podem ser contornadas adotando-se modelos de defasagens com restrições nos coeficientes, como os modelos polinomiais apresentados na subseção 3.3.1 ou, ainda mais parcimoniosamente, utilizando-se funções de transferência com termos autoregressivos. A primeira aplicação, apresentada na seção 5.1, versa sobre efeitos (presentes e acumulados) de poluentes sobre contagens de óbitos de crianças com idades inferiores a 5 anos, devido a doenças respiratórias, na região metropolitana de São Paulo.

Outra questão importante na avaliação do impacto da poluição atmosférica é a construção de modelos para previsão dos próprios níveis de poluentes, uma vez que é bastante comum a existência de períodos sem observação neste tipo de série temporal, devido a falhas nas estações de mensuração ou mesmo no registro de tais mensurações. A previsão do nível de poluentes é uma tarefa bastante complexa, uma vez que a propagação dos mesmos depende - além de variáveis climáticas como temperatura e umidade - de correntes de ar, efeitos topográficos, bem como de suas próprias características físico-químicas. Na aplicação desenvolvida na seção 5.2, utiliza-se um banco de dados referente à região metropolitana do Rio de Janeiro, que contém registros de níveis de poluentes, chuva, temperatura e umidade, modelando-se a evolução do impacto da chuva sobre níveis de material particulado.

Na seção 5.3, faz-se um breve resumo dos resultados obtidos nas aplicações.

5.1 Efeito de Monóxido de Carbono sobre Óbitos de Crianças em São Paulo

Nesta seção é apresentada uma aplicação envolvendo contagens diárias de óbitos por doenças respiratórias de crianças e níveis de poluentes atmosféricos, registrados em São Paulo, no período de 01 de janeiro de 1994 a 31 de dezembro de 1997, aos quais serão aplicados os modelos Poisson propostos no capítulo 3.

Há algum tempo tem-se percebido que o número de internações por doenças respiratórias em São Paulo parece estar correlacionado às condições atmosféricas e níveis de poluição na região metropolitana. Essa hipótese vem servindo como motivação para o desenvolvimento de diversos estudos que buscam quantificar o efeito de diferentes poluentes sobre o número de internações e/ou óbitos. O banco de dados utilizado apresenta apenas registros de óbitos (e não de internações). Obviamente as causas que levam ao óbito são várias, entretanto, utilizando a informação disponível, restringimo-nos ao efeito de poluentes sobre a contagem de desfechos epidemiológicos, utilizando, como covariáveis: temperatura, umidade e termos para ajuste de sazonalidade. Em análises preliminares (Alves, 2003; Ornelas, 2004), verificamos significância (e persistência) do impacto de Monóxido de Carbono (CO) sobre óbitos de crianças menores de 5 anos em São Paulo. Passamos a descrever os resultados das análises dos efeitos deste poluente.

De acordo com Estrella et al. (2005), o Monóxido de Carbono resulta da queima incompleta de combustíveis e a principal fonte emissora desse poluente é veicular, com altas concentrações ocorrendo em áreas sujeitas a tráfego intenso e congestionamentos. O Monóxido de Carbono compete com o Oxigênio, tendo uma capacidade de combinação com a hemoglobina muito superior a este último. Forma-se, a partir dessa combinação, a carboxihemoglobina, reduzindo o transporte de oxigênio pelo sangue e debilitando a função respiratória.

Os padrões de qualidade do ar definem legalmente o limite máximo para a concentração de um poluente na atmosfera, que garanta a proteção da saúde e o bem estar da população. Os padrões nacionais foram estabelecidos pelo IBAMA - Instituto Brasileiro de Meio Ambiente e aprovados pelo CONAMA - Conselho Nacional de Meio Ambiente. Na figura 5.1, encontra-se a série temporal de registros do poluente Monóxido de Carbono, no período avaliado. Toma-se, como valor diário do poluente, a média aritmética sobre as oito estações de mensuração de CO disponíveis no banco de dados. A linha horizontal indica o limite superior da fronteira de níveis aceitáveis para composição de um índice de qualidade do ar, conforme registra a Companhia de Tecnologia de Saneamento Ambiental (CETESB). Observa-se que nos meses de inverno, quando se agrava o fenômeno da inversão térmica, dificultando a dispersão de poluentes, os níveis de CO ultrapassam o limite aceitável.

A figura 5.2 exibe a série de contagens diárias de óbitos por doenças respiratórias para crianças com idades inferiores a 5 anos, na cidade de São Paulo, no período de 1994 a 1997. Nota-se a existência de sazonalidade bem marcada, com aumento de casos durante o inverno. Além da natureza discreta do número de óbitos, pode-se observar que o número médio de contagens é bastante baixo, invalidando qualquer possibilidade de aproximação normal para tal desfecho.

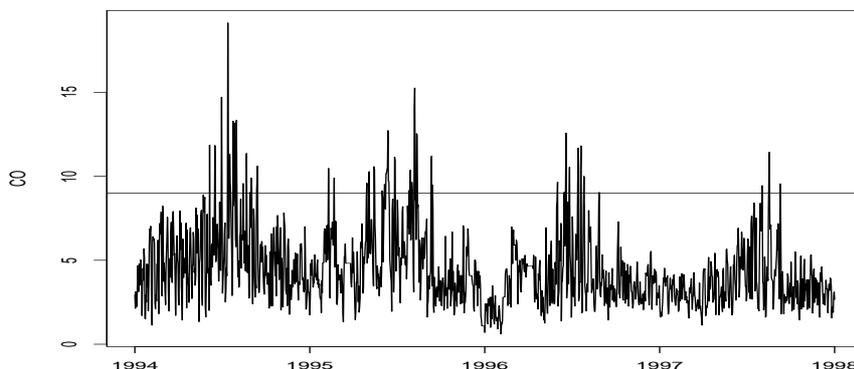


Figura 5.1 Níveis diários de CO (ppm) em SP - 01/Jan/1994 a 31/Dez/1997.

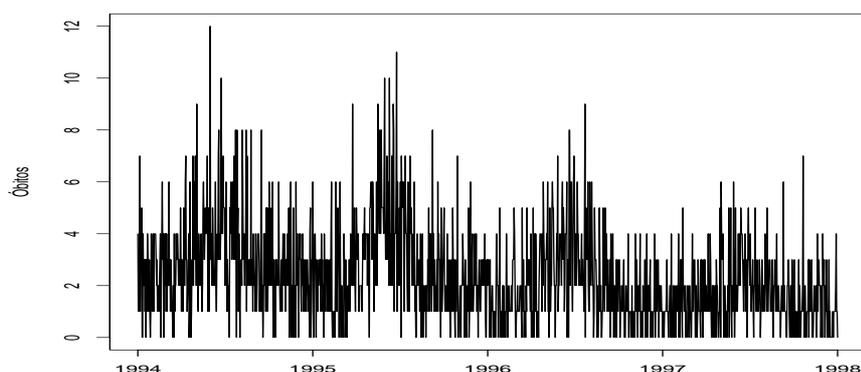


Figura 5.2 Contagem de óbitos por doenças respiratórias de menores de 5 anos em SP - 01/Jan/1994 a 31/Dez/1997.

Ao analisar estes mesmos dados, por meio de modelos aditivos generalizados, Conceição et al. (2001) constataram a existência de associação significativa entre mortalidade infantil por doenças respiratórias e níveis diários de CO, SO₂ e PM₁₀. Ao incluir os poluentes simultaneamente no modelo, observaram relação significativa apenas para o poluente CO e significância marginal do poluente SO₂. Botter et al. (2000) modelaram as contagens totais de óbitos de maiores de 65 anos, também em São Paulo, de 1991 a 1993, utilizando modelos de espaço de estados. Embora tenham encontrado maior efeito associado ao Dióxido de Enxofre (SO₂), restringem o horizonte de influência dos poluentes a apenas alguns dias. Ao utilizar, para este mesmo desfecho, no período de 1994 a 1997, modelos aditivos generalizados (com temperatura e umidade consideradas de forma não-paramétrica e defasagens distribuídas com restrição polinomial para os efeitos de poluentes), Alves (2005) observa que, embora o efeito imediato de

SO_2 seja maior que o de CO para os óbitos de idosos, o efeito de CO é mais duradouro, superando o de SO_2 para horizontes além de 4 dias. Ferreira e Gamerman (2000) avaliaram a associação entre poluentes e mortalidade infantil em São Paulo durante o ano de 1991, utilizando, como variáveis explicativas, os níveis de NO_2 e CO, relacionados às contagens de óbitos por meio de Modelos Dinâmicos Lineares Generalizados, como proposto no presente trabalho. O modelo adotado por Ferreira e Gamerman, entretanto, não considera os efeitos de médio e longo prazo das covariáveis, avaliando apenas o impacto instantâneo dos poluentes sobre o número de óbitos.

A forma geral dos modelos adotados na presente análise é:

$$\begin{aligned} \text{Óbitos}_t &\sim \text{Poisson}(\lambda_t) \\ \log(\lambda_t) = \eta_t &= \alpha + E_t + \delta_1 \text{Temp}_t + \delta_2 \text{Umid}_t \\ &\quad + \delta_3 \text{Cos}\left(\frac{2\pi t}{365}\right) + \delta_4 \text{Sen}\left(\frac{2\pi t}{365}\right) + \delta_5 \text{Cos}\left(\frac{4\pi t}{365}\right) + \delta_6 \text{Sen}\left(\frac{4\pi t}{365}\right) \end{aligned}$$

em que E_t é o bloco estrutural representando o efeito acumulado de CO a cada instante, expresso por diferentes formas propostas de função de transferência. Além do poluente, utilizam-se como covariáveis em todos os modelos ajustados a temperatura mínima e a umidade média, cujos valores diários são exibidos na figura 5.3.

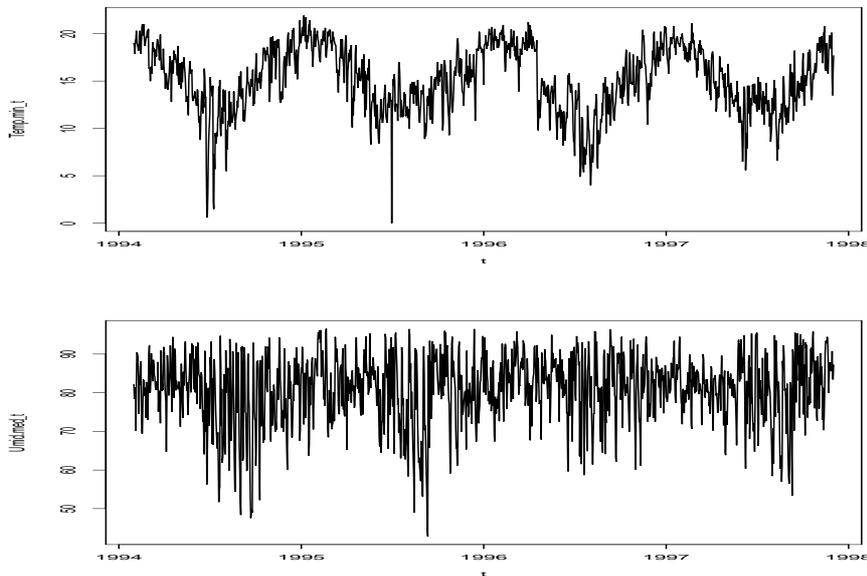


Figura 5.3 Registros diários de temperatura mínima e umidade média em SP - 01/Jan/1994 a 31/Dez/1997.

Na tabela 5.1, são detalhadas as especificações do preditor η_t que define cada um dos modelos estimados, via MCMC, como descrito no capítulo 3 e sob a adoção de prioris vagas para todos os parâmetros. Na mesma tabela, são apresentados os valores obtidos para dois critérios de comparação de modelos: DIC e EPD, ambos descritos na seção 2.5.

Tabela 5.1 Comparação dos modelos ajustados para óbitos de crianças, com avaliação do impacto cumulativo do poluente Monóxido de Carbono.

Modelo (Especificações do Preditor)	Função de Transferência	DIC	EPD
1: $(r = 1, s = 0), \delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \rho E_{t-1} + \beta CO_t$	$(p_D=4,1)$ 560,85	2825,98
2: $(r = 1, s = 1), \delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \rho E_{t-1} + \beta_0 CO_t + \beta_1 CO_{t-1}$	$(p_D=4,6)$ 561,76	2830,47
3: $(r = 0, s = 30, d = 2),$ $\delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \sum_{j=0}^{30} \beta_j CO_{t-j}$ $\beta_j = \sum_{k=0}^2 \zeta_k j^k$	$(p_D=3,5)$ 559,30	2828,69
4: $(r = 0, s = 30, d = 3),$ $\delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \sum_{j=0}^{30} \beta_j CO_{t-j}$ $\beta_j = \sum_{k=0}^3 \zeta_k j^k$	$(p_D=4,1)$ 560,20	2832,95
5: $(r = 1, s = 0),$ erros iid, $\delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \rho E_{t-1} + \beta CO_t + \varepsilon_t$ $\varepsilon_t \sim N(0, Q_\varepsilon)$	$(p_D=229,9)$ 508,09	2521,05
6: $(r = 1, s = 1),$ erros iid, $\delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \rho E_{t-1} + \beta_0 CO_t + \beta_1 CO_{t-1} + \varepsilon_t$ $\varepsilon_t \sim N(0, Q_\varepsilon)$	$(p_D=233,2)$ 508,68	2534,36
7: $(r = 0, s = 30, d = 2),$ erros iid, $\delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \sum_{j=0}^{30} \beta_j CO_{t-j} + \varepsilon_t$ $\varepsilon_t \sim N(0, Q_\varepsilon)$ $\beta_j = \sum_{k=0}^2 \zeta_k j^k$	$(p_D=233,0)$ 508,58	2521,58
8: $(r = 0, s = 30, d = 3),$ erros iid, $\delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$	$E_t = \sum_{j=0}^{30} \beta_j CO_{t-j} + \varepsilon_t$ $\varepsilon_t \sim N(0, Q_\varepsilon)$ $\beta_j = \sum_{k=0}^3 \zeta_k j^k$	$(p_D=229,8)$ 506,54	2523,66
9: $(r = 1, s = 0), \delta_3 = \delta_4 = \delta_5 = \delta_6 = 0$ choque dinâmico	$E_t = \rho E_{t-1} + \beta_t CO_t$ $\beta_t = \beta_{t-1} + v_t, v_t \sim N(0, Q)$	$(p_D=29,2)$ 425,77	2649,18
10: $(r = 1, s = 0)$ choque dinâmico e termos sazonais	$E_t = \rho E_{t-1} + \beta_t CO_t$ $\beta_t = \beta_{t-1} + v_t, v_t \sim N(0, Q)$	$(p_D=24,9)$ 441,24	2673,07

r =ordem da autoregressão, s =número de defasagens em CO, d =grau do polinômio de aproximação dos parâmetros de defasagem

O critério DIC privilegia os modelos em que se permite variação estocástica do parâmetro que governa o impacto imediato de CO sobre a resposta média, sendo o modelo 9 o escolhido. Já o critério EPD apresenta menores valores para os modelos cujos preditores contam com erros aleatórios e independentes. De acordo com tal critério, os modelos com defasagens sujeitas a restrições polinomias de graus 2 e 3 (modelos 7 e 8) são praticamente equivalentes ao modelo de ordem $(r = 1, s = 0)$. Como a parametrização deste último é mais parcimoniosa e não está sujeita a truncamento no horizonte de influência do poluente, escolheu-se, com base no critério EPD, o modelo 5. Passamos a descrever os resultados obtidos em cada um destes modelos.

Análise do Modelo 5

A figura 5.4 exibe histogramas das amostras das distribuições a posteriori de parâmetros envolvidos no modelo 5 e a tabela 5.2 exibe estatísticas associadas a tais amostras.

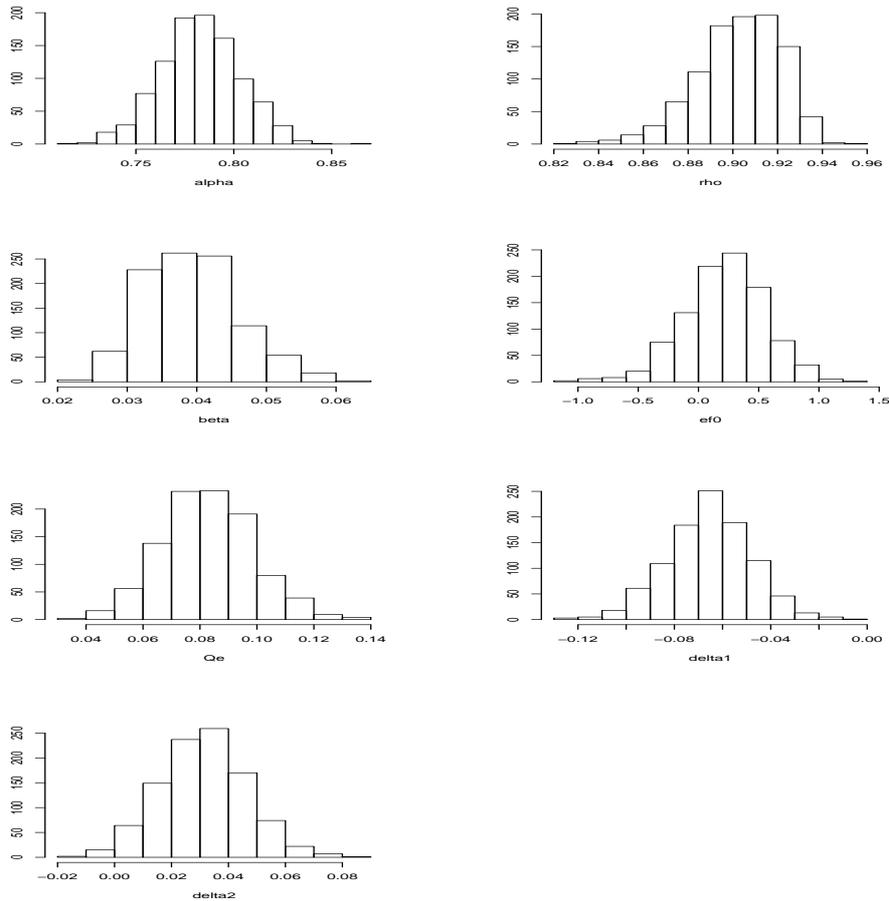


Figura 5.4 Efeito de CO sobre óbitos de crianças em SP - modelo 5: Histogramas das amostras da posteriori.

Tem-se elevada precisão associada ao processo de estimação, o que pode ser constatado ao se comparar a magnitude dos erros-padrão reportados à magnitude média de cada um dos parâmetros. As distribuições a posteriori, com exceção do parâmetro E_0 (cujas posteriori apresenta coeficiente de variação estimado - dado por $\sqrt{\hat{V}(E_0|D_T)}/\hat{E}(E_0|D_T)$ - igual a 120%), são bastante concentradas, com coeficientes de variação abaixo de 10%. Elevações no nível de CO têm efeito imediato significativo sobre a resposta esperada, como indicam o histograma e o sumário da posteriori de β e, ainda, de acordo com a estimação do parâmetro autoregressivo ρ , tal efeito é duradouro. A temperatura mínima tem efeito negativo sobre o número esperado de óbitos, ou seja, maiores registros de internações estão associados a temperaturas mais baixas.

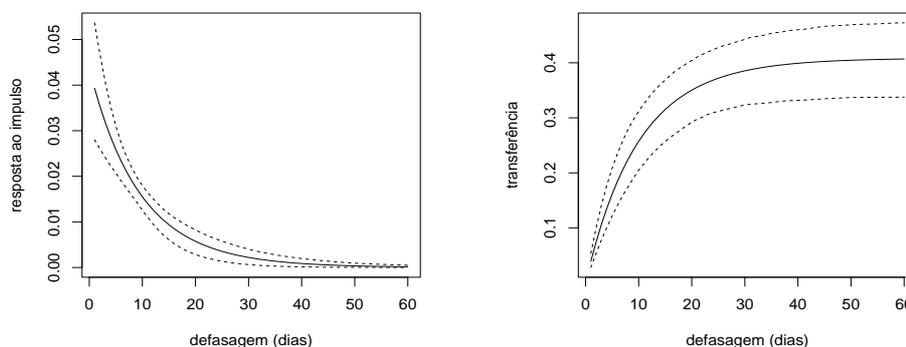
Tabela 5.2 Efeito de CO sobre óbitos de crianças em SP - modelo 5: Sumário das amostras da posteriori.

	α	ρ	β	E_0	Q_ε	δ_1	δ_2
Mín.	0,71590	0,82930	0,0218	-1,09200	0,03689	-0,12660	-0,01528
Q1	0,76970	0,89150	0,0342	0,00836	0,07191	-0,07652	0,02083
Mediana	0,78300	0,90490	0,0390	0,22350	0,08238	-0,06517	0,03157
Média	0,78300	0,90300	0,0393	0,21440	0,08273	-0,06556	0,03124
Q3	0,79610	0,91690	0,0435	0,44620	0,09347	-0,05368	0,04074
Máx.	0,86160	0,95180	0,0636	1,22600	0,13480	-0,00717	0,08063
e.p.	0,00055	0,00093	0,0003	0,00637	0,00046	0,00020	0,00019

Pode-se considerar o efeito de umidade positivo, com intervalo de credibilidade a 95% para δ_2 estimado por $(0,002; 0,061)$ e $\hat{P}(\delta_2 > 0 | D_T) = 0,983$.

Os gráficos das funções de resposta ao impulso e de transferência estimadas com base nas amostras das distribuições a posteriori de ρ e β encontram-se na figura 5.5. De acordo com tais estimativas, o efeito de um pulso de um desvio-padrão no nível de CO extingue-se após aproximadamente 30 dias.

Exponenciando-se a função de resposta ao impulso obtém-se a evolução dos riscos relativos associados aos níveis máximos de CO registrados em cada ano do período de análise, comparados ao nível médio ao longo de todo o período. como exhibe a figura 5.6. O nível máximo observado para o Monóxido de Carbono no período de análise foi 19,14 ppm, registrado no dia 14/07/1994. Segundo o modelo 5, uma elevação no valor de CO do nível médio observado ao longo de todo o período de análise (em torno de 4,44 ppm) para o nível máximo observado, provocaria um acréscimo inicial em torno de 20% no número esperado de óbitos, com tal efeito decrescente nos dias subsequentes. A magnitude do efeito, junto ao longo período de persistência observado para o poluente servem como alerta para a necessidade de implementação e manutenção de políticas públicas eficazes no combate à emissão de CO.

**Figura 5.5** Efeito de CO sobre óbitos de crianças em SP - modelo 5: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível do poluente.

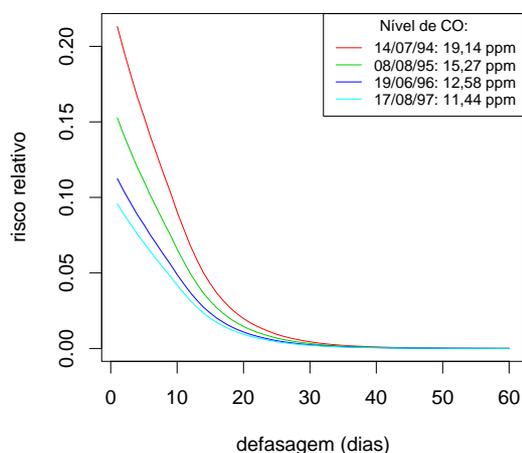


Figura 5.6 Efeito de CO sobre óbitos de crianças em SP - modelo 5: Evolução dos riscos relativos associados aos níveis máximos de CO observados a cada ano, quando comparados ao nível médio observado no período de análise.

A figura 5.7 apresenta a trajetória da função de resposta média, λ_t (estimada pela média a posteriori, com intervalos de credibilidade a 90%), junto aos valores observados para a série de contagens de óbitos. Pode-se ver que o modelo foi capaz de capturar bastante bem as flutuações da variável resposta.

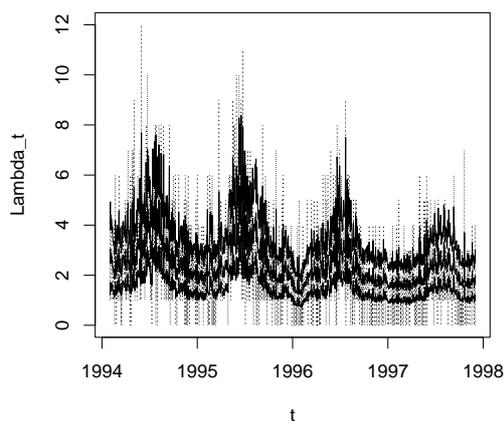


Figura 5.7 Efeito de CO sobre óbitos de crianças em SP - modelo 5: Contagens diárias de óbitos (linha tracejada) e função de resposta média, λ_t , estimada, com intervalos de credibilidade a 90%.

As 30 últimas observações disponíveis foram reservadas para avaliação da capacidade preditiva do modelo, tendo sido este ajustado utilizando as observações até o dia 30/11/1997. Na figura 5.8 encontram-se os histogramas das amostras das distribuições preditivas para o período de 01/12/1997 a 30/12/1997. As linhas verticais traçadas em cada histograma indicam o número de óbitos observado em cada um desses dias. Com exceção de alguns poucos dias, as distribuições preditivas concentram-se em torno dos valores observados. As distribuições preditivas são assimétricas e previsões baseadas na moda fornecem menor erro quadrático médio (1,43) que aquelas baseadas na mediana e na média, que apresentam erros quadráticos médios 1,63 e 2,05, respectivamente.

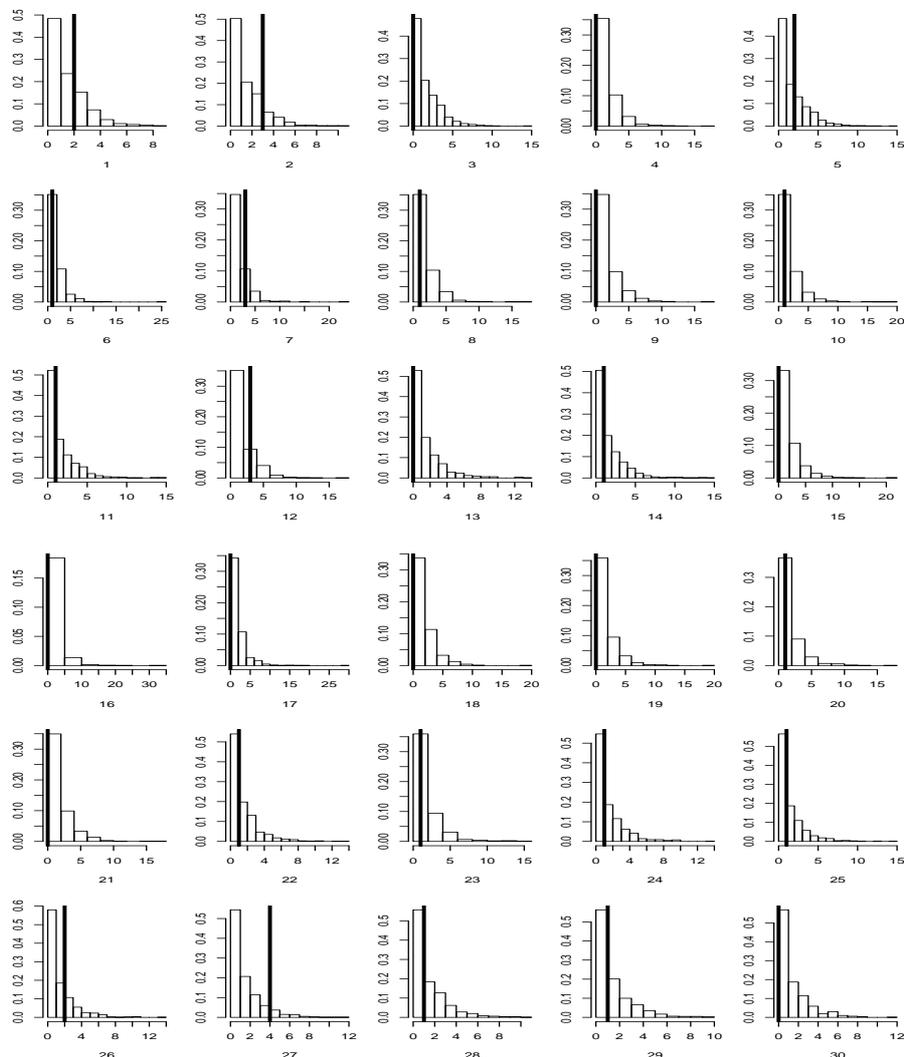


Figura 5.8 Efeito de CO sobre óbitos de crianças em SP - modelo 5: Histogramas das amostras das distribuições preditivas para o número de óbitos, com horizontes variando de 1 a 30 dias.

Modelo 9

Após convergência do procedimento MCMC, obteve-se a trajetória estimada para β_t , como exibe a figura 5.9.

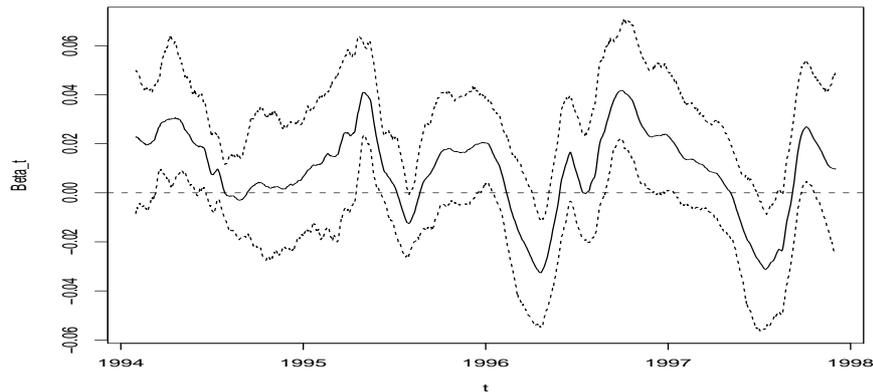


Figura 5.9 Efeito de CO sobre óbitos de crianças em SP - modelo 9: Média e intervalos de credibilidade a 90% para β_t .

Ainda que o critério DIC tenha apontado para o modelo 9, a estimativa dos coeficientes β_t implica, em alguns momentos, efeito negativo, o que é contra-intuitivo. Tais estimativas parecem influenciadas por algum resquício de sazonalidade não tratada neste modelo, o que nos leva à análise do modelo 10, que conta com termos sazonais e também apresenta valor do critério DIC inferior aos demais.

Análise do Modelo 10

Na figura 5.10 são exibidos histogramas das amostras das distribuições a posteriori dos parâmetros estáticos envolvidos no modelo 10. As amostras finais foram constituídas por 2000 pontos, obtidos mesclando-se 1000 iterações finais de duas cadeias paralelas, tomadas a intervalos de 10 iterações. Observa-se valor elevado para o parâmetro autoregressivo, indicando lento decaimento do efeito de Monóxido de Carbono sobre o número esperado de óbitos. Temperatura mínima apresenta tendência a efeito negativo, como esperado, com $\hat{P}(\delta_1 < 0 | D_T) = 0,94$. Embora o histograma da amostra de δ_2 indique não significância da variável umidade média, cujo coeficiente tem intervalos de credibilidade a 95% estimado por $(-0,029; 0,025)$, a variável é mantida no modelo, uma vez que é reconhecida por epidemiologistas como um fator relevante para a explicação do desfecho em questão. Assim, é desejável que a incerteza refletida por esta regressora seja considerada.

Estatísticas obtidas das amostras da posteriori encontram-se na tabela 5.2. Ao se comparar a magnitude dos erros-padrão à média obtida para cada parâmetro, percebe-se que as maiores

incertezas são encontradas na estimação de Q , δ_2 e δ_3 . A variância de evolução, Q , não está presente na verossimilhança e, portanto, toda a informação disponível para sua estimação é obtida de sua distribuição a priori (especificada de forma vaga) e da atualização dos erros de evolução, sendo natural esperar-se alguma dificuldade em sua estimação. É possível tentar melhorar a precisão das estimativas destes parâmetros utilizando-se cadeias mais longas e tomando-se amostras maiores. De forma geral, entretanto, o processo de estimação foi bastante preciso, como se pode ver pelos erros-padrão reportados para os demais parâmetros. Como já mencionado, a estimativa do parâmetro autoregressivo ρ indica que um pulso no nível de CO

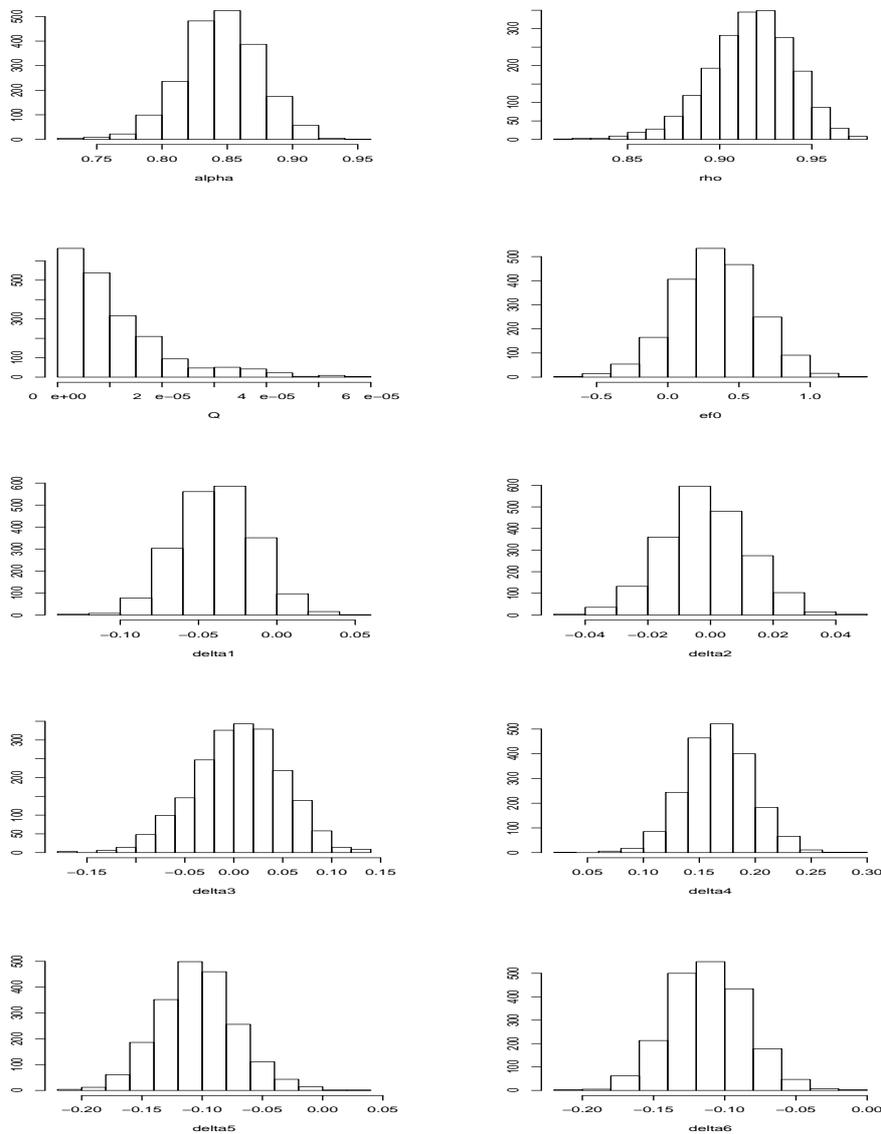
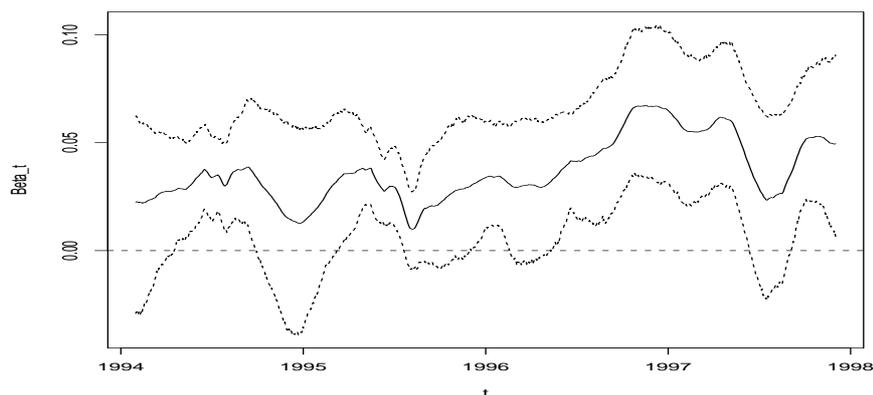


Figura 5.10 Efeito de CO sobre óbitos de crianças em SP - modelo 10: Histogramas das amostras da posteriori.

Tabela 5.2 Efeito de CO sobre óbitos de crianças em SP - modelo 10: Sumário das amostras da posteriori.

	α	ρ	Q	E_0	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6
Mín	0,73	0,81	1,36E-06	-0,68	-0,13	-0,049	-0,161	0,03	-0,21	-0,204
Q1	0,83	0,90	4,08E-06	0,14	-0,06	-0,011	-0,024	0,15	-0,13	-0,129
Mediana	0,85	0,92	7,90E-06	0,33	-0,04	-0,002	0,006	0,17	-0,11	-0,113
Média	0,85	0,92	1,10E-05	0,34	-0,04	-0,002	0,005	0,17	-0,10	-0,112
Q3	0,87	0,93	1,46E-05	0,53	-0,02	0,008	0,037	0,19	-0,08	-0,094
Máx.	0,95	0,98	5,95E-05	1,36	0,04	0,043	0,139	0,29	0,02	-0,003
e.p.	0,1%	0,1%	1,94E-06	0,5%	0,1%	0,03%	0,3%	0,1%	0,1%	0,04%

tem efeito duradouro sobre a resposta média. A variância dos erros de sistema apresenta-se bastante baixa, resultando em movimentos suaves para os coeficientes β_t cuja trajetória estimada é exibida na figura 5.11. Observa-se que o efeito de CO não se apresenta significativo em todo o período de análise.

**Figura 5.11** Efeito de CO sobre óbitos de crianças em SP - modelo 10: Média e intervalos de credibilidade a 90% para β_t .

Na figura 5.12 apresenta-se a trajetória estimada das funções de resposta ao impulso e de transferência associadas à elevação de um desvio-padrão no nível de CO, para quatro instantes, especificamente, aqueles em que foram observados os níveis máximos do poluente a cada ano. As estimativas pontuais são dadas por sua média a posteriori e registram-se também os limites de credibilidade estimados a 90%. Consta-se que tal pulso apresentaria-se significativo nos dias 14/07/94 e 19/06/96, com efeito anulando-se após aproximadamente 30 dias. Já nos dias 08/08/95 e 17/08/97, a variação sobre o nível de CO não tem impacto significativo sobre o número esperado de óbitos.

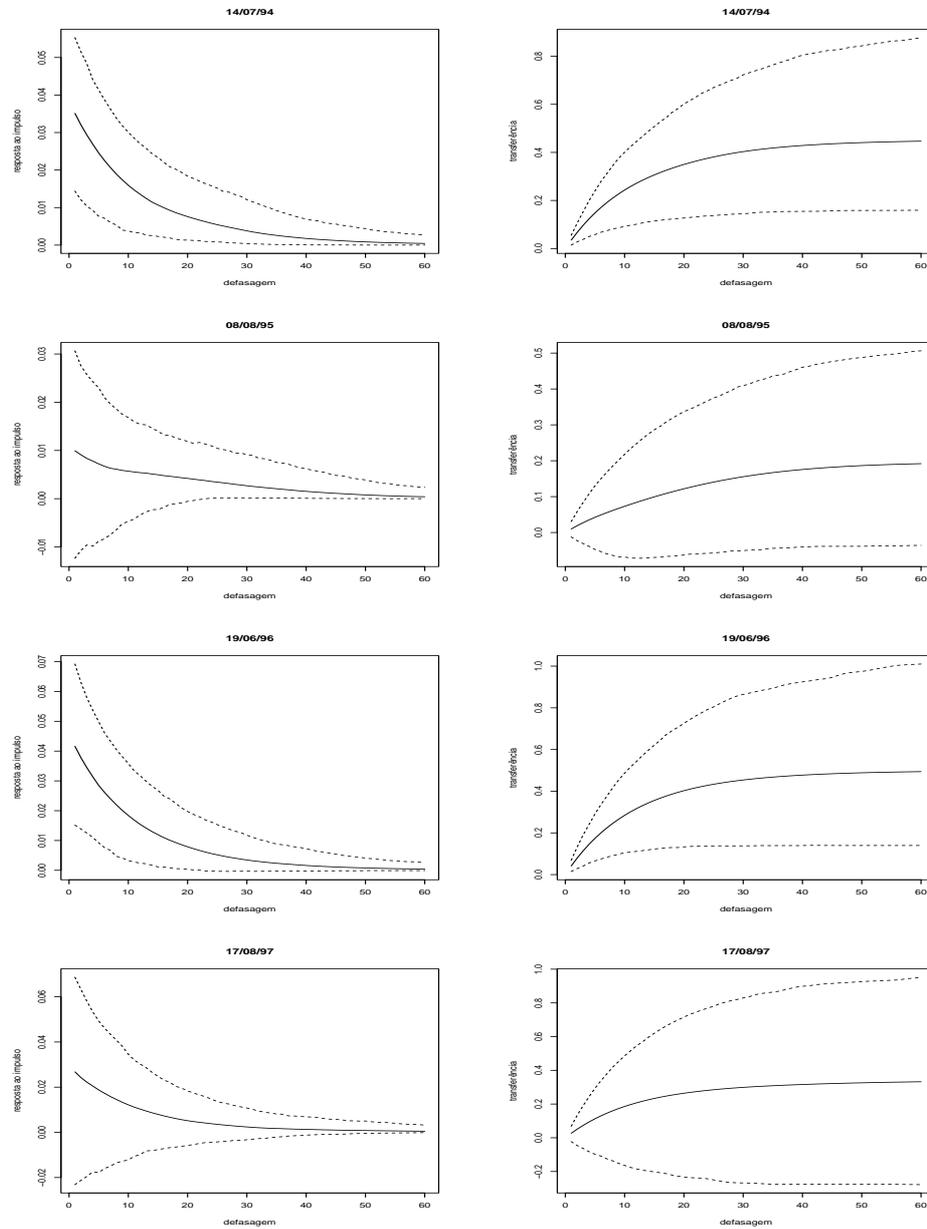


Figura 5.12 Efeito de CO sobre óbitos de crianças em SP - modelo 10: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível do poluente, estimadas em quatro instantes - estimativas da média a posteriori e de limites de credibilidade a 90%.

O nível máximo de CO registrado nos dois últimos anos de análise ficou em torno de 12 ppm. Na figura 5.13 apresenta-se a evolução temporal dos riscos relativos - obtidos exponenciando-se a função de resposta ao impulso - associados ao aumento no nível de CO do valor médio observado (4,44 ppm) para o 12,00 ppm, supondo-se a ocorrência de tal variação em diferentes instantes, nos quais o efeito de CO apresenta-se significativo, como indicado na figura 5.11. Como se pode ver, o mesmo pulso aplicado à regressora em diferentes momentos tem impactos com magnitudes bastante distintas, provocando um aumento inicial no número esperado de óbitos que varia de 2,57% a 10,64%. O efeito dessa elevação decai gradualmente, extinguindo-se após aproximadamente 30 dias.

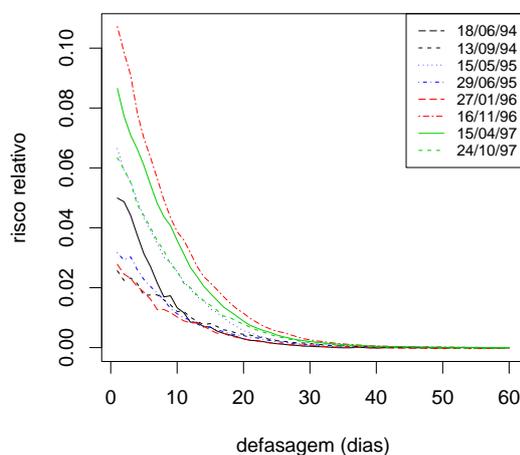


Figura 5.13 Efeito de CO sobre óbitos de crianças em SP - modelo 10: Evolução dos riscos relativos associados a 12 ppm de CO, comparado ao nível médio observado no período de análise, em diferentes instantes.

A figura 5.14 apresenta a trajetória da função de resposta média, λ_t (estimada pela média a posteriori, com intervalos de credibilidade a 90%), junto aos valores observados para a série de contagens de óbitos. Percebe-se que o modelo acompanha as flutuações da variável resposta, com incerteza menor que aquela registrada pelo modelo 5.

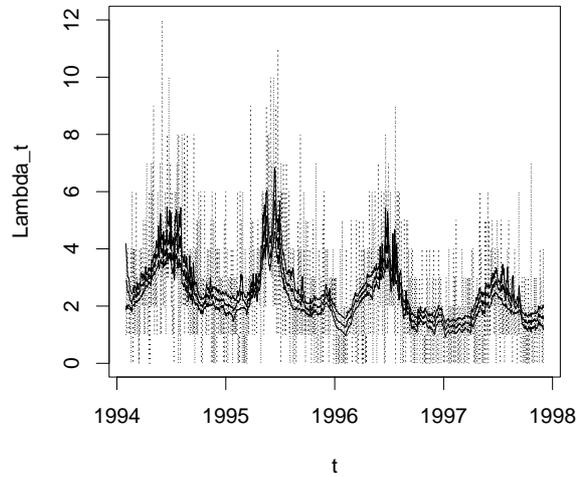


Figura 5.14 Efeito de CO sobre óbitos de crianças em SP - modelo 10: Contagens diárias de óbitos (linha tracejada) e função de resposta média, λ_t , estimada, com intervalos de credibilidade a 90%.

A figura 5.15 apresenta histogramas das distribuições preditivas para y_{T+1}, \dots, y_{T+30} , condicionais às regressoras. Mais uma vez, com exceção feita a alguns poucos pontos, as distribuições preditivas concentram-se em torno dos valores observados e os limites de credibilidade a 95% compreendem os valores observados em todo o horizonte de previsão. Média, mediana e moda estimadas a posteriori apresentam, respectivamente, erros quadráticos médios: 1,46, 1,27 e 1,13, sendo, portanto, a melhor estimativa pontual dada pela moda a posteriori.

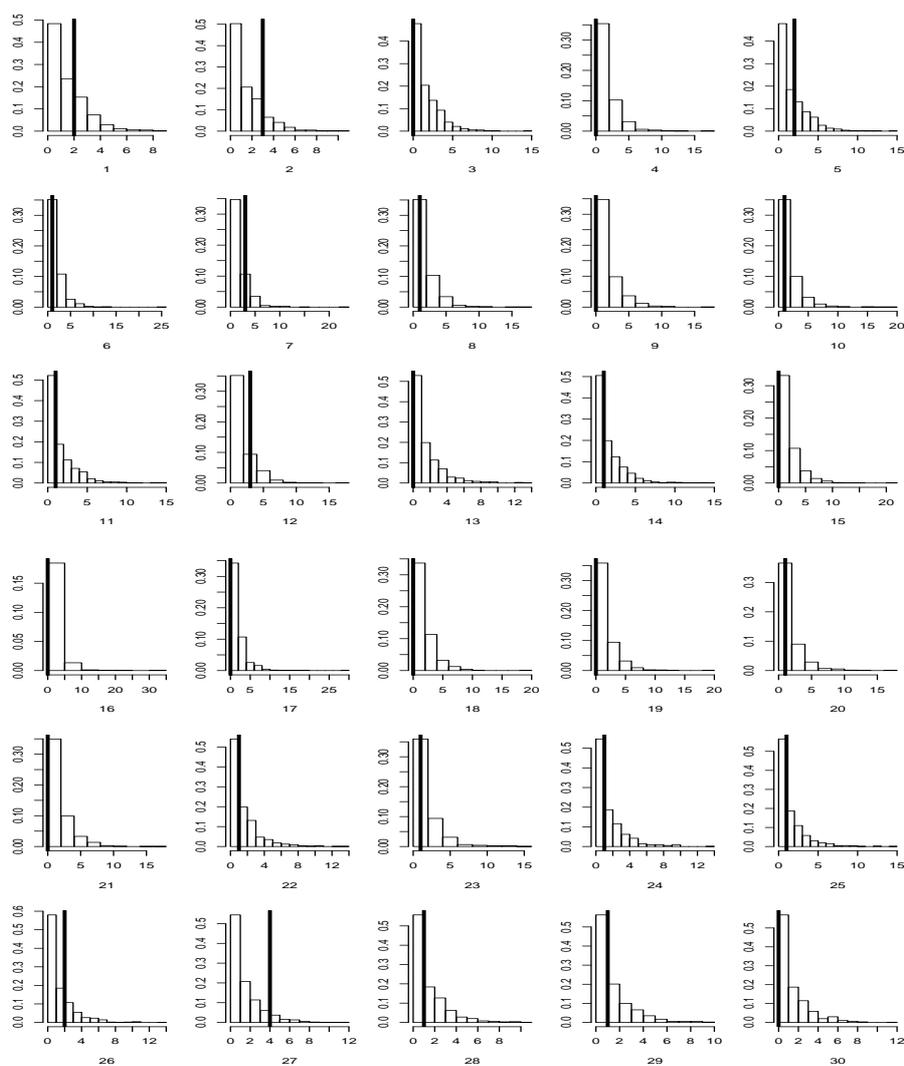


Figura 5.15 Efeito de CO sobre óbitos de crianças em SP - modelo 5: Histogramas das amostras das distribuições preditivas para o número de óbitos, com horizontes variando de 1 a 30.

Comparação entre os Modelos 5 e 10

Como visto, os critérios para seleção de modelos adotados não são conclusivos, apontando modelos diferentes. Existem diversos outros métodos para seleção de modelos e deve-se levar em conta que tais métodos dependem fortemente das funções de perda utilizadas. Assim, a escolha de outro método poderia levar a um outro modelo. Por exemplo, supondo-se que um modelo M tenha sido ajustado com base em (y_1, \dots, y_T) pode-se avaliar sua capacidade preditiva. Os valores obtidos para as estimativas das verossimilhanças preditivas (2.24) dos modelos 5 e 9 foram, respectivamente: $2,63 \times 10^{-19}$ e $1,73 \times 10^{-19}$, fornecendo uma razão de verossimilhanças preditivas igual a 1,5, favorável ao modelo 5.

De acordo com as estimativas pontuais para as previsões, entretanto, o modelo 10 seria o escolhido, como se verifica pelos erros quadráticos médios associados a média, mediana e moda reportados nas análises de cada um dos modelos.

A grande vantagem do modelo 5 em relação ao modelo 10 reside no tempo computacional necessário a seu ajuste, muito inferior a este último. Entretanto, com os avanços tecnológicos, os tempos computacionais tendem a diminuir cada vez mais, não sendo este, a médio prazo, um obstáculo a modelos dinâmicos. Ainda, julgamos a formulação do modelo 10 muito mais atraente e realista, uma vez que toda a incerteza é estruturada por meio da evolução temporal do parâmetro β_t , ao contrário do modelo 5, em que uma parcela considerável da incerteza associada à resposta é atribuída a termos aleatórios independentes, dando conta de efeitos não presentes no modelo. Concluimos, portanto, que é válido o investimento no sentido de melhorar o preditor do modelo 10, talvez incorporando efeito de outros poluentes ou considerando efeitos passados para as variáveis climáticas, não abandonando, assim, a flexibilidade de que diferentes situações atmosféricas observadas ao longo do tempo impliquem em impactos com diferentes magnitudes sobre a resposta esperada.

5.2 Efeito de Chuva sobre Níveis de Material Particulado no Rio de Janeiro

Segundo informações coletadas no sítio da Companhia de Tecnologia de Saneamento Ambiental (CETESB), o material particulado é um conjunto de poluentes constituídos de poeiras, fumaças e todo o tipo de material sólido e líquido que se mantém suspenso na atmosfera devido a seu pequeno tamanho (nesta aplicação, consideram-se partículas com diâmetro inferior a $10\mu\text{m}$). Esse poluente resulta da queima incompleta de combustíveis e de seus aditivos, de processos industriais e do desgaste de pneus e freios. Em geral provem da fumaça emitida pelos veículos movidos a óleo diesel, da fumaça expelida pelas chaminés das indústrias ou pelas queimadas, da poeira depositada nas ruas, dos resíduos de processos industriais que utilizam material granulado, de obras viárias ou que movimentam terra, areia, etc. O material particulado serve de meio de transporte para outras substâncias, como hidrocarbonetos e metais, que se agregam às partículas. Entre as partículas inaláveis, as mais grossas ficam retidas na parte superior do sistema respiratório, enquanto as mais finas, devido ao seu tamanho diminuto, podem atingir os alvéolos pulmonares. Entre os sintomas relacionados com a inalação do material particulado estão as alergias, asma e bronquite crônica. Causa também irritação nos olhos e garganta, reduzindo a resistência às infecções. Assim, são comuns os estudos epidemiológicos considerando níveis de material particulado como regressores. É bastante usual, entretanto, encontrar períodos sem registros, devido a defeitos nos equipamentos de mensuração ou outros motivos. Assim, modelos para previsão de níveis de poluentes podem ser muito úteis na formulação de modelos hierárquicos em que, num primeiro nível, determine-se o comportamento do poluente (em função de fatores meteorológicos, por exemplo) e, num segundo nível, o desfecho epidemiológico, dado o nível de poluente. Além disso, modelos que tenham níveis de poluentes como variável resposta são úteis na medida em que podem contribuir para a identificação de fontes controláveis de emissão, auxiliando na definição de políticas públicas para melhoria da qualidade do ar.

Nesta segunda aplicação, propõem-se modelos para avaliação da associação entre níveis do poluente material particulado (PM_{10}) no município do Rio de Janeiro e condições climáticas, incluindo-se temperatura mínima e umidade média, com especial atenção à propagação do efeito de volumes de chuva sobre os níveis desse poluente. Utilizam-se, ainda, como variáveis de controle, indicadoras de feriado e de dia da semana, tomando como base o domingo.

De acordo com o relatório técnico do projeto *Qualidade do Ar e Efeitos na Saúde da População do Município do Rio de Janeiro* (2005), os dados diários de poluição do ar para o período de setembro de 2000 a agosto de 2002 foram obtidos na Fundação Estadual de Engenharia do Meio Ambiente (FEEMA) - cujos monitores automáticos localizam-se em Jacarepaguá e no Centro - e na Secretaria Municipal do Meio Ambiente (SMAC), que tem monitores automáticos nos bairros de Copacabana, Centro, São Cristóvão e Tijuca. O nível de poluente utilizado nas aplicações é a média diária proveniente dessas diferentes estações de mensuração. Ainda segundo o mesmo relatório, as informações sobre variáveis meteorológicas foram coletadas no Departamento de Climatologia da UERJ, que forneceu dados da região do Maracanã, e no Departamento de Meteorologia da Aeronáutica, que disponibilizou as informações dos aeródros

mos de Santa Cruz, Campo dos Afonsos, Galeão e Santos Dumont. Mais uma vez utiliza-se a média das medidas diárias das variáveis meteorológicas. A figura 5.16 exhibe as concentrações médias diárias de PM_{10} ($\mu g/m^3$) registradas no período de análise e a figura 5.17, as séries diárias de volume de chuva, umidade média e temperatura mínima no período.

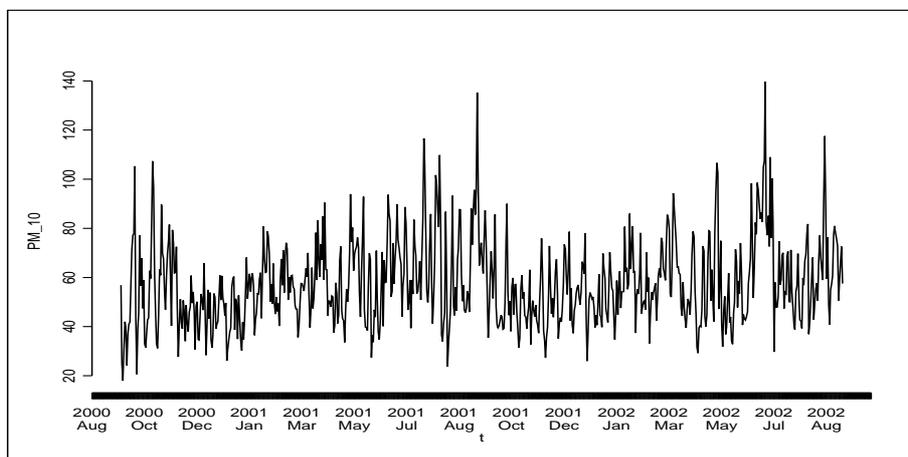


Figura 5.16 Níveis diários de PM_{10} ($\mu g/m^3$) no RJ - 01/Set/2000 a 31/Ago/2002.

Os modelos ajustados seguem a seguinte estrutura:

$$\text{Poluente}_t \sim \text{Gama}(\varphi, \lambda_t)$$

$$\log\left(\frac{\varphi}{\lambda_t}\right) = \eta_t = \alpha + E_t + \delta'_C \mathbf{C}_t + \delta'_D \mathbf{D}_t + \delta'_S \mathbf{S}_t,$$

em que:

- E_t é o bloco estrutural representando o efeito acumulado de chuva a cada instante, expresso por diferentes formas propostas de função de transferência;
- $\mathbf{C}'_t = (\text{Umidade}_t, \text{Temperatura}_t)$;
- $\mathbf{D}'_t = (\text{Seg}_t, \text{Ter}_t, \text{Qua}_t, \text{Qui}_t, \text{Sex}_t, \text{Sab}_t, \text{Fer}_t)$ é um vetor de variáveis indicadoras para dias da semana e feriados;
- $\mathbf{S}'_t = (\text{Cos}(\frac{2\pi t}{365}), \text{Sen}(\frac{2\pi t}{365}))$.

5.2 EFEITO DE CHUVA SOBRE NÍVEIS DE MATERIAL PARTICULADO NO RIO DE JANEIRO 128

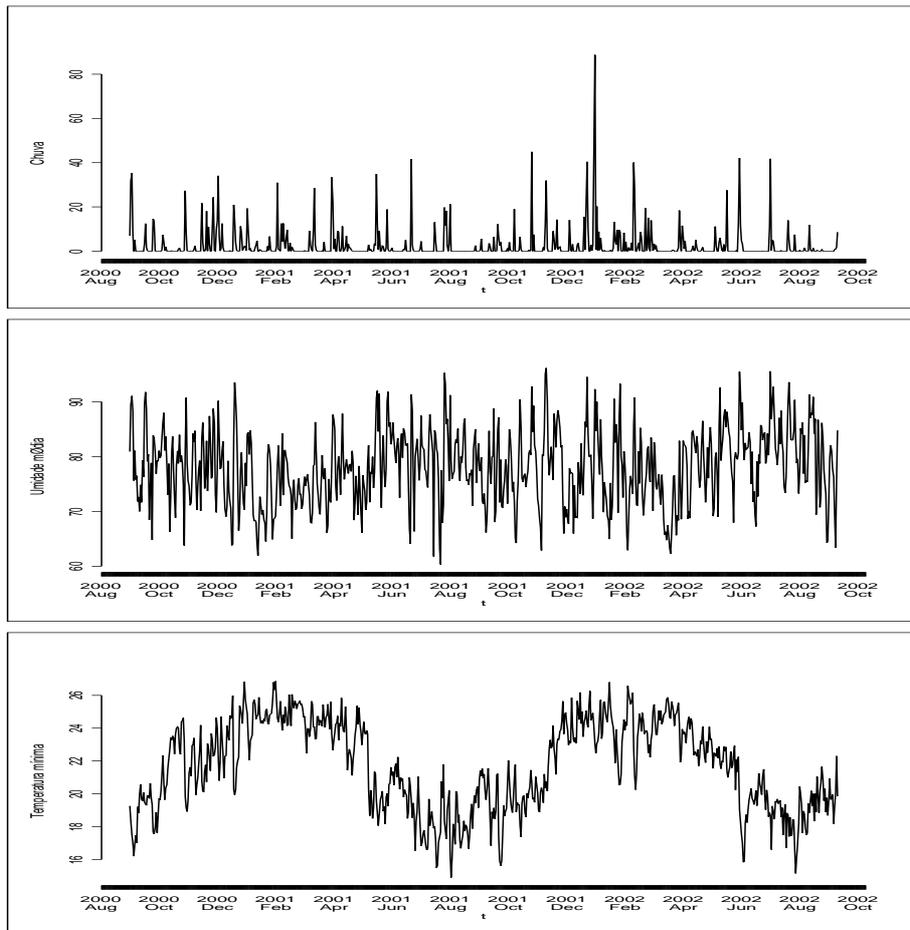


Figura 5.17 Variáveis meteorológicas no RJ - 01/Set/2000 a 31/Ago/2002.

A tabela 5.3 apresenta os valores obtidos para os critérios DIC e EPD, em função de diferentes especificações do preditor η_t .

Tabela 5.3 Comparação dos modelos ajustados para nível de PM_{10} , com avaliação do impacto cumulativo do volume de chuva.

Modelo (Especificações do Preditor)	Função de Transferência	DIC	EPD
1: $(r = 1, s = 0), \delta_S = \mathbf{0}$	$E_t = \rho E_{t-1} + \beta Chuva_t$	$(p_D=13,1)$ 5419,7	74,3
2: $(r = 1, s = 1), \delta_S = \mathbf{0}$	$E_t = \rho E_{t-1} + \beta_0 Chuva_t + \beta_1 Chuva_{t-1}$	$(p_D=13,6)$ 5416,9	74,0
3: $(r = 1, s = 0), \delta_S \neq \mathbf{0}$	$E_t = \rho E_{t-1} + \beta Chuva_t$	$(p_D=13,9)$ 5301,5	62,4
4: $(r = 1, s = 1), \delta_S \neq \mathbf{0}$	$E_t = \rho E_{t-1} + \beta_0 Chuva_t + \beta_1 Chuva_{t-1}$	$(p_D=14,2)$ 5303,8	62,7
5: $(r = 0, s = 30, d = 2), \delta_S = \mathbf{0}$	$E_t = \sum_{j=0}^{30} \beta_j Chuva_{t-j}$ $\beta_j = \sum_{k=0}^2 \zeta_k j^k$	$(p_D=13,1)$ 5431,2	75,7
6: $(r = 0, s = 30, d = 3), \delta_S = \mathbf{0}$	$E_t = \sum_{j=0}^{30} \beta_j Chuva_{t-j}$ $\beta_j = \sum_{k=0}^3 \zeta_k j^k$	$(p_D=13,4)$ 5415,4	74,0
7: $(r = 1, s = 0), \delta_S = \mathbf{0}$ erros iid	$E_t = \rho E_{t-1} + \beta Chuva_t + \varepsilon_t$ $\varepsilon_t \sim N(0, Q_\varepsilon)$	$(p_D=236,6)$ 5399,9	51,3

r =ordem da autoregressão, s =número de defasagens para o volume de chuva, d =grau do polinômio de aproximação dos parâmetros de defasagem

Segundo o critério DIC, os melhores modelos são aqueles que apresentam termos sazonais, com ligeira vantagem para o modelo 3, de ordem $(r = 1, s = 0)$. Já de acordo com o critério EPD, o melhor modelo é o 7: aquele em que foram introduzidos termos de erro aleatório no preditor η_t . Passamos a descrever os resultados de ambos os modelos.

Análise do Modelo 3

A figura 5.18 exibe os histogramas das amostras das posteriores marginais dos parâmetros envolvidos no modelo, bem como dos níveis de PM_{10} estimados para os instantes $t = 478$ e $t = 479$, em que não havia registro do poluente. Algumas estatísticas referentes a estas amostras são exibidas na tabela 5.4.

Tabela 5.4 Efeito de chuva sobre PM_{10} no RJ - modelo 3: Sumário das amostras da posteriori.

	φ	α	E_0	ρ	β	δ_{Umid}	δ_{Temp}	δ_{Cos}	δ_{Sen}
Mín	16,30	3,15	-0,13	0,35	-0,053	-0,08	0,083	0,015	-0,28
Q1	19,09	3,48	0,53	0,64	-0,039	-0,06	0,121	0,037	-0,25
Mediana	19,80	3,57	0,70	0,68	-0,035	-0,06	0,130	0,047	-0,24
Média	19,82	3,57	0,72	0,67	-0,035	-0,06	0,129	0,048	-0,24
Q3	20,53	3,66	0,89	0,71	-0,031	-0,05	0,137	0,057	-0,23
Máx.	22,91	3,92	1,87	0,82	-0,018	-0,03	0,162	0,107	-0,18
e.p.	0,0295	0,0184	0,0263	0,0020	0,0003	0,0007	0,0019	0,0021	0,0016
	δ_{Seg}	δ_{Ter}	δ_{Qua}	δ_{Qui}	δ_{Sex}	δ_{Sab}	δ_{Fer}	y_{478}	y_{479}
Mín	0,00	0,07	0,08	0,11	0,115	-0,01	-0,211	21,950	24,34
Q1	0,08	0,14	0,16	0,18	0,196	0,08	-0,109	39,420	45,95
Mediana	0,10	0,16	0,18	0,20	0,217	0,11	-0,080	45,050	52,45
Média	0,10	0,16	0,18	0,20	0,218	0,11	-0,079	45,850	52,73
Q3	0,13	0,18	0,20	0,22	0,241	0,13	-0,050	51,350	58,98
Máx.	0,20	0,25	0,28	0,30	0,308	0,20	0,038	87,280	90,51
e.p.	0,0009	0,0008	0,0008	0,0009	0,0009	0,0009	0,0011	0,2539	0,2771

Os erros-padrão estimados indicam bastante precisão no processo de estimação. As estimativas associadas aos dias da semana são significativas, com efeitos comparados ao domingo crescendo à medida em que se aproxima a sexta-feira e decrescendo novamente no sábado. Tal efeito deve-se, possivelmente, ao grande fluxo de veículos usual no município do Rio de Janeiro às sextas-feiras. Feriados apresentam efeito negativo, quando comparados a dias normais, provavelmente devido à diminuição de atividades industriais e de tráfego. A unidade média apresenta efeito negativo e causa certa estranheza a estimativa positiva do efeito da temperatura, que pode estar ocorrendo devido à forte correlação a posteriori apresentada pelo coeficiente desta variável com o nível α (-72%) e com os termos sazonais: δ_{Cos} (68%) e δ_{Sen} (-74%). O efeito imediato da chuva, β , é negativo, indicando decréscimo nos níveis de PM_{10} em função de elevações nos volumes de chuva.

5.2 EFEITO DE CHUVA SOBRE NÍVEIS DE MATERIAL PARTICULADONO RIO DE JANEIRO 131

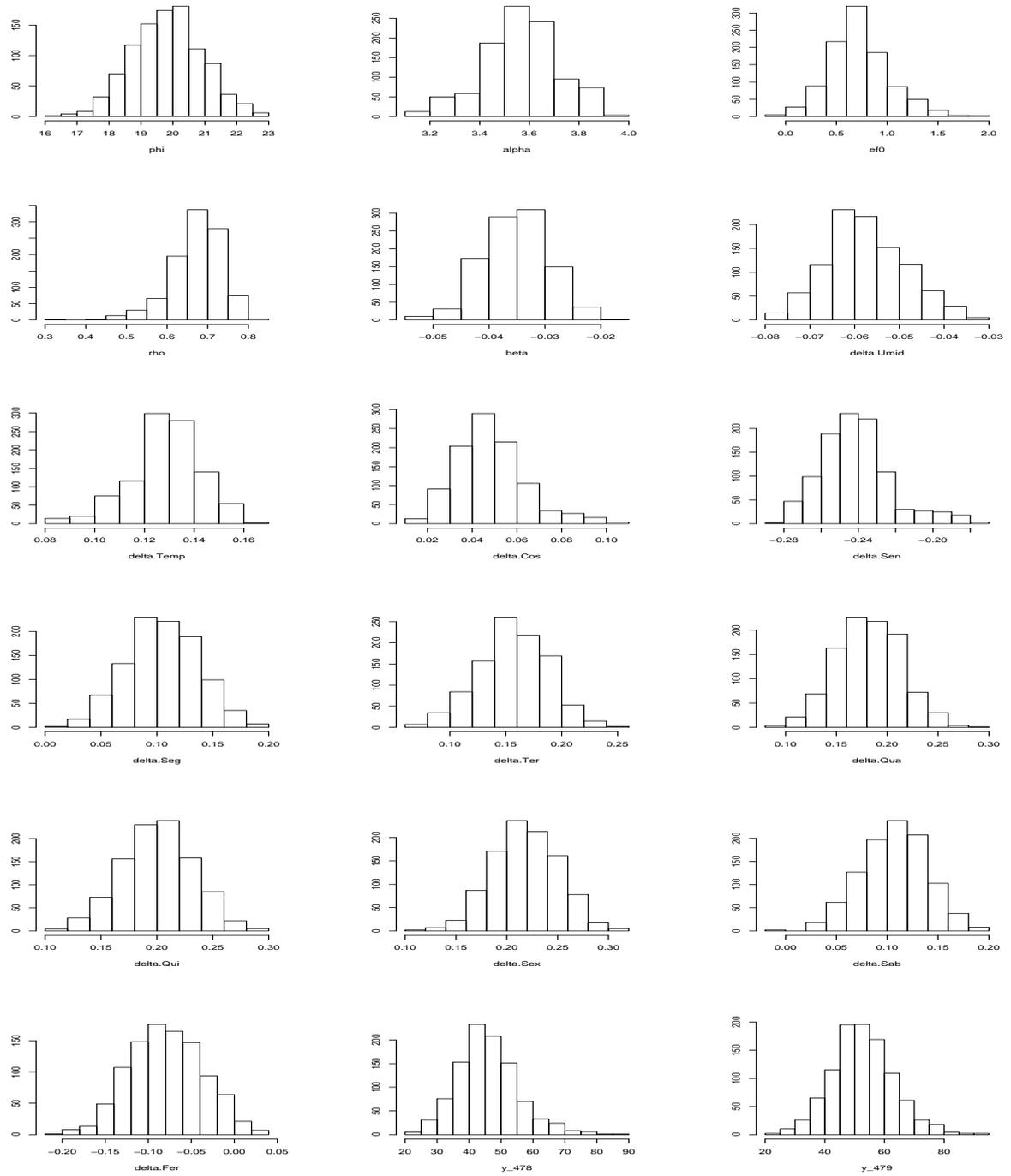


Figura 5.18 Efeito de chuva sobre PM_{10} no RJ - modelo 3: Histogramas das amostras da posteriori.

As funções de resposta ao impulso e de transferência estimadas pela média a posteriori junto a seus limites de credibilidade a 95%, com base em elevações de um desvio-padrão no nível de chuva, são exibidas na figura 5.19. O efeito da chuva sobre o nível de PM_{10} extingue-se após aproximadamente 10 dias. Na figura 5.20, observa-se que o nível máximo de chuva registrado no período de análise conduziria a uma redução imediata da ordem de 35% no nível de material particulado e, considerando-se que não houvesse chuva nos dias subsequentes a este volume, os níveis de PM_{10} tornariam a subir gradualmente.

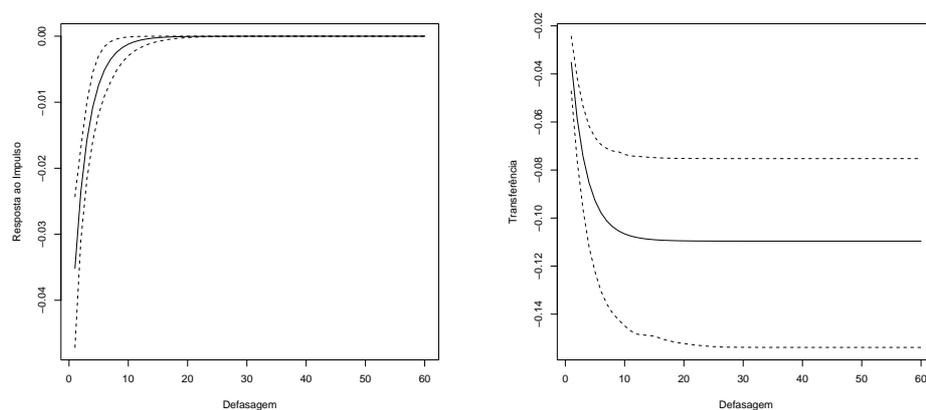


Figura 5.19 Efeito de chuva sobre PM_{10} no RJ - modelo 3: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível de chuva.

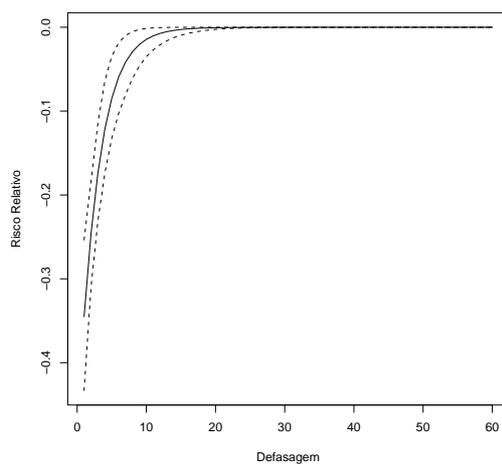


Figura 5.20 Efeito de chuva sobre PM_{10} no RJ - modelo 3: Evolução do risco relativo associado ao nível máximo de chuva observado no período de análise, quando comparado a dias sem chuva.

As 30 últimas observações disponíveis foram reservadas para avaliação da capacidade preditiva do modelo. Na figura 5.21 encontram-se os histogramas das amostras das distribuições preditivas para o período de 02/08/2002 a 31/08/2002. As linhas verticais traçadas em cada histograma indicam o nível de PM_{10} observado em cada um desses dias. Com exceção de alguns poucos dias, as distribuições preditivas concentram-se em torno dos valores observados. As previsões baseadas na média fornecem erro quadrático médio 215,7 (correspondente a um erro relativo médio de 17,7%), praticamente equivalente àquelas baseadas na mediana, que apresentam erro quadrático médio 220,9, o que se traduz em um erro relativo médio de 17,5%.

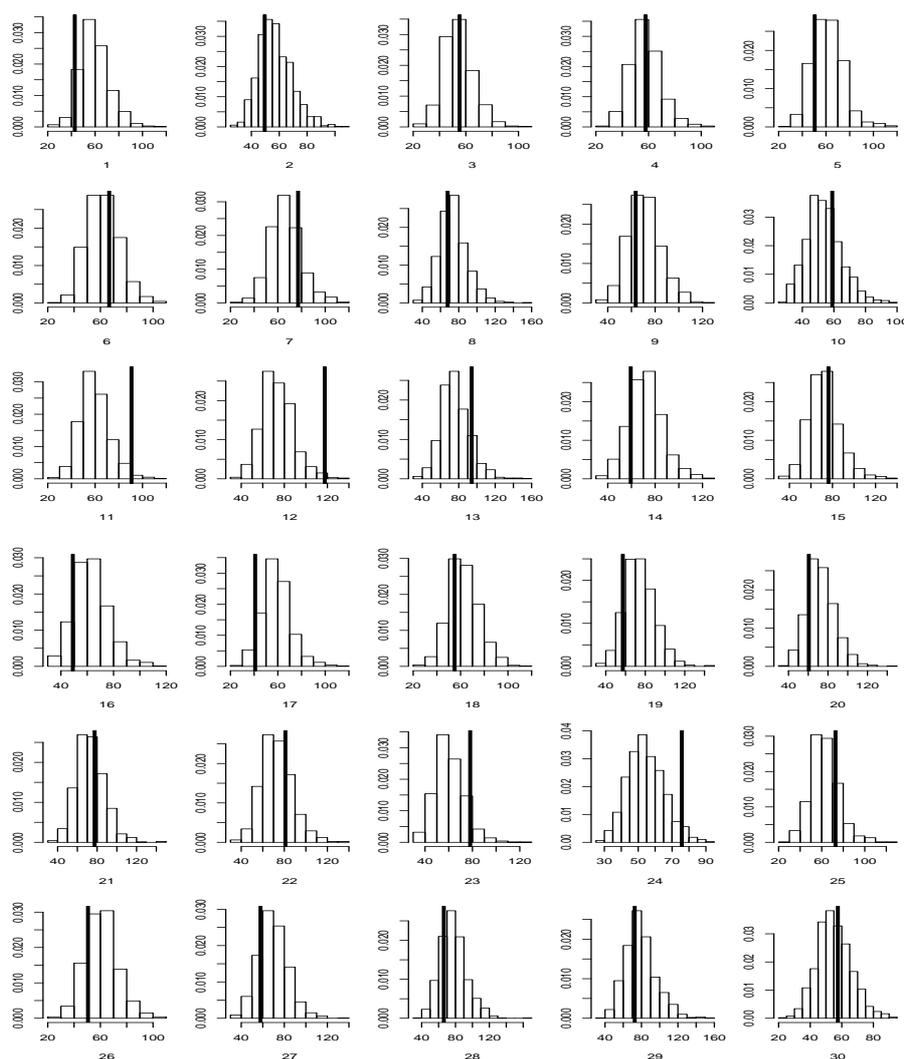


Figura 5.21 Efeito de CO sobre órbitos de crianças em SP - modelo 3: Histogramas das amostras das distribuições preditivas para o número de órbitos, com horizontes variando de 1 a 30.

Análise do Modelo 7

Os histogramas referentes às amostras das distribuições a posteriori dos parâmetros envolvidos no modelo 7 são exibidos na figura 5.22. Estatísticas sumárias associadas a estas amostras encontram-se registradas na tabela 5.5.

Tabela 5.5 Efeito de chuva sobre PM_{10} no RJ - modelo 3: Sumário das amostras da posteriori.

	φ	α	β	ρ	E_0	Q_ε	$\delta.Umid$	$\delta.Temp$	$\delta.Seg$
Mín	16,10	4,11	-0,0881	0,572	-0,07	0,005	-0,05	-0,027	0,00
Q1	21,43	4,27	-0,0693	0,696	0,35	0,015	-0,04	-0,004	0,10
Mediana	24,11	4,32	-0,0651	0,722	0,51	0,019	-0,03	0,002	0,12
Média	24,81	4,33	-0,0652	0,719	0,53	0,019	-0,03	0,001	0,12
Q3	27,43	4,38	-0,0608	0,746	0,68	0,023	-0,03	0,008	0,14
Máx.	48,85	4,66	-0,0447	0,834	1,47	0,041	-0,01	0,032	0,21
e.p.	1,1659	0,0065	0,0005	0,0020	0,0191	0,0016	0,0004	0,0008	0,0020
	$\delta.Ter$	$\delta.Qua$	$\delta.Qui$	$\delta.Sex$	$\delta.Sab$	$\delta.Fer$	y_{478}	y_{479}	
Mín	0,05	0,08	0,10	0,15	0,024	-0,24	25,540	29,260	
Q1	0,15	0,18	0,19	0,21	0,091	-0,13	45,030	50,220	
Mediana	0,17	0,20	0,21	0,23	0,117	-0,10	52,810	59,200	
Média	0,17	0,20	0,21	0,24	0,116	-0,10	53,830	60,480	
Q3	0,19	0,22	0,23	0,26	0,141	-0,08	60,460	69,160	
Máx.	0,27	0,31	0,32	0,34	0,210	0,02	116,100	119,200	
e.p.	0,0021	0,0021	0,0021	0,0020	0,0021	0,0025	0,7420	0,8586	

Assim como ocorreu no modelo 3, observa-se um efeito de aumento do poluente à medida em que a sexta-feira se aproxima, com decréscimo no sábado. Tem-se também um decaimento significativo nos níveis de PM_{10} em feriados. A temperatura aparece, no presente modelo, como variável não significativa e umidade apresenta mais uma vez efeito negativo. A persistência do efeito da chuva não se altera significativamente em relação ao modelo 3, mas o impacto imediato da chuva passa a ser maior (em módulo).

5.2 EFEITO DE CHUVA SOBRE NÍVEIS DE MATERIAL PARTICULADONO RIO DE JANEIRO 135

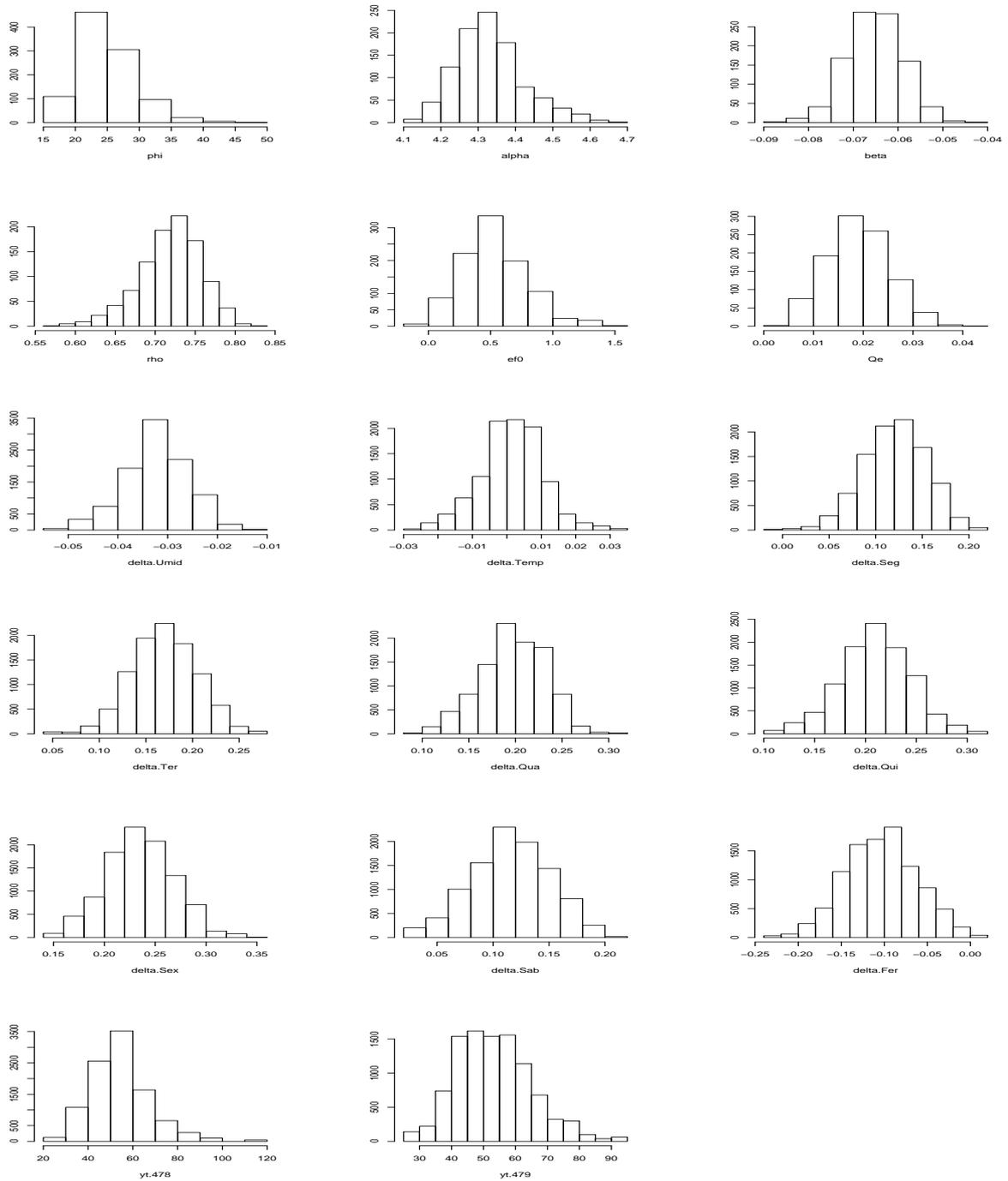


Figura 5.22 Efeito de chuva sobre PM_{10} no RJ - modelo 8: Histogramas das amostras da posteriori.

A figura 5.23 exibe as funções de resposta ao impulso e de transferência associadas a elevações de um desvio-padrão nos níveis de chuva e, na figura 5.24, observa-se o risco relativo associado ao nível máximo de chuva observado no período de análise, comparado a um dia sem chuva. Como se pode ver, de acordo com o modelo 7, tal volume de chuva ocasionaria uma queda imediata no nível de PM_{10} de mais que 50%, com os níveis do poluente voltando a seu patamar usual após cerca de 20 dias.

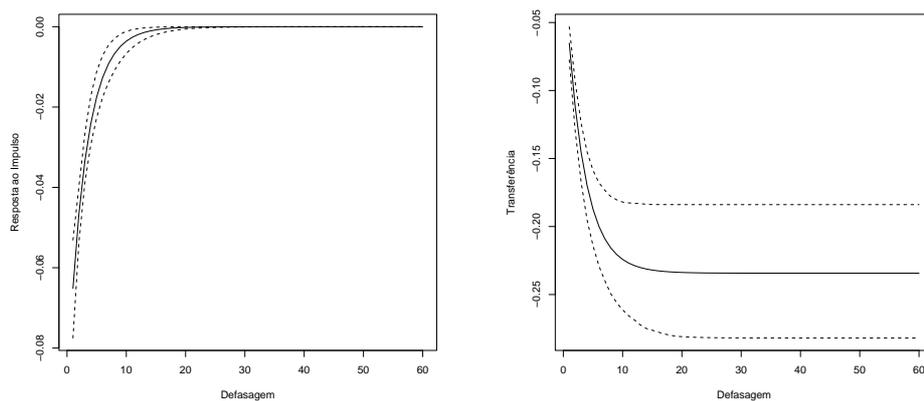


Figura 5.23 Efeito de chuva sobre PM_{10} no RJ - modelo 7: Funções de resposta ao impulso e de transferência associadas a elevações de 1 desvio-padrão no nível de chuva.

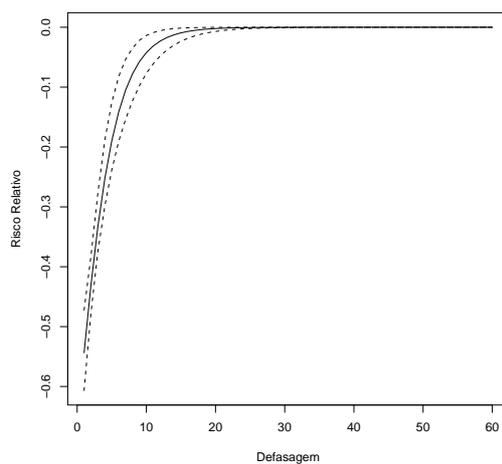


Figura 5.24 Efeito de chuva sobre PM_{10} no RJ - modelo 7: Evolução do risco relativo associado ao nível máximo de chuva observado no período de análise, quando comparado a dias sem chuva.

A figura 5.25 exibe histogramas das amostras das distribuições preditivas para os níveis de PM_{10} nos 30 últimos dias do período de análise. As linhas verticais em cada histograma indicam os reais valores observados. Com exceção do 12º dia no horizonte de previsão, em que se observou um nível atípico para o poluente, bem superior à média deste período final, as distribuições preditivas capturaram bastante bem os verdadeiros valores observados para o poluente. O erro quadrático médio associado à média foi 254,6 (correspondente a um erro relativo médio de 17,0%). Já para a mediana das distribuições preditivas, observou-se um erro quadrático médio de 262,5 e erro relativo médio de 16,8%.

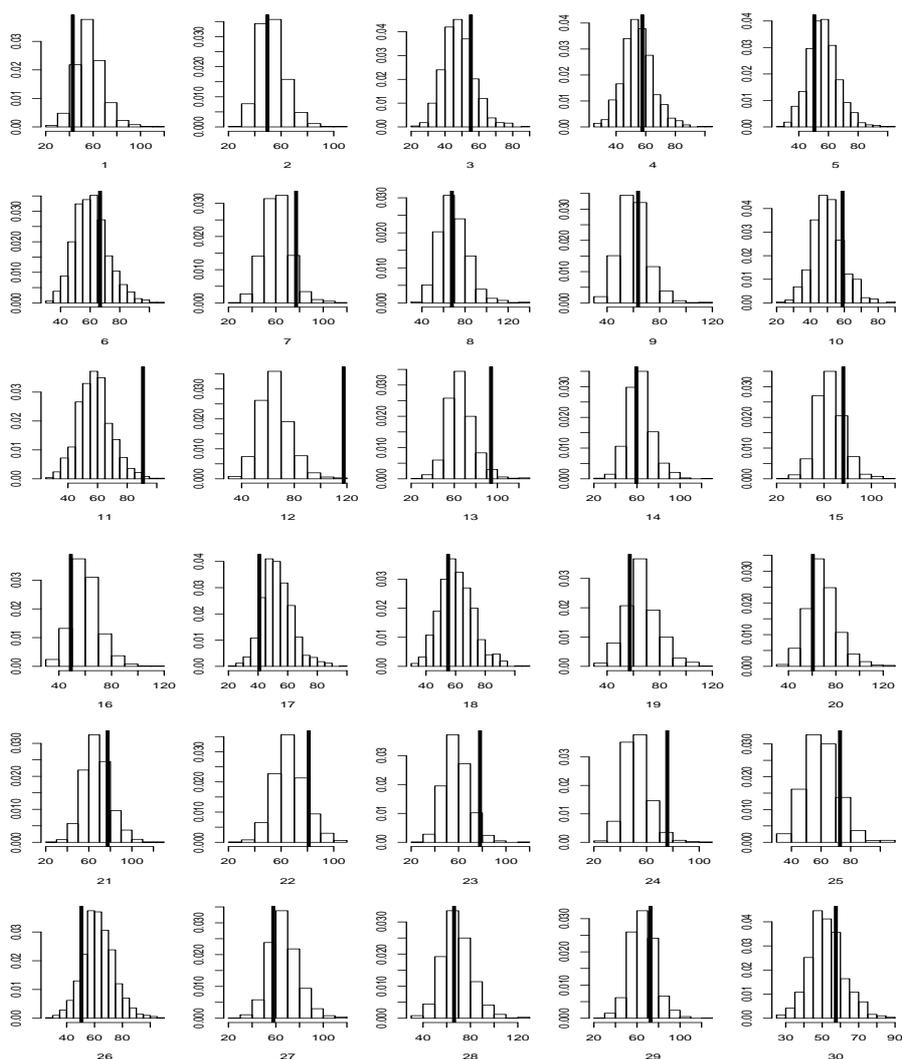


Figura 5.25 Efeito de CO sobre órbitos de crianças em SP - modelo 7: Histogramas das amostras das distribuições preditivas para o número de órbitos, com horizontes variando de 1 a 30.

Como o modelo 3 apresentou menor erro preditivo e é bem mais parcimonioso que o modelo 7, julgamos que este seja o melhor modelo entre os ajustados.

5.3 Conclusões

Na primeira aplicação apresentada, fica claro que o efeito de elevações nos níveis Monóxido de Carbono sobre o desfecho analisado perdura por diversos dias. De acordo com o modelo 5, a elevação do nível médio ao nível máximo do poluente observados no período de análise chegaria a acarretar um aumento imediato de 20% no número esperado de óbitos. Ao se tornar a estrutura que governa o impacto do poluente mais flexível, no modelo 10, verifica-se que tais elevações apresentam-se significativas em alguns instantes, e em outros não, o que provavelmente ocorre devido aos mecanismos de interação do poluente com outras fontes associadas ao desfecho (incorporadas ou não no modelo). Julgamos que seja válido investigar a inclusão de novos termos no preditor desse modelo, mantendo a evolução temporal do coeficiente β_t , de forma a capturar outras fontes de incerteza que talvez sejam a razão para os bons resultados associados ao modelo 5, que conta com termos de erro não estruturado no preditor. Cabe ressaltar que, embora os resultados de ajuste e preditivos do modelo 5 sejam satisfatórios, a incerteza representada pelos erros não tem sua fonte explicitada, ao contrário do que ocorre no modelo 10.

Na aplicação sobre níveis de material particulado, além de se verificar a significância do efeito imediato da chuva, tendo-se chegado a obter uma estimativa de decréscimo em torno de 35% no nível esperado do poluente, associado ao volume de chuva máximo observado, constatou-se também a propagação desse efeito nos dias subsequentes.

Um aspecto bastante interessante nesta segunda aplicação é o fato de terem sido as observações omissas preenchidas durante o procedimento de estimação e, portanto, os resultados obtidos refletem essa fonte adicional de incerteza.

Em ambas as aplicações, pode-se verificar a relevância na estimação de modelos com preditores não-lineares associados a respostas não-gaussianas: discreta, no primeiro caso, e contínua, no segundo, tendo sido possível a identificação de todos os modelos propostos. No caso da segunda aplicação, modelos que apresentam funções de transferência dinâmicas ainda apresentam problemas de convergência. Atualmente estamos investigando as causas desses problemas, que provavelmente ocorrem devido à quantidade de informação disponível (700 dias), bem menor que na aplicação de óbitos em São Paulo, em que se dispôs de 4 anos de observações diárias, totalizando 1461 dias.

No próximo capítulo, faz-se um breve resumo dos aspectos desenvolvidos ao longo deste trabalho, apontando-se alguns tópicos para pesquisa futura.

Conclusões e Trabalhos Futuros

As possíveis causas para violações das hipóteses básicas de um modelo dinâmico linear são diversas. Ao longo deste trabalho, tratou-se da estimação Bayesiana de modelos dinâmicos com estrutura observacional pertencente à família exponencial - portanto, não necessariamente normal - e contendo, em seus preditores, funções de transferência - o que torna tais preditores não-lineares. Tais modelos encaixam-se, dessa forma, na categoria dos modelos dinâmicos não normais e não lineares. Como visto na subseção 2.2.2, até o final da década de 1980, a obtenção da distribuição a posteriori para esse tipo de modelo restringia-se a uma quantidade limitada de características associadas a esta: momentos de primeira e segunda ordens e modas a posteriori ou, alternativamente, obtinham-se distribuições aproximadas, ao custo de transformações, de forma a se atingir normalidade da resposta e/ou linearidade do preditor. A partir da década de 1990, com o advento de métodos que postulam a obtenção de amostras da posteriori, adquiriu-se a flexibilidade de, uma vez obtida uma amostra representativa da distribuição a posteriori, construir-se estimativas de Monte Carlo para quaisquer quantidades associadas a tal distribuição. Em suma: é possível fazer inferência Bayesiana completa sob não-normalidade e não-linearidade. Os métodos de Monte Carlo via cadeias de Markov foram adotados, neste trabalho, para este fim. Especificamente, aplicou-se o amostrador de Gibbs (com passos Metropolis-Hastings) para a construção de cadeias de Markov que tivessem, como distribuição estacionária, a posteriori conjunta dos parâmetros envolvidos nos modelos propostos. Passamos a discorrer, na seção a seguir, sobre aspectos relevantes ao longo do desenvolvimento do trabalho. Finalmente, na seção 6.2 são levantados tópicos para pesquisa futura.

6.1 Comentários Gerais

6.1.1 Aspectos Computacionais

Do ponto de vista teórico, sob certas condições impostas sobre os núcleos de transição de forma a produzir cadeias homogêneas, ergódicas e irredutíveis - atendidas pelas propostas adotadas neste trabalho - os algoritmos MCMC certamente conduzirão à distribuição estacionária.

Portanto, a eficiência computacional desses algoritmos está ligada ao tempo para que a cadeia de Markov construída atinja a distribuição a posteriori. Assim sendo, do ponto de vista prático, devem-se buscar estratégias para que a convergência ocorra o mais brevemente possível. Em modelos dinâmicos, a dimensão do vetor paramétrico costuma ser bastante elevada e, ainda, os parâmetros de estado (aqueles que variam estocasticamente) são serialmente correlacionados, o que dificulta a convergência. Deve-se, então, na estratégia de amostragem, decidir pela atualização destes parâmetros individualmente ou em blocos. Tem-se o seguinte impasse: ao se tomar blocos muito grandes, torna-se difícil fazer movimentos que sejam adequados para todas as componentes do bloco, o que leva a baixas taxas de aceitação no algoritmo Metropolis-Hastings e, conseqüentemente, à possibilidade de que a cadeia gerada fique limitada a subregiões do espaço de estados. Em contrapartida, os movimentos individuais sobre os parâmetros de estado ficam comprometidos pela autocorrelação inerente a estes. A estratégia aqui tomada foi guiada pela comparação, feita por Gamerman (1998), desses dois esquemas a um terceiro, que se mostrou mais favorável: a reparametrização dos modelos dinâmicos em termos dos erros de evolução e amostragem individual destes últimos, que, a priori, não são correlacionados.

Outro ponto que merece destaque é a forma adotada para a densidade proposta no algoritmo Metropolis-Hastings. Neste trabalho, ao se amostrar parâmetros fixos no tempo, oscilou-se entre propostas passeio-aleatório, com variâncias calibradas de forma a se otimizar as taxas de aceitação, e a proposta sugerida por Gamerman (1998), dada pela condicional completa obtida a partir do modelo normal postulado para as observações ajustadas (2.10). Nesse caso, como a dimensão do vetor paramétrico é pequena, é bastante simples calibrar as propostas e como estas são simétricas, tem-se uma simplificação das razões de aceitação (2.15), que reduzem-se a razões de posteriors, o que facilita o código para implementação do algoritmo. Por outro lado, ao se permitir a evolução estocástica para pelo menos um dos parâmetros envolvidos no modelo, torna-se muito difícil calibrar as variâncias de propostas passeio-aleatório de forma a garantir movimentação por todo o espaço de estados. Nessas condições, a proposta baseada em (2.10), que é uma aproximação para a posteriori, forneceu altas taxas de aceitação, mostrando-se bastante superior à proposta passeio-aleatório e conduzindo à convergência bem mais rapidamente. De fato, houve casos em que, ao tentar utilizar propostas passeio-aleatório não obtivemos convergência.

Quanto à verificação de convergência, não foi realizado qualquer teste "formal" de convergência, mesmo porque tais testes são capazes apenas de fornecer evidências de convergência, mas não são conclusivos. Em todas as aplicações feitas, buscou-se gerar cadeias bastante longas, de forma a tentar minimizar a possibilidade de se obter amostras de distribuições ótimas apenas localmente. Optou-se por gerar duas cadeias paralelas e avaliar se seus traços mesclavam-se bem.

6.1.2 Identificação dos Modelos Propostos

De forma geral, os exercícios com dados simulados apontaram as seguintes evidências: funções de transferência de ordem ($r = 1, s \leq 1$) foram estimadas bastante bem, mas para val-

ores $s \geq 2$, houve dificuldade na identificação dos parâmetros associados às defasagens da regressora X . A adição de termos de erro independentes, visando capturar fontes de variação não presentes no preditor, não comprometeu a estimação dos demais parâmetros estruturais. Quanto aos modelos que apresentaram parâmetros evoluindo estocasticamente, houve certa dificuldade, para dados Poisson, na estimação do parâmetro autoregressivo, ρ , ao se permitir evolução do nível α_t , obtendo-se subestimativas para ρ , o que, evidentemente, acarretou efeitos de X (medidos pela função de resposta ao impulso) menos duradouros que o esperado. Presumimos que este tipo de problema possa ser causado pelo fato de que a evolução do nível segundo um passeio aleatório carrega, a cada instante, a memória do processo sob análise, "competindo", assim, com o parâmetro ρ . Ao aplicar a mesma estrutura a dados com distribuição Binomial, entretanto, o parâmetro autoregressivo foi muito bem estimado, o que sugere que uma investigação mais profunda desse tipo de estrutura seja recomendável, trabalhando-se com diferentes especificações probabilísticas para a variável resposta, aliadas a variadas especificações de valores para as componentes do vetor paramétrico. É possível que, dada uma estrutura observacional, o modelo proposto não seja identificável para determinada magnitude dos parâmetros subjacentes, mas que o seja para outra. Já os modelos com parâmetro de choque instantâneo da regressora, β_t , variando segundo um passeio aleatório, apresentaram-se muito bem estimados, recuperando-se bastante bem tanto as funções de resposta ao impulso e transferência teóricas quanto as funções de resposta média. Este tipo de modelo mostra-se bastante versátil, uma vez que determina, a cada instante, funções de resposta ao impulso que, apesar de sujeitas à mesma forma analítica, podem ter suas magnitudes atualizadas de acordo com variações ocorridas nos sistemas analisados. Sua estimação para diferentes respostas, inclusive de natureza discreta, constitui-se, assim, em uma das maiores contribuições deste trabalho.

Cabe registrar que, em todos os exercícios simulados em que algum dos parâmetros evoluiu dinamicamente, as variâncias dos erros de evolução apresentaram-se subestimadas, fornecendo trajetórias para estes parâmetros mais suaves que suas trajetórias teóricas. Embora as trajetórias estimadas tenham capturado satisfatoriamente o padrão teórico, esforços no sentido de melhor estimar as variâncias de evolução são válidos.

Quanto à natureza da distribuição de Y_t , não se observou problemas além dos já relatados associados a respostas Poisson e Gama. Para respostas distribuídas segundo o modelo Binomial(n, p_t), parece haver maior facilidade de identificação à medida em que n aumenta. De fato, chegamos a fazer alguns exercícios para dados simulados de acordo com o modelo Bernoulli, para o qual $n = 1$ e, nesse caso, o parâmetro autoregressivo apresentou-se severamente subestimado. Uma possível explicação para tais resultados é a seguinte: observe-se que o coeficiente de variação em um modelo Binomial é proporcional a $1/\sqrt{n}$. Tomando-se tal coeficiente como uma medida de incerteza associada aos dados (já livre de sua escala de mensuração), tem-se a incerteza tendendo a zero quando $n \rightarrow \infty$. Em suma: a informação contida nos dados aumenta com n e daí os melhores resultados obtidos com valores maiores para este parâmetro. O coeficiente de variação depende, ainda, de p_t , e, portanto, dos valores atribuídos ao vetor paramétrico, uma vez que estes determinam o preditor atribuído a $\text{logito}(p_t)$. Assim, a discussão acima somente faz sentido ao se aumentar o valor de n para valores de p_t fixos.

Para a distribuição Gaussiana Inversa, houve problemas na identificação do parâmetro σ^2 . Como toda a estrutura preditiva é colocada sobre a média, μ_t , e σ^2 é ortogonal a este parâmetro,

a estimação dos parâmetros estruturais que compõem o preditor não traz informação para a estimação de σ^2 . Sendo assim, o investimento no sentido de se atribuir estruturas preditivas também a σ^2 pode ser promissor.

6.1.3 Alguns aspectos sobre as aplicações desenvolvidas

Em ambas as aplicações apresentadas, a primeira envolvendo o efeito de um poluente sobre um desfecho epidemiológico e na segunda, a modelagem dos níveis de um poluente em função de variáveis meteorológicas e de calendário, observou-se que o efeito de um pulso em uma das regressora perdura por vários dias. Assim, para avaliação do real efeito acumulado pela regressora, a utilização de modelos de função de transferência (contrapondo-se à abordagem usual nas aplicações desenvolvidas nessa área de avaliação de efeitos de curto prazo) mostrou-se extremamente propícia. Ainda, a flexibilidade intrínseca aos modelos dinâmicos adotados permite que sejam ajustados modelos que melhor reflitam a realidade.

Uma das premissas na filosofia de inferência Bayesiana reside no fato de que, idealmente, deve-se trazer à luz e conjuntamente todas as incertezas envolvidas no problema. Na segunda aplicação desenvolvida no capítulo 5, havia dois valores omissos para a variável resposta, os quais foram estimados pelo próprio procedimento MCMC, incorporando-se às distribuições a posteriori a variabilidade devida ao desconhecimento de tais valores. Na primeira aplicação, entretanto, havia 56 instantes sem registro para a regressora Monóxido de Carbono. Nessa aplicação, as observações omissas foram preenchidas utilizando o aplicativo desenvolvido por Junger (2002), com metodologia baseada no algoritmo EM e fortemente influenciada pelo ajuste de um modelo ARIMA que dê conta do padrão temporal existente nas séries de poluentes em cada estação de mensuração. Como a série temporal disponível em cada estação de mensuração é bastante longa, obteve-se valores de preenchimento que parecem bastante razoáveis. Esta é apenas uma dentre as diversas possibilidades para preenchimento de dados faltantes e, como todo método de inserção prévia de dados omissos, apresenta uma forte limitação, principalmente no contexto Bayesiano, e com o algoritmo utilizado, a incerteza associada às estimativas obtidas é subestimada, uma vez que os dados de preenchimento são, de fato, tratados como valores observados. O preenchimento via algoritmo MCMC contornaria este problema, obviamente envolvendo algum custo computacional. Tentou-se estimar tais quantidades utilizando passos adicionais no algoritmo MCMC, entretanto, com a estrutura proposta para os modelos, não foi possível obter informação suficiente para estimar estes valores omissos. Uma alternativa para se obter informação suficiente para estimação de valores de X_t seria a construção de modelos hierárquicos em que se determine uma estrutura preditiva para X_t e, em um segundo nível, a resposta esperada seja modelada condicionalmente a X_t . Na aplicação desenvolvida, poderia-se pensar, por exemplo, na modelagem dos níveis de Monóxido de Carbono em função de variáveis meteorológicas e, completando a estrutura hierárquica, o número esperado de óbitos seria modelado em função dos níveis do poluente e também de variáveis climáticas.

Outra questão relevante quanto às aplicações desenvolvidas é o fato de só termos atribuído efeitos cumulativos através do tempo a uma das regressoras presentes no preditor. Em ambas as

aplicações, os efeitos de temperatura e umidade foram considerados instantâneos, contrariando a intuição de que oscilações sobre estas variáveis apresentem uma dinâmica própria de impacto sobre a resposta esperada. Inicialmente, tentaram-se modelagens do tipo:

$$E_t = \rho E_{t-1} + (X_t, Temperatura_t, Umidade_t)\beta, \quad \beta' = (\beta_1, \beta_2, \beta_3).$$

Não foi possível, na formulação acima, identificar o parâmetro autoregressivo, uma vez que se tentou atribuir a mesma medida de memória a variáveis que podem apresentar diferentes dinâmicas. Pode-se relaxar estas condições, adotando-se abordagens como a descrita na subseção 6.2.3

6.2 Tópicos para Trabalho Futuro

6.2.1 Modelagem de efeitos que apresentem pontos de inflexão

Ao se analisar problemas nos quais o efeito da regressora sobre a resposta esperada leve algum tempo até atingir seu ápice, faz-se necessário, de acordo com a formulação introduzida por Koyck (1954), aplicar diversas defasagens em X . Em aplicações nas quais os dados encontram-se agregados temporalmente, dificilmente se observa esse tipo de padrão, mas se a unidade temporal for mais refinada, é bastante razoável supor-se que o efeito de X possa aumentar por vários instantes, até que finalmente atinja seu valor mais extremo, para então tornar a decair. Como visto ao longo deste trabalho, modelos definidos em termos de defasagens na regressora X são propensos a problemas de identificação. Como discutimos anteriormente, para evitar problemas de estimação acarretados pela autocorrelação entre X_t, X_{t-1}, \dots , uma abordagem alternativa consiste em assumir-se, para um modelo de defasagens distribuídas (ou modelo de função de transferência de forma livre) que os coeficientes $\beta_j, j = 0, 1, \dots, s$, associados às diversas defasagens estejam relacionados por alguma forma funcional. Neste sentido, aplicamos, em particular, restrições polinomiais propostas por Almon (1965). Quanto maior o grau do polinômio suavizador, mais flexível é o formato da trajetória dos coeficientes β_j , obtendo-se, já com graus $d = 2$ ou $d = 3$, trajetórias de decaimento geométrico ou crescimento até um valor máximo e posterior decaimento do efeito de X . Entretanto, além da necessidade de escolha do grau do polinômio de aproximação para os coeficientes β_j , uma severa limitação associada a essa abordagem encontra-se na necessidade de se especificar um número máximo de defasagens, s , assumindo-se assim que $\beta_j = 0, j > s$ e correndo-se o risco de subestimar o efeito total acumulado pela regressora a cada instante. Solow (1960) propõe uma forma de restrição aos coeficientes $\beta_j, j = 1, 2, \dots$, supondo-se que os mesmos sigam o padrão ditado pela distribuição Pascal. Tem-se, então:

$$\beta_j = \beta \binom{r+j-1}{j} (1-\rho)^r \rho^j.$$

Ao se substituir os valores de β_j assim determinados em $E_t = \beta_0 X_t + \beta_1 X_{t-1} + \dots$ e fazendo-se $E_t - (1 - \rho B)^r E_t$, B o operador defasagem, chega-se a:

$$E_t - r\rho E_{t-1} + \binom{r}{2} \rho^2 E_{t-2} - \dots + (-1)^r \rho^r E_{t-r} = \beta(1 - \rho)^r X_t.$$

A estrutura acima (à qual, em geral, adiciona-se um termo de erro) provê diversos perfis possíveis para os efeitos de X : a distribuição Pascal é sempre assimétrica positiva, e quanto maior o valor de ρ e menor o valor de r , mais acentuada é a assimetria, sendo o modelo de Koyck obtido como caso particular quando $r = 1$.

Ravines et al. (2006) utiliza as restrições de Solow em modelos gaussianos ajustados para dados trimestrais e, ao permitir valores 1, 2 e 3 para r (ora fixando estes valores e escolhendo o melhor modelo por critérios de seleção, ora atribuindo priori discreta aos três valores propostos para r e avaliando sua posteriori), obtém melhores resultados para $r = 1$. Seria interessante avaliar a estimação desse tipo de modelo para respostas não Gaussianas e, ainda, para dados com menor nível de agregação, como os dados diários das aplicações desenvolvidas neste trabalho, onde haveria maiores chances de se obter valores $r > 1$. Ainda, podem-se buscar parametrizações alternativas e parcimoniosas para efeitos sujeitos a pontos de inflexão.

6.2.2 Modelagem espaço-temporal

Em muitas aplicações, é importante modelar adequadamente não só a forma de propagação de efeitos de uma regressora sobre a resposta esperada no tempo, mas também o comportamento desses efeitos no espaço. Nas aplicações desenvolvidas nesse trabalho, por exemplo, tem-se registros de níveis de poluentes em várias estações de mensuração. Ao trabalharmos com uma única medida, representada pela média diária desses registros sobre todas as estações, implicitamente postulamos que os comportamentos assumidos propaguem-se de forma homogênea em toda a região observada. Na modelagem referente ao número de óbitos de crianças em São Paulo, por exemplo, restringimo-nos a essa hipótese devido ao fato de que, embora os níveis de poluentes sejam tomados em diferentes pontos no espaço, o desfecho em questão (número de óbitos), foi registrado apenas como a contagem total na cidade de São Paulo.

Suponha-se, entretanto, que as informações tanto para o desfecho quanto para as regressoras estejam disponíveis para m localizações, registradas por T períodos de tempo. Diversas são as formas possíveis para a modelagem espaço-temporal da propagação dos efeitos de uma regressora X sobre a resposta esperada. Supondo-se, inicialmente, funções de transferência de ordem ($r = 1, s = 0$), pode-se, por exemplo, assumir:

- que o efeito de X se propague da mesma maneira em todas as localizações $i, i = 1, \dots, m$, como em Fernandes (2006):

$$E_{i,t} = \rho E_{i,t-1} + \beta X_{i,t};$$

- que, em cada estação, os impactos imediatos de X sobre a resposta possam variar com o tempo:

$$E_{i,t} = \rho E_{i,t-1} + \beta_t X_{i,t};$$

- que os impactos imediatos de X sobre a resposta possam variar de acordo com a localização:

$$E_{i,t} = \rho E_{i,t-1} + \beta_i X_{i,t};$$

- que a velocidade do decaimento do efeito de X sobre a resposta varie segundo a localização:

$$E_{i,t} = \rho_i E_{i,t-1} + \beta X_{i,t};$$

- que a memória do processo seja fixa no tempo e no espaço, mas que os choques imediatos de X sobre a resposta variem temporal e espacialmente:

$$E_{i,t} = \rho E_{i,t-1} + \beta_{i,t} X_{i,t}.$$

Obviamente, se o número de localizações é razoavelmente elevado e se as séries temporais observadas em cada uma das localizações são extensas, algumas das especificações acima podem implicar um vetor paramétrico a estimar com dimensão bastante alta e a viabilidade da identificação de tais modelos deve ser avaliada caso a caso.

Todas as especificações acima pressupõem a mesma forma de propagação: decaimento geométrico dos efeitos de X ao longo do tempo. Uma outra tentativa que pode ser feita é a de se supor que mesmo o formato da propagação possa variar de acordo com a localização, em determinados pontos decaindo sempre, em outros apresentando inflexão, o que poderia ser feito de diferentes maneiras: uma primeira forma seria a especificação de:

$$E_{i,t} = \rho E_{i,t-1} + \beta_{i0} X_{i,t} + \beta_{i1} X_{i,t-1} + \beta_{is} X_{i,t-s},$$

com a estimação dos parâmetros β indicando a significância ou não dos mesmos (o que determinaria diferentes formatos). Como já relatado, entretanto, para valores altos de s teria-se problemas na identificação dos modelos assim especificados. Para contornar tal problema, já se mencionou a possibilidade de utilização de abordagens como a de Solow. Pode-se tentar especificar modelos de Solow em que o parâmetro r , que determina se a função de resposta ao impulso é monótona ou não, varie espacialmente, atribuindo-se a este parâmetro prioris discretas definidas sobre poucos pontos, como faz Ravines et al. (2006) para uma única localização. Se for possível determinar a posteriori de $r_i, i = 1, \dots, s$, pode-se chegar a diferentes formatos de propagação através do espaço.

6.2.3 Tratamento do efeito acumulado de mais que uma regressora

Muitas vezes é necessário construir-se preditores nos quais presume-se que mais que uma regressora apresente efeitos sobre a resposta que se propagam por alguns períodos. Se a idéia de

se trabalhar com modelos de defasagens distribuídas para uma única regressora já traz complicações devido à autocorrelação serial associada a tal regressora, mais complexo ainda torna-se o problema ao se tentar atribuir defasagens distribuídas a outra(s) regressora(s), uma vez que, além de autocorrelações seriais, tem-se então que lidar com a possível correlação entre regressoras ao longo do tempo. Assim, parametrizações parcimoniosas, como aquelas adotadas nas funções de transferência com forma funcional específica, podem ser mais adequadas.

Já iniciamos a análise de modelos contendo mais que uma função de transferência, mas ainda em um estágio incipiente. A figura 6.1 exibe os histogramas de amostras das posteriores marginais das componentes de $\theta = (\alpha, \rho_1, \rho_2, \beta_1, \beta_2, E_{0,1}, E_{0,2})$, obtidas ao se ajustar, para dados simulados, o modelo:

$$\begin{aligned} y_t &\sim \text{Poisson}(\lambda_t) \\ \log(\lambda_t) &= \alpha + E_{1,t} + E_{2,t} \\ E_{1,t} &= \rho_1 E_{1,t-1} + \beta_1 X_{1,t} \\ E_{2,t} &= \rho_2 E_{2,t-1} + \beta_2 X_{2,t}. \end{aligned}$$

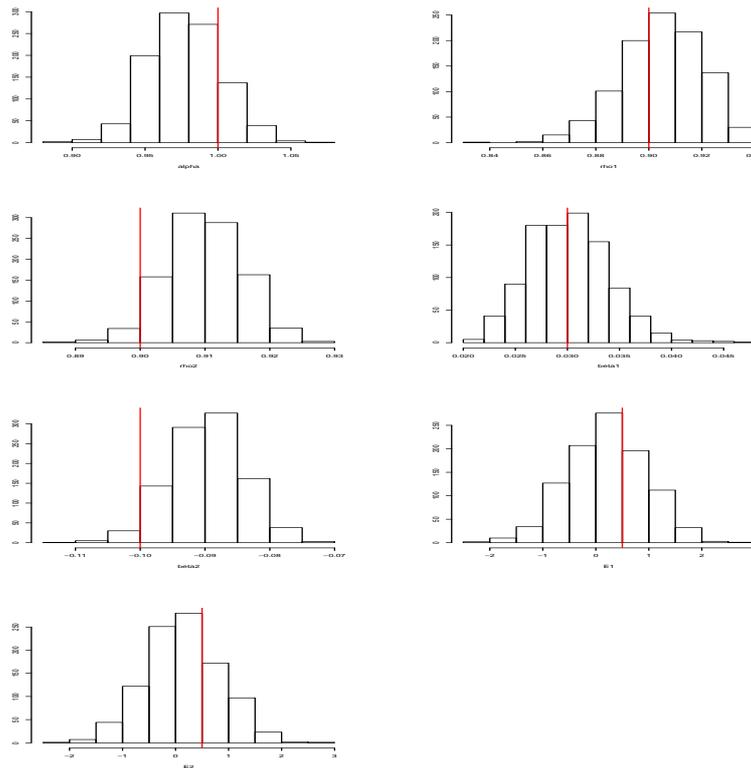


Figura 6.1 Estimação de um modelo contendo duas funções de transferência: Histogramas das amostras da posteriori.

Assumiram-se as seguintes distribuições a priori: $\rho_1, \rho_2 \stackrel{iid}{\sim} U(0, 1)$, $\alpha, \beta_1, \beta_2 \stackrel{iid}{\sim} N(0, 10^5)$. Ao se utilizar prioris tão vagas quanto estas para os efeitos iniciais das regressoras, após 30000

iterrações, as cadeias referentes a estes efeitos ainda não haviam convergido. Adotaram-se, então, para estes parâmetros, as priors: $E_{0,1}, E_{0,2} \stackrel{iid}{\sim} N(0, 1)$. As linhas vermelhas nos histogramas indicam os valores teóricos utilizados para a geração do conjunto de dados. Como se pode ver, foi possível identificar todos os parâmetros envolvidos. Pretendemos, em breve, adotar estruturas variadas para as funções de transferência associadas a cada regressora.

Referências Bibliográficas

- Almon, S. (1965). The distributed lag between capital appropriations and expenditures. *Econometrica*, 33:178–96.
- Alves, M. B. (2003). Funções de transferência em modelos dinâmicos lineares generalizados bayesianos. *Cadernos do IME - Série Estatística*, 15:36–52.
- Alves, M. B. (2005). Comparação de efeitos de poluentes sobre óbitos de crianças e idosos em São Paulo. Parte dos resultados do projeto de extensão Poluição versus Saúde, desenvolvido na Universidade do Estado do Rio de Janeiro.
- Botter, D. A., Jørgensen, B., e Peres, A. A. Q. (2000). A longitudinal study on mortality and air pollution for São Paulo, Brazil. Relatório técnico, Department of Statistics and Demography, Odense University.
- Box, G. E. P., Jenkins, G. M., e Reinsel, G. C. (1994). *Time Series Analysis: Forecasting and Control*. Prentice-Hall, Englewood Cliffs, Nova Jersey, 3ª edição.
- Carlin, B. P., Polson, N. G., e Stoffer, D. S. (1992). A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association*, 87(418): 493–500.
- Carter, C. K. e Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81 (3):541–53.
- Conceição, G. M. S., Miraglia, S. G. E. K., Kishi, H. S., Saldiva, P. H. N., e Singer, J. M. (2001). Air pollution and children mortality - a time series study in São Paulo, Brazil. *Environmental Health Perspectives*, 109:347–50.
- De Jong, P. e Shephard, N. (1995). The simulation smoother for time series models. *Biometrika*, 82(2):339–50.
- Durbin, J. e Koopman, S. J. (2000). Time series analysis of non-Gaussian observations based on state space models from both classical and Bayesian perspectives (com discussão). *Journal of the Royal Statistical Society, série B*, 62(1):3–56.
- Durbin, J. e Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, 89:603–15.

- Edlund, P.O. (1987). On identification of transfer function models. Relatório técnico, Departamento de Estatística, Universidade de Estocolmo.
- Estrella, B., Estrella, R., Oviedo, J., Narváez, X., Reyes, M. T., Gutiérrez, M., e Naumova, E. N. (2005). Acute respiratory diseases and carboxyhemoglobin status in school children of Quito, Ecuador. *Environmental Health Perspectives*, 113(5):607–11.
- Fahrmeir, L. (1992). Posterior mode estimation by extended Kalman filtering for multivariate dynamic generalized linear models. *Journal of the American Statistical Association*, 87(418):501–9.
- Fahrmeir, L. (1997). Penalized likelihood estimation and iterative Kalman smoothing for non-Gaussian dynamic regression models. *Computational Statistics and Data Analysis*, 24: 295–320.
- Fernandes, M. V. M. (2006). Modelos para processos espaço-temporais inflacionados de zeros. Dissertação de mestrado, Universidade Federal do Rio de Janeiro.
- Ferreira, M. A. R. e Gamerman, D. (2000). Dynamic generalized linear models. Em Dey, D. K., Ghosh, S. K., e Mallick, B. K., editores, *Generalized Linear Models*, pages 57–72. Marcel Dekker Inc.
- Fisher, I. (1937). Note on a short-cut method for calculating distributed lags. *Bulletin de L'Institut International de Statistique*, XXIX(3):323–27.
- Frühwirth-Schnatter, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15:183–202.
- Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall, Londres.
- Gamerman, D. (1998). Markov chain Monte Carlo for dynamic generalized linear models. *Biometrika*, 85:215–27.
- Gelfand, A. E. e Ghosh, S. K. (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika*, 85:1–11.
- Gelfand, A. E. e Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409.
- Gelman, A. e Rubin, D. R. (1992). Inference from iterative simulation using multiple sequences (com discussão). *Statistical Science*, 7:457–511.
- Geman, S. e Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6: 721–41.

- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (com discussão). Em Bernardo, J. M., Berger, J. O., Dawid, A. P., e Smith, A. F. M., editores, *Bayesian Statistics 4*, pages 169–193. Oxford University Press.
- Geweke, J. e Tanizaki, H. (2001). Bayesian estimation of state-space models using the Metropolis-Hastings algorithm within Gibbs sampling. *Computational Statistics and Data Analysis*, 37:151–70.
- Gordon, N. J., Salmond, D. J., e Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings-F*, 140(2):107–13.
- Greene, W. H. (1997). *Econometric Analysis*. Prentice Hall, Upper Saddle River, Nova Jersey.
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press, 3^a edição.
- Harrison, P. J. e Stevens, C. (1976). Bayesian forecasting (com discussão). *Journal of the Royal Statistical Society, série B*, 38:205–47.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, New York.
- Hastie, T. J. e Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall, Londres.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.
- Junger, W. L. (2002). *Imputação de dados faltando em séries temporais multivariadas via algoritmo EM*. Monografia de graduação, Universidade do Estado do Rio de Janeiro.
- Kitagawa, G. (1987). Non-Gaussian state-space modeling of nonstationary time series (com discussão). *Journal of the American Statistical Association*, 82(400):1032–63.
- Klein, L. R. (1958). The estimation of distributed lags. *Econometrica*, 26(4):553–65.
- Knorr-Held, L. (1999). Conditional prior proposals in dynamic models. *Scandinavian Journal of Statistics*, 26:129–44.
- Koyck, L. (1954). *Distributed Lags and Investment Analysis*. North Holland, Amsterdam.
- Leon, A. P., Mendonça, G. A. S., Junger, W. L., e Cruz, J. C. *et al.* (2005). PROGRAMA ARESRIO. Relatório de conclusão: Qualidade do ar e efeitos na saúde da população do município do Rio de Janeiro. Relatório técnico, Ministério do Meio Ambiente, Secretaria de Qualidade Ambiental nos Assentamentos Humanos, Instituto de Medicina Social/UERJ, Centro de Estudos da Saúde do Trabalhador e Ecologia Humana/ENSP-FIOCRUZ.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., e Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(410):1087–92.

- Migon, H. S. (1984). *An Approach to Non-linear Bayesian Forecasting Problems with Applications*. Tese de doutorado, Universidade de Warwick.
- Migon, H. S. e Harrison, P. J. (1985). An application of nonlinear bayesian forecasting to television advertising. Em Bernardo, J., DeGroot, M., Lindley, D., e Smith, A., editores, *Bayesian Statistics*, volume 2, pages 681–96. Elsevier Science Publishers B. V.
- Migon, H. S., Gamerman, D., Lopes, H. F., e Ferreira, M. A. R. (2005). Bayesian dynamic models. Em Dey, D. e Rao, C.R., editores, *Handbook of Statistics*, volume 25, pages 553–88.
- Moreira, M. C. B. (1994). Modelos dinâmicos generalizados: Métodos aproximados. Dissertação de mestrado, Universidade Federal do Rio de Janeiro.
- Nelder, J. A. e Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, série A*, 135(3):370–84.
- Ornelas, C. P. (2004). *Análise da evolução temporal de poluentes sobre órbitos - uma aplicação de modelos semi-paramétricos de defasagens distribuídas com restrição polinomial*. Monografia de graduação, Universidade do Estado do Rio de Janeiro.
- Pole, A. (1988). Transfer response models: a numerical approach. Em Bernardo, J., DeGroot, M., Lindley, D., e Smith, A., editores, *Bayesian Statistics*, volume 3, pages 733–46. Oxford University Press.
- Ravines, R., Migon, H., e Schmidt, A. (2006). Revisiting distributed lag models through a bayesian perspective (a ser publicado). *Applied Stochastic Models in Business and Industry*.
- Schwartz, J. (2000). The distributed lag between air pollution and daily deaths. *Epidemiology*, 11(3):320–6.
- Shephard, N. (1994). Partial non-Gaussian state space. *Biometrika*, 81(1):115–31.
- Shephard, N. e Pitt, M. K. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika*, 84:653–67.
- Singh, A. C. e Roberts, G. R. (1992). State space modelling of cross-classified time series of counts. *International Statistics Review*, 60:321–36.
- Smith, A. F. M. e Gelfand, A. E. (1992). Bayesian statistics without tears: a sampling-resampling perspective. *The American Statistician*, 46(2):84–8.
- Solow, R. M. (1960). On a family of lag distributions. *Econometrica*, 28(2):393–406.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., e Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, série B*, 64:583–639.
- West, M. e Harrison, P. J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, Londres, 2^a edição.

- West, M., Harrison, P. J., e Migon, H. (1985). Dynamic generalized linear model and bayesian forecasting (com discussão). *Journal of the American Statistical Association*, 80(389):73–97.
- Zanobetti, A., Schwartz, J., Samoli, E., Gryparis, A., Touloumi, G., Atkinson, R., Le Tertre, A., Bobros, J., Celko, M., Goren, A., Forsberg, B., Michelozzi, P., Rabczenko, D., Ruiz, E. A., e Katsouyanni, K. (2002). The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement. *Epidemiology*, 13(1):87–93.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)