



UFAM

UM MODELO DINÂMICO PARA SÉRIES TEMPORAIS CONTÍNUAS COM
MASSA EM ZERO

Jhonata da Silva Pereira

Dissertação de Mestrado apresentada ao
Programa de Pós-graduação em Matemática,
da Universidade Federal do Amazonas, como
parte dos requisitos necessários à obtenção do
título de Mestre em Matemática

Orientador: James Dean Oliveira dos Santos
Junior D.Sc.

Manaus
Março de 2017

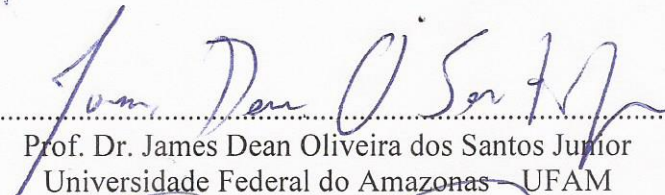
Jhonata da Silva Pereira


UM MODELO DINÂMICO PARA SÉRIES TEMPORAIS CONTÍNUAS COM MASSA EM
ZERO


Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Amazonas, como requisito final para obtenção do título de Mestre em Matemática, área de concentração em Estatística.

Aprovado por:

BANCA EXAMINADORA


Prof. Dr. James Dean Oliveira dos Santos Junior
Universidade Federal do Amazonas – UFAM
Orientador


Prof. Dr. José Mir Justino da Costa
Universidade Federal do Amazonas – UFAM


Prof. Dr. Plínio Lucas Dias Andrade
Universidade Tecnológica Federal do Paraná – UTFPR

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

P436u Pereira, Jhonata da Silva Pereira
Um Modelo Dinâmico para Séries Temporais Contínuas com
Massa em Zero / Jhonata da Silva Pereira Pereira. 2017
77 f.: il.; 31 cm.

Orientador: James Dean Oliveira dos Santos Junior
Dissertação (Mestrado em Matemática - Estatística) -
Universidade Federal do Amazonas.

1. Séries Temporais. 2. Dados Contínuos com Massa em Zero. 3.
mld. 4. mlgd. 5. Filtro de Kalman. I. Santos Junior, James Dean
Oliveira dos II. Universidade Federal do Amazonas III. Título

*À minha família em especial
minha mãe, Silvana, pelo seu
amor e cuidado que tem por mim
até hoje.*

*À minha avó, Ivanide, por ter
me ensinado muito dessa vida e
pela saudades de seus mimos (in
memorian).*

*Ao meu bisavô, "papai Chico",
pelo seu amor paterno nos
primeiros anos de minha vida
(In memorian).*

Agradecimentos

Aqui encerra um grande ciclo da minha vida e dele muitas pessoas fizeram parte, as quais levarei em meus pensamentos, mesmo que não os vejam mais, para o resto da minha vida. A estas pessoas agradeço de todo o coração.

Agradeço aos amigos da UFAM, meu professor e orientador James Dean por ter me confiado esse trabalho e aos amigos da sala da pós graduação por terem feito os dias mais divertidos, seja pelos jogos ou seja pelas brincadeiras e piadas. E desses destacam-se: "Ciço", Danilo, Thiago, Alex, Guilherme kiss e Vinícius.

Agradeço, também, as meninas do mestrado: Alice, Diana, Renata, Sara e "Dona"Regina.

Agradeço à minha família em especial minha mãe, Silvana, minhas irmãs, Karina e Kallyandra, e minha esposa, Mel, por ter me aturado até hoje (que por sinal não é nada fácil) e zelar pela nossa família.

Enfim, agradeço a Deus pelo seu amor infinito e pelo dom da perseverança. Agradeço a Deus pelo Seu melhor presente que me destes, a minha filha Sophia Gabriela, que é minha vida, minha razão de viver e meu tudo. É por ti, minha princesa, que luto para ser uma pessoa melhor a cada dia. És por ti, meu amor, que sou o que sou hoje. És por ti, minha vida, que vou mais além. Obrigado minha filha por amolecer esse coração de pedra, por fazer derramar lágrimas desses olhos secos, obrigado por, apesar de ser tão nova, ensinar-me tantas coisas. Meu amor por você é tão grande que mal consigo expressá-lo, mas saiba que te amo de todo o meu coração e como você fala: "te amo do tamanho do universo".

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Matemática, da Universidade Federal do Amazonas, como parte dos requisitos necessários para a obtenção do grau de Mestre em Matemática. (M.Sc.)

UM MODELO DINÂMICO PARA SÉRIES TEMPORAIS CONTÍNUAS COM
MASSA EM ZERO

Jhonata da Silva Pereira

Março/2017

Orientador: James Dean Oliveira dos Santos Junior D.Sc.

Área de Concentração: Estatística

Apresentamos nesta dissertação um modelo para séries temporais contínuas em que os valores observados são não negativos, mas tem uma proporção de zeros (massa em zero). O modelo é baseado na teoria do Modelo Linear Dinâmico (MLD) e Modelo Linear Generalizado Dinâmico (MLGD) que nos permite fazer inferências dos parâmetros de forma recursiva através do filtro de Kalman. Para avaliar o modelo proposto foi desenvolvido um estudo de simulação e aplicações em séries temporais de precipitação pluviométrica.

Abstract of Dissertation presented to Postgraduate in Mathematics, of the Federal University of Amazonas, as a partial fulfillment of the requirements for the degree of Master of Mathematics. (M.Sc.)

THESIS TITLE

Jhonata da Silva Pereira

March/2017

Advisor: James Dean Oliveira dos Santos Junior D.Sc.

Research lines: Statistics

We present in this dissertation a model for continuous time series in which the observed values are nonnegative but have a proportion of zeros (mass in zero). The model is based on the theory of Dynamic Linear Model (DLM) and Dynamic Generalized Linear Model (DGLM) Which allows us to make inferences of the parameters in a recursive way through the Kalman filter. To evaluate the proposed model was developed a simulation study and applications in time series of pluviometric precipitation.

Sumário

| | |
|--|------------|
| Lista de Figuras | x |
| Lista de Tabelas | xii |
| 1 Introdução | 1 |
| 1.1 Séries Temporais | 1 |
| 1.2 Objetivos da Dissertação | 2 |
| 1.2.1 Objetivo Geral | 2 |
| 1.3 Objetivo Específico | 2 |
| 1.4 Organização do Trabalho | 3 |
| 2 Modelos Dinâmicos | 4 |
| 2.1 Modelo de Espaço de Estado | 4 |
| 2.2 Modelo Linear Dinâmico | 5 |
| 2.2.1 Filtro para o MLD | 7 |
| 2.2.2 Distribuições de Previsão | 9 |
| 2.2.3 Modelo Polinomial de 1ª Ordem | 10 |
| 2.2.4 Modelo Polinomial de 2ª Ordem | 11 |
| 2.2.5 Representação da Sazonalidade | 12 |
| 2.2.6 Modelos com Variância Desconhecida | 14 |
| 2.2.7 Fatores de Desconto | 16 |
| 2.2.8 Superposição de Modelos | 17 |
| 2.3 Modelo Linear Generalizado Dinâmico | 17 |
| 2.3.1 Filtro para o MLGD | 19 |
| 2.3.2 Exemplo: Modelo Dinâmico Bernoulli (MDB) | 22 |
| 2.4 Multiprocessos | 25 |
| 2.5 Métricas de Performance de Previsão | 27 |
| 3 Modelo Dinâmico para Séries Temporais Contínuas com Massa em Zero | 29 |
| 3.1 Representação Estocástica | 29 |
| 3.2 Modelo para séries com dados contínuos e massa em zero | 30 |

| | | |
|----------|---|-----------|
| 3.2.1 | Modelo com V_t Desconhecido | 39 |
| 4 | Análise de Desempenho da Nossa Proposta | 46 |
| 4.1 | Modelo 1 | 47 |
| 4.2 | Modelo 2 | 49 |
| 4.3 | Modelo 3 | 50 |
| 4.4 | Modelo 4 | 52 |
| 4.5 | Modelo 5 | 53 |
| 4.6 | Modelo 6 | 54 |
| 4.7 | Modelo 7 | 57 |
| 4.8 | Discussão | 57 |
| 5 | Aplicação em Séries Reais | 59 |
| 5.1 | Manaus | 59 |
| 5.2 | Tarauacá | 64 |
| 5.3 | São Gonçalo | 67 |
| 6 | Considerações Finais e Trabalhos Futuros | 69 |
| | Referências Bibliográficas | 71 |
| A | Propriedades da Distribuição Normal | 74 |
| A.1 | Transformações lineares: | 74 |
| A.2 | Distribuições Marginais | 74 |
| A.3 | Distribuições Condicionais | 74 |
| A.4 | Reconstrução da Conjunta | 75 |
| B | Modelo com Variância Desconhecida | 76 |
| C | Superposição de Modelos | 77 |

Lista de Figuras

| | | |
|------|---|----|
| 2.1 | Representação gráfica dos harmônicos. | 12 |
| 4.1 | Série com tendências constantes. | 47 |
| 4.2 | Multiprocesso Classe I: Série com tendências constantes. | 48 |
| 4.3 | Série com tendência constante na parte contínua e linear na parte discreta. | 49 |
| 4.4 | Multiprocesso Classe I: Série com tendência constante(contínua) e linear (discreta). | 50 |
| 4.5 | Série com tendência linear na parte contínua e constante na parte discreta. | 51 |
| 4.6 | Multiprocesso Classe I: Série com tendência linear (contínua) e constante (discreta). | 51 |
| 4.7 | Modelo com tendências lineares | 52 |
| 4.8 | Multiprocesso Classe I: Série com tendências lineares. | 53 |
| 4.9 | Série com tendências constante e sazonal na parte contínua e tendência constante na parte discreta | 54 |
| 4.10 | Multiprocesso Classe I: Série com tendência constante "+"sazonal na parte contínua e constante na discreta. | 54 |
| 4.11 | Série com tendência constante na parte contínua e sazonal na parte discreta | 55 |
| 4.12 | Multiprocesso Classe I: Série com tendência constante (contínua) e sazonal (discreta). | 56 |
| 4.13 | Série com tendência crescente na parte contínua e sazonal na discreta | 57 |
| 4.14 | Multiprocesso Classe I: Série com tendência crescente (contínua) e sazonal (discreta) | 58 |
| 5.1 | Série da Precipação de Manaus 1 | 60 |
| 5.2 | Multiprocesso para os quatro modelos ajustados em Manaus 1 | 60 |
| 5.3 | Modelagem da série usando o modelo proposto com $V_t = V$ desconhecido. | 61 |
| 5.4 | Precipação no município de Manaus no período de 2012 a 2015. | 61 |

| | | |
|------|---|----|
| 5.5 | Multiprocesso para os modelos da série de Manaus 2. | 63 |
| 5.6 | Ajuste do modelo à série de Manaus no período de 2012 a 2015. . . . | 63 |
| 5.7 | Ajuste do modelo à série de Manaus por ano. | 64 |
| 5.8 | Série da precipitação pluviométrica de Tarauacá. | 65 |
| 5.9 | Multiprocesso - Tarauacá | 65 |
| 5.10 | Ajuste do modelo à série de Tarauacá no período de 2012 a meado de 2015. | 66 |
| 5.11 | Ajuste do modelo à série de Tarauacá por ano. | 66 |
| 5.12 | Série da precipitação pluviométrica de São Gonçalo. | 67 |
| 5.13 | Multiprocesso para os modelos de São Gonçalo. | 68 |
| 5.14 | Ajuste do modelo à série de São Gonçalo no período de 2012 a meado de 2015 | 68 |

Lista de Tabelas

| | | |
|-----|---|----|
| 4.1 | Medidas de desempenho: Série com tendências constantes | 48 |
| 4.2 | Medidas de desempenho: Série com tendência constante(contínua) e linear (discreta) | 49 |
| 4.3 | Medidas de desempenho: Série com tendência linear (contínua) e constante (discreta). | 50 |
| 4.4 | Medidas de desempenho: Série com tendências lineares. | 52 |
| 4.5 | Medidas de desempenho: Série com tendência constante "+" sazonal na parte contínua e constante na discreta. | 53 |
| 4.6 | Medidas de desempenho: Série com tendência constante (contínua) e sazonal (discreta). | 55 |
| 4.7 | Medidas de desempenho: Série com tendência crescente (contínua) e sazonal (discreta) | 58 |
| 5.1 | Fatores de descontos para a Série de Manaus 1. | 60 |
| 5.2 | Fatores de descontos para a série de Manaus 2. | 62 |
| 5.3 | Fatores de descontos para a Série de Tarauacá. | 65 |
| 5.4 | Fatores de descontos para a Série de São Gonçalo. | 67 |

Capítulo 1

Introdução

1.1 Séries Temporais

Qualquer coleção de observações ordenadas sequencialmente ao longo do tempo é denominada série temporal. Este tipo de estrutura é muito comum em áreas como economia (preços diários de ações, taxa mensal de desemprego, produção industrial), medicina (eletrocardiograma, eletroencefalograma), epidemiologia (número mensal de novos casos de meningite) e meteorologia (precipitação pluviométrica, temperatura diária, velocidade do vento). Neste tipo de dados, as observações vizinhas são estocasticamente dependentes. Em geral, os objetivos de uma análise de série temporal são:

- compreender os mecanismos geradores da série:
 - ⇒ descrever eficientemente o comportamento da série (verificar tendência, componentes sazonais/cíclicos, valores discrepantes e mudanças estruturais).
- prever o comportamento da série:
 - ⇒ construir planos a curto, médio e longo prazo;
 - ⇒ obter direcionamento para tomada de decisões.

As séries temporais são formadas pelo conjunto de informações $\{Y_t, t \in T\}$, onde Y_t é a variável de interesse observada no tempo t e T é o conjunto de tempos, podendo ser classificado como:

- Discreto: $T = \{t_1, t_2, \dots, t_n\}$;
- Contínuo: $T = \{t, t_1 < t < t_2\}$;

Séries compostas de variáveis aleatórias exclusivamente contínuas são denominadas contínuas. Séries compostas de variáveis aleatórias exclusivamente discretas são

denominadas discretas. A série é dita multivariada se as observações são $\{\mathbf{Y}_t, t \in T\}$, onde \mathbf{Y}'_t s são vetores $n \times 1$ de variáveis aleatórias.

Harrison & Stevens [8] e West et al. [26] apresentaram uma ampla classe de modelos para tratar séries temporais, conhecidos como modelos dinâmicos. Modelos dinâmicos são usados principalmente por sua versatilidade, sendo capazes de incorporar variáveis explicativas, acompanhar tendências e efeitos sazonais, além de modelar a variância do processo.

Há aplicações em que variáveis podem tomar qualquer valor não negativo, mas tem uma proporção de resultado zero. Chamamos essas variáveis de semicontínuas (Agresti & Min [1]), ou contínua com massa em zero, quando tem uma distribuição contínua exceto para uma massa de probabilidade em 0. Dados semicontínuos são comuns em muitas áreas. Por exemplo, quando cada observação é um registro da precipitação diária, muitos dias não têm precipitação. Num estudo sobre as despesas das famílias, algumas famílias não gastam nada em determinada mercadoria durante o período de investigação. Em um estudo de custos médicos anuais, uma parcela da população tem despesa médica zero. Nesse trabalho estamos interessados em séries temporais que possuam essa característica.

1.2 Objetivos da Dissertação

1.2.1 Objetivo Geral

Construir um modelo dinâmico que seja capaz de modelar séries temporais contínuas que tenham massa em zero.

1.3 Objetivo Específico

Os objetivos específicos são:

1. apresentar as definições dos modelos dinâmicos com base na revisão bibliográfica;
2. construir o modelo para séries temporais com massa em zero que assumam a forma dos modelos dinâmicos e fazer os devidos processos de inferência;
3. simular cenários e verificar se o modelo construído obtém bons resultados;
4. aplicar o novo modelo em dados reais de precipitação pluviométrica.

1.4 Organização do Trabalho

Esta dissertação está dividida em seis capítulos sendo que no primeiro apresentamos uma introdução e os objetivos deste trabalho.

No Capítulo 2, apresentamos a definição do Modelo Linear Dinâmico, do Modelo Linear Generalizado Dinâmico e os métodos de inferência para estes modelos.

No Capítulo 3, introduzimos o Modelo Dinâmico para Séries Temporais Contínuas com Massa em Zero. Nesse capítulo, descrevemos a estrutura do modelo e apresentamos os passos necessários para inferência dos parâmetros.

No Capítulo 4 apresentamos um estudo do modelo proposto em séries geradas através de cenários computacionais. O objetivo é verificar se o modelo é capaz de descrever o comportamento das séries simuladas.

A aplicação do modelo proposto em dados reais é apresentado no Capítulo 5, utilizando três séries distintas. A primeira é referente a precipitação pluviométrica do município de Manaus, a segunda no de Tarauacá/AC e a terceira série pluviométrica em São Gonçalo, um distrito do município de Souza/PB.

As considerações finais e propostas para trabalhos futuros são apresentados no Capítulo 6.

Capítulo 2

Modelos Dinâmicos

Modelos de Espaço de Estado (MEE) são amplamente utilizados pela sua flexibilidade na análise de séries temporais, uma vez que acomoda uma classe geral de modelos cujos dados possuem natureza e estrutura variada. Tais modelos foram introduzidos por Kalman [12] com o objetivo de estimar os estados de sistemas dinâmicos.

A grande vantagem do MEE é que este nos permite modelar processos univariados ou multivariados, estacionários ou não estacionários, com mudanças estruturais e padrões irregulares sem a preocupação de fazermos transformações nos dados originais.

Neste capítulo é feita uma breve revisão deste modelo, abordando com mais ênfase dois casos particulares: o Modelo Linear Dinâmico e o Modelo Linear Generalizado Dinâmico.

2.1 Modelo de Espaço de Estado

Um MEE consiste de uma série temporal p -dimensional $(\boldsymbol{\theta}_t, t = 0, 1, 2, \dots)$ e uma série temporal Y_t cuja observação é denotada por y_t , para $t = 1, 2, 3, \dots$, onde as seguintes condições são satisfeitas:

1. $(\boldsymbol{\theta}_t, t = 0, 1, 2, \dots)$ é um processo de Markov, de modo que para qualquer $t \geq 1$,

$$f(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{0:t-1}) = f(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1}).$$

Isto significa que a informação sobre $\boldsymbol{\theta}_t$ vinda de $(\boldsymbol{\theta}_0, \dots, \boldsymbol{\theta}_{t-1}) = \boldsymbol{\theta}_{0:t-1}$ é exatamente a mesma informação vinda apenas de $\boldsymbol{\theta}_{t-1}$. Assim, a lei de probabilidade do processo $(\boldsymbol{\theta}_t, t = 0, 1, 2, \dots)$ é especificada atribuindo a densidade inicial de $\boldsymbol{\theta}_0$, $f(\boldsymbol{\theta}_0)$, e as densidades de transição de $\boldsymbol{\theta}_t$ condicionado a $\boldsymbol{\theta}_{t-1}$, $f(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1})$.

2. condicional a $\boldsymbol{\theta}_t$, os Y'_t s são independentes entre si e Y_t depende apenas de $\boldsymbol{\theta}_t$. Portanto, segue que para qualquer $n \geq 1$, $(y_1, \dots, y_n) | (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_n)$ possui densidade condicional conjunta $\prod_{t=1}^n f(y_t | \boldsymbol{\theta}_t)$.

Como consequência das condições listadas acima, o MEE é completamente especificado pela:

1. distribuição inicial a priori dos estados $f(\boldsymbol{\theta}_0)$;
2. densidade condicional da evolução dos estados $f(\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1})$;
3. densidade condicional $f(y_t | \boldsymbol{\theta}_t)$, de forma que a distribuição conjunta das observações e dos estados, para qualquer $t \geq 1$, seja dada por

$$f(\boldsymbol{\theta}_{0:t}, y_{1:t}) = f(\boldsymbol{\theta}_0) \prod_{j=1}^t f(\boldsymbol{\theta}_j | \boldsymbol{\theta}_{j-1}) f(y_j | \boldsymbol{\theta}_j). \quad (2.1)$$

A vantagem dos MEE's é que qualquer modelo linear de séries temporais pode ser expresso na forma de espaço de estados, relacionando as observações Y_t a erros ν_t através do processo de Markov $\boldsymbol{\theta}_t$, denominado vetor de estados latentes do sistema. Assim, a estimação e predição dos estados latentes no tempo t podem ser feitas recursivamente, baseado no conjunto de informações disponíveis no instante t . Em geral, quando trabalhamos com séries temporais ($Y_t, t = 1, 2, \dots$), o MEE pode ser escrito da seguinte forma:

$$Y_t = h_t(\boldsymbol{\theta}_t, \nu_t), \quad (2.2)$$

$$\boldsymbol{\theta}_t = g_t(\boldsymbol{\theta}_{t-1}, \boldsymbol{\omega}_t). \quad (2.3)$$

Onde h_t e g_t são funções reais, ν_t e $\boldsymbol{\omega}_t$ são erros dos processos que assumimos independentes entre e dentre si.

Com o avanço tecnológico nos anos 90, o uso dos MEE's tornou-se ainda mais viável, e a aproximação bayesiana mostrou-se bastante favorável para o cálculo recursivo de tal metodologia. Trabalhos mais recentes para a análise de séries temporais baseados nos MEE's podem ser encontrados em Gomes [7], Liu & West [13] e Petris et al.[18].

2.2 Modelo Linear Dinâmico

Os Modelos Lineares Dinâmicos (MLD's) são um caso especial do MEE, que ocorrem quando h_t e g_t das Equações (2.2) e (2.3), respectivamente, são lineares no primeiro argumento e os erros associados no segundo argumento possuem distribuição normal.

Considere uma s rie temporal univariada $(Y_t, t = 1, 2, \dots)$.

Defini o 1 Para cada  ndice de tempo t , o Modelo Linear Din mico univariado   definido por

$$\text{Equa o de Observa o: } Y_t = \mathbf{F}_t^T \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N[0, V_t], \quad (2.4)$$

$$\text{Equa o de Evolu o: } \boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim N[\mathbf{0}, \mathbf{W}_t], \quad (2.5)$$

$$\text{Informa o Inicial: } \boldsymbol{\theta}_0 | y_0 \sim N[\mathbf{m}_0, \mathbf{C}_0], \quad (2.6)$$

para alguma m dia e vari ncia a priori \mathbf{m}_0 e \mathbf{C}_0 . Assume-se que as sequ ncias de erros observacionais ν_t e de evolu o $\boldsymbol{\omega}_t$ s o independentes ao longo do tempo e entre si, e independentes da priori $\boldsymbol{\theta}_0 | y_0$. Para $t = 1, 2, \dots$,

- $\boldsymbol{\theta}_t$   um vetor p -dimensional denominado par metro de estado ou simplesmente estado do modelo din mico;
- \mathbf{F}_t   um vetor p -dimensional de constantes conhecidas, podendo ser um conjunto de vari veis explicativas;
- \mathbf{G}_t   uma matriz $p \times p$ conhecida que descreve a evolu o temporal dos par metros de estado, conhecida como matriz de evolu o;
- V_t   a vari ncia associada ao erro observacional ν_t ;
- \mathbf{W}_t   uma matriz de covari ncia $p \times p$ conhecida associada ao erro de evolu o (ou erro de estado) $\boldsymbol{\omega}_t$.

A equa o de observa o (2.4) relaciona o vetor de observa es Y_t ao par metro de estado $\boldsymbol{\theta}_t$ onde assumimos que as observa es Y_t s o condicionalmente independentes dado $\boldsymbol{\theta}_t$ e que Y_t depende apenas de $\boldsymbol{\theta}_t$, para $t = 1, 2, \dots$, enquanto que a equa o de evolu o (2.5)   respons vel pela evolu o dos par metros de estado atrav s do tempo. Estas duas equa es podem ser reescritas, para $t = 1, 2, \dots$, como

$$Y_t | \boldsymbol{\theta}_t \sim N[\mathbf{F}_t^T \boldsymbol{\theta}_t, V_t], \quad (2.7)$$

$$\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1} \sim N[\mathbf{G}_t \boldsymbol{\theta}_{t-1}, \mathbf{W}_t]. \quad (2.8)$$

Devido   estrutura markoviana do modelo, $\boldsymbol{\theta}_t$ dado $\boldsymbol{\theta}_{t-1}$, para $t = 1, 2, \dots$, tem uma distribui o normal e dado $\boldsymbol{\theta}_{t-1}$ e os valores de \mathbf{G}_t e \mathbf{W}_t , $\boldsymbol{\theta}_t$   independente da informa o anterior ao tempo t , conforme podemos ver nas Equa es (2.7) e (2.8). Denotamos por $y_{0:t} = (y_0, y_{1:t}) = (y_0, y_1, y_2, \dots, y_t)$ o conjunto de informa es dispon veis at  o instante de tempo t , para $t = 1, 2, \dots$, em que y_0 denota o conjunto de informa es no instante inicial $t = 0$.

O modelo descrito em (2.4) e (2.5) é completamente especificado pela quádrupla $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}_t$ e pela distribuição a *priori* Normal assumida para os parâmetros de estado, dada na Equação (2.6) em que \mathbf{m}_0 e \mathbf{C}_0 são, respectivamente, a média e a variância da distribuição a *priori* Normal para $\boldsymbol{\theta}_0$ que reflete nossa incerteza a respeito do processo sob estudo no instante de origem $t = 0$.

Note que o erro ν_t é simplesmente uma perturbação aleatória no processo de medida das observações Y_t . O erro de evolução $\boldsymbol{\omega}_t$, por outro lado, influencia no desenvolvimento do sistema ao longo do tempo. A suposição de que estes erros são independentes dentre e entre si, claramente separa estas duas fontes de variação estocástica e torna mais nítido o papel que cada uma representa.

2.2.1 Filtro para o MLD

Na análise de séries temporais um dos objetivos é a estimação dos parâmetros do modelo e a previsão. No MLD é possível empregar o filtro de Kalman, que é um procedimento recursivo com a finalidade de fazer inferências sobre a estrutura dos estados e previsão de observações futuras da série, baseado nas informações disponíveis. Podemos sintetizar o processo do filtro de Kalman da seguinte maneira:

- em um tempo $t - 1$ dispomos da distribuição a *posteriori* $f(\boldsymbol{\theta}_{t-1}|y_{0:t-1})$;
- dado que conhecemos $y_{0:t-1}$, podemos resumir a informação contida neste conjunto através da distribuição a *priori* $f(\boldsymbol{\theta}_t|y_{0:t-1})$;
- baseado numa estimativa de $\boldsymbol{\theta}_t$, calculamos a distribuição de previsão um passo a frente $f(y_t|y_{0:t-1})$;
- uma vez que y_t foi observado, o sistema é atualizado com a entrada de mais informações e calculamos a distribuição a *posteriori* no tempo t , $f(\boldsymbol{\theta}_t|y_{0:t})$.

Para cada tempo t , dadas as distribuições de $y_t|\boldsymbol{\theta}_t$, $\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}$ e $\boldsymbol{\theta}_{t-1}|y_{0:t-1}$, e assumindo $\boldsymbol{\theta}_t$, quando condicionado a $\boldsymbol{\theta}_{t-1}$, independente de $y_{0:t-1}$, a distribuição a *priori*, de previsão um passo a frente e *posteriori* são, respectivamente:

$$\begin{aligned}
f(\boldsymbol{\theta}_t|y_{0:t-1}) &= \int f(\boldsymbol{\theta}_t, \boldsymbol{\theta}_{t-1}|y_{0:t-1})d\boldsymbol{\theta}_{t-1} = \int f(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, y_{0:t-1})f(\boldsymbol{\theta}_{t-1}|y_{0:t-1})d\boldsymbol{\theta}_{t-1} \\
&= \int f(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1})f(\boldsymbol{\theta}_{t-1}|y_{0:t-1})d\boldsymbol{\theta}_{t-1}; \tag{2.9}
\end{aligned}$$

$$\begin{aligned}
f(y_t|y_{0:t-1}) &= \int f(y_t, \boldsymbol{\theta}_t|y_{0:t-1})d\boldsymbol{\theta}_t = \int f(y_t|\boldsymbol{\theta}_t, y_{0:t-1})f(\boldsymbol{\theta}_t|y_{0:t-1})d\boldsymbol{\theta}_t \\
&= \int f(y_t|\boldsymbol{\theta}_t)f(\boldsymbol{\theta}_t|y_{0:t-1})d\boldsymbol{\theta}_t; \tag{2.10}
\end{aligned}$$

$$\tag{2.11}$$

$$\begin{aligned}
f(\boldsymbol{\theta}_t|y_{0:t}) &= \frac{f(\boldsymbol{\theta}_t|y_{0:t-1})f(y_t|\boldsymbol{\theta}_t, y_{0:t-1})}{f(y_t|y_{0:t-1})} \\
&\propto f(\boldsymbol{\theta}_t|y_{0:t-1})f(y_t|\boldsymbol{\theta}_t), \tag{2.12}
\end{aligned}$$

sendo a Equação (2.12) obtida via Teorema de Bayes.

Teorema 1 *No MLD, a previsão um passo a frente e as distribuições a posteriori, para todo t , são dadas por:*

(a) *Posteriori no tempo $t - 1$:*

Para alguma média \mathbf{m}_{t-1} e variância \mathbf{C}_{t-1} ,

$$\boldsymbol{\theta}_{t-1}|y_{0:t-1} \sim N[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}].$$

(b) *Priori no tempo t :*

$$\boldsymbol{\theta}_t|y_{0:t-1} \sim N[\mathbf{a}_t, \mathbf{R}_t],$$

onde,

$$\begin{aligned}
\mathbf{a}_t &= \mathbf{G}_t\mathbf{m}_{t-1}; \\
\mathbf{R}_t &= \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t^T + \mathbf{W}_t.
\end{aligned}$$

(c) *Previsão um passo a frente:*

$$Y_t|y_{0:t-1} \sim N[f_t, Q_t],$$

onde,

$$\begin{aligned}
f_t &= \mathbf{F}_t^T\mathbf{a}_t; \\
Q_t &= \mathbf{F}_t^T\mathbf{R}_t\mathbf{F}_t + V_t.
\end{aligned}$$

(d) *Posteriori no tempo t :*

$$\boldsymbol{\theta}_t | y_{0:t} \sim N[\mathbf{m}_t, \mathbf{C}_t],$$

onde,

$$\begin{aligned} \mathbf{m}_t &= \mathbf{a}_t + \mathbf{A}_t e_t; \\ \mathbf{C}_t &= \mathbf{R}_t - \mathbf{A}_t \mathbf{Q}_t \mathbf{A}_t^T; \end{aligned}$$

com,

$$\begin{aligned} \mathbf{A}_t &= \mathbf{R}_t \mathbf{F}_t / \mathbf{Q}_t; \\ e_t &= y_t - f_t. \end{aligned}$$

Demonstração: A demonstração é feita por indução usando as propriedades da normal multivariada do Apêndice A (vide West & Harrison [25]).

2.2.2 Distribuições de Previsão

Para fazer previsão da série dado que temos a informação até o instante t , ou seja, fazer inferência sobre y_{t+k} , para algum $k \geq 1$, condicionado ao conhecimento da série $y_{1:t}$, precisamos da distribuição de $y_{t+k} | y_{1:t}$ que é denominada função de previsão k passos à frente.

Definição 2 Para qualquer tempo corrente t , a função de previsão, $f_t(k)$, é definida para todos os inteiros $k \geq 0$, como

$$f_t(k) = E[\mu_{t+k} | y_{0:t}] = E[\mathbf{F}_{t+k}^T \boldsymbol{\theta}_{t+k} | y_{0:t}], \quad (2.13)$$

onde,

$$\mu_{t+k} = \mathbf{F}_{t+k}^T \boldsymbol{\theta}_{t+k}, \quad (2.14)$$

é a função de resposta média.

Para k estritamente maior que 0, a função de previsão fornece valores esperados de futuras observações dada a informação atual,

$$f_t(k) = E[Y_{t+k} | y_{0:t}], \quad \text{para } k \geq 1. \quad (2.15)$$

Os seguintes resultados fornecem as distribuições de previsões.

Teorema 2 Para cada tempo t e $k \geq 1$, as distribuições para $\boldsymbol{\theta}_{t+k}$ e Y_{t+k} dado $y_{0:t}$, k passos a frente, são dadas por:

a) distribuição de estado:

$$\boldsymbol{\theta}_{t+k}|y_{0:t} \sim N[\mathbf{a}_t(k), \mathbf{R}_t(k)], \quad (2.16)$$

b) distribuição de Previsão:

$$Y_{t+k}|y_{0:t} \sim N[\mathbf{f}_t(k), Q_t(k)], \quad (2.17)$$

com momentos definidos recursivamente por:

$$f_t(k) = \mathbf{F}_{t+k}^T \mathbf{a}_t(k) \quad e \quad Q_t(k) = \mathbf{F}_{t+k}^T \mathbf{R}_t(k) \mathbf{F}_{t+k} + V_{t+k},$$

onde

$$\mathbf{a}_t(k) = \mathbf{G}_{t+k} \mathbf{a}_t(k-1) \quad e \quad \mathbf{R}_t(k) = \mathbf{G}_{t+k} \mathbf{R}_t(k-1) \mathbf{G}_{t+k}^T + \mathbf{W}_{t+k},$$

com valores iniciais

$$\mathbf{a}_t(0) = \mathbf{m}_t \quad e \quad \mathbf{R}_t(0) = \mathbf{C}_t.$$

A demonstração pode ser vista em West & Harrison [25] e Petris [18].

Colorário 1 No caso especial em que a matriz de evolução \mathbf{G}_t é constante, $\mathbf{G}_t = \mathbf{G}$ para todo t , então, para $k \geq 0$,

$$\mathbf{a}_t(k) = \mathbf{G}^k \mathbf{m}_t, \quad (2.18)$$

tal que

$$f_t(k) = \mathbf{F}_{t+k}^T \mathbf{G}^k \mathbf{m}_t. \quad (2.19)$$

Além disso, se $\mathbf{F}_t = \mathbf{F}$ para todo t , o modelo é chamado de MLD de séries temporais (MLDST), e a função de previsão tem a forma

$$f_t(k) = \mathbf{F}^T \mathbf{G}^k \mathbf{m}_t. \quad (2.20)$$

2.2.3 Modelo Polinomial de 1ª Ordem

O mais simples MLD é o modelo polinomial de 1ª ordem, também chamado de passeio aleatório. Ele é composto apenas de um nível que varia segundo um passeio aleatório. Este modelo é definido pela quádrupla

$$\{1, 1, V, W\}_t,$$

Definição 3 Para cada tempo t , o Modelo Linear Dinâmico de primeira ordem é definido por:

$$\begin{aligned} \text{Equação de Observação:} & \quad Y_t = \theta_t + \nu_t, & \nu_t & \sim N[0, V_t], \\ \text{Equação de evolução:} & \quad \theta_t = \theta_{t-1} + \omega_t, & \omega_t & \sim N[0, W_t], \\ \text{Informação Inicial:} & \quad \theta_0 | y_0 \sim N[m_0, C_0], \end{aligned}$$

para alguma média e variância a priori m_0 e C_0 . Assume-se que as sequências de erros observacionais ν_t e de evolução ω_t são independentes ao longo do tempo e entre si, e independentes da priori $\theta_0 | y_0$.

O MLD polinomial de 1ª ordem é denominado constante se $V_t = V$ e $W_t = W$ para todo t . A função de previsão para este modelo é

$$f_t(k) = m_t. \quad (2.21)$$

Como $f_t(k) = m_t$ não depende de k , este modelo não é interessante para previsões em longo prazo.

2.2.4 Modelo Polinomial de 2ª Ordem

O modelo polinomial de 2ª ordem ou modelo de tendência linear é usado para descrever séries que apresentam tendências lineares, sendo caracterizado pela quádrupla:

$$\left\{ \left(\begin{array}{c} 1 \\ 0 \end{array} \right), \left(\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array} \right), V, \left(\begin{array}{cc} W_{\theta_1} & 0 \\ 0 & W_{\theta_2} \end{array} \right) \right\}_t. \quad (2.22)$$

Definição 4 Para cada tempo t , o Modelo Linear Dinâmico de 2ª ordem é definido por:

$$\begin{aligned} \text{Equação de Observação:} & \quad Y_t = \theta_{1,t} + \nu_t, & \nu_t & \sim N[0, V_t], \\ \text{Equação de Evolução:} & \quad \theta_{1,t} = \theta_{1,t-1} + \theta_{2,t-1} + \omega_{1,t}, & \omega_{1,t} & \sim N[0, W_{\theta_{1,t}}], \\ & \quad \theta_{2,t} = \theta_{2,t-1} + \omega_{2,t}, & \omega_{2,t} & \sim N[0, W_{\theta_{2,t}}], \\ \text{Informação Inicial:} & \quad \boldsymbol{\theta}_0 | y_0 \sim N[\mathbf{m}_0, \mathbf{C}_0], \end{aligned}$$

para alguma média e variância a priori \mathbf{m}_0 e \mathbf{C}_0 . Assume-se que as sequências de erros observacionais ν_t e de evolução $\boldsymbol{\omega}_t$ são independentes ao longo do tempo e

entre si, e independentes da priori $\boldsymbol{\theta}_0|y_0$. Com

$$\boldsymbol{\theta}_t = \begin{pmatrix} \theta_{1,t} \\ \theta_{2,t} \end{pmatrix}, \boldsymbol{\omega}_t = \begin{pmatrix} \omega_{1,t} \\ \omega_{2,t} \end{pmatrix}. \quad (2.23)$$

$\theta_{1,t}$ é o nível da série e $\theta_{2,t}$ representa o crescimento incremental (inclinação do nível), mas $\theta_{2,t}$ pode ser negativo caracterizando um decrescimento. O nível permanece localmente linear, mas a forma da reta pode variar com o tempo. A função de previsão para esse modelo é

$$\begin{aligned} f_t(k) &= \mathbf{F}_t^T \mathbf{G}^k \mathbf{m}_t \\ &= m_{1,t} + km_{2,t}. \end{aligned} \quad (2.24)$$

2.2.5 Representação da Sazonalidade

A sazonalidade de p períodos de uma série temporal pode ser modelada através de combinações lineares de harmônicos, ou seja, por meio de funções periódicas. Essa forma de modelar a sazonalidade é denominada de representação de coeficientes de Fourier e utiliza funções trigonométricas como seno e cosseno.

Considere o seguinte componente, denominado r -ésimo harmônico, definido por

$$\phi_r(t) = A_r \cos\left(\frac{2r\pi}{p}t + \psi_r\right), \quad (2.25)$$

onde A_r é denominado amplitude e ψ_r é a fase do harmônico de ordem r , $r = 1, 2, \dots$. A função $\phi_r(t)$ dá uma volta completa em p/r unidade de tempo.

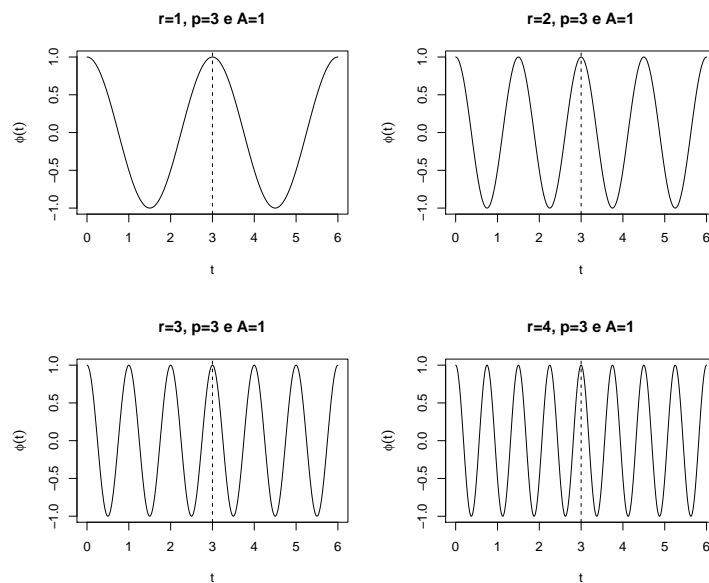


Figura 2.1: Representação gráfica dos harmônicos.

Na Figura 2.1 está a representação gráfica dos harmônicos para $r = 1, 2, 3, 4$ com $p = 3$, $\psi_r = 0$ e $A_r = 1$. Observa-se que o harmônico 1 completa uma volta em 3 unidades de tempo, o harmônico 2 completa uma volta em 3/2 unidades de tempo e duas voltas em $p=3$.

A Equação (2.25) pode ser escrita como soma de seno e cosseno

$$\phi_r(t) = a_r \cos\left(\frac{2r\pi}{p}\right) + b_r \text{sen}\left(\frac{2r\pi}{p}\right), \quad (2.26)$$

onde a_r e b_r são denominados coeficientes de Fourier (West & Harrison [25]). Um resultado importante dessa nova representação é que diferente da Equação (2.25), esta é linear nos termos desconhecidos. Assim, os fatores sazonais $\theta_1, \theta_2, \dots, \theta_p$ podem ser representados por combinações lineares de funções trigonométricas, pois podemos escrever θ_j , $j = 1, 2, \dots, p$, como

$$\theta_j = \begin{cases} a_0 + \sum_{r=1}^{q-1} [a_r \cos(wrt) + b_r \text{sen}(wrt)], & \text{se } p \text{ é ímpar, } q = (p+1)/2; \\ a_0 + \sum_{r=1}^{q-1} [a_r \cos(wrt) + b_r \text{sen}(wrt)] + a_q \cos(\pi t), & \text{se } p \text{ par, } q = p/2; \end{cases} \quad (2.27)$$

onde o ciclo t corresponde ao j -ésimo período no ciclo sazonal $w = 2\pi/p$. A Equação (2.27) é conhecida como representação de Fourier (West & Harrison [25], Petris et al [18]).

A representação trigonométrica do MLD é completada definindo os componentes:

- $\theta_{r,t} = a_r \cos(wrt) + b_r \text{sen}(wrt)$;
- $\gamma_{r,t} = -a_r \text{sen}(wrt) + b_r \cos(wrt)$, $r = 1, \dots, q-1$;
- $\theta_t^T = \begin{cases} (a_0 \theta_{1,t} \gamma_{1,t} \theta_{2,t} \gamma_{2,t} \cdots \theta_{q-1,t} \gamma_{q-1,t}), & \text{se } p \text{ é ímpar, } q = (p+1)/2; \\ (a_0 \theta_{1,t} \gamma_{1,t} \theta_{2,t} \gamma_{2,t} \cdots \theta_{q-1,t} \gamma_{q-1,t} a_q \cos(\pi t)), & \text{se } p \text{ é par, } q = p/2; \end{cases}$
- $F_t^T = \begin{cases} (1 \ 1 \ 0 \ 1 \ 0 \ \cdots \ 1 \ 0), & \text{se } p \text{ é ímpar, } q = (p+1)/2; \\ (1 \ 1 \ 0 \ 1 \ 0 \ \cdots \ 1 \ 0 \ 1), & \text{se } p \text{ é par, } q = p/2; \end{cases}$.

dadas as definições acima, podemos reescrever (2.27), como

$$\theta_j = F_t^T \theta_t, \quad j = 1, 2, \dots, p,$$

e

$$\theta_{t+1} = G\theta_t,$$

onde,

$$\mathbf{G} = \begin{cases} \text{diag}(1, \mathbf{G}_1, \dots, \mathbf{G}_{q-1}), & \text{se } p \text{ é ímpar, } q = (p+1)/2; \\ \text{diag}(1, \mathbf{G}_1, \dots, \mathbf{G}_{q-1}, 1), & \text{se } p \text{ é par, } q = p/2; \end{cases},$$

com

$$\mathbf{G}_r = \begin{pmatrix} \cos(wr) & \text{sen}(wr) \\ -\text{sen}(wr) & \cos(wr) \end{pmatrix}, \quad r = 1, 2, \dots, q-1.$$

A representação é completada através da incorporação dos erros de observação e de evolução levando a seguinte definição.

Definição 5 Para cada tempo t , a representação trigonométrica da sazonalidade é definida por:

$$\begin{aligned} \text{Equação de Observação:} \quad & Y_t = \mathbf{F}_t^T \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim N[0, V_t], \\ \text{Equação do Sistema:} \quad & \boldsymbol{\theta}_{t+1} = \mathbf{G}\boldsymbol{\theta}_t + \boldsymbol{\omega}_{t+1}, \quad \boldsymbol{\omega}_{t+1} \sim N[\mathbf{0}, \mathbf{W}_{t+1}], \\ \text{Informação Inicial:} \quad & \boldsymbol{\theta}_0 | y_0 \sim N[\mathbf{m}_0, \mathbf{C}_0], \end{aligned}$$

para alguma média e variância a priori \mathbf{m}_0 e \mathbf{C}_0 . Assume-se que as sequências de erros observacionais ν_t e de evolução $\boldsymbol{\omega}_t$ são independentes ao longo do tempo e entre si, e independentes da priori $\boldsymbol{\theta}_0 | y_0$. Onde

$$\mathbf{G} = \begin{cases} \text{diag}(1, \mathbf{G}_1, \dots, \mathbf{G}_{q-1}), & \text{se } p \text{ é ímpar, } q = (p+1)/2; \\ \text{diag}(1, \mathbf{G}_1, \dots, \mathbf{G}_{q-1}, 1), & \text{se } p \text{ é par, } q = p/2; \end{cases},$$

e

$$\mathbf{F}_t^T = \begin{cases} (1 \ 1 \ 0 \ 1 \ 0 \ \dots \ 1 \ 0), & \text{se } p \text{ é ímpar, } q = (p+1)/2; \\ (1 \ 1 \ 0 \ 1 \ 0 \ \dots \ 1 \ 0 \ 1), & \text{se } p \text{ é par, } q = p/2. \end{cases}$$

Para este modelo a função de previsão é dada por:

$$f_t(k) = \sum_{r=1}^q [m_{t,2r-1} \cos(wrk) + m_{t,2r} \text{sen}(wrk)], \quad (2.28)$$

onde, $\mathbf{m}_t = E[\boldsymbol{\theta}_t | Y_{0:t}] = (m_{t,1}, m_{t,2}, \dots, m_{t,p-1}, 0)$.

2.2.6 Modelos com Variância Desconhecida

Até agora, foi assumido em todos os modelos que a variância das observações, dada por V_t , era conhecida para todos os tempos, algo pouco comum na prática. Considerando V_t constante em t , podemos estimar a variância seguindo a mesma estrutura do MLD

Definição 6 Para cada tempo t , definindo $\phi = V^{-1}$, o DLM é definido por

$$\begin{aligned} \text{Equação de Observação:} \quad Y_t &= \mathbf{F}_t^T \boldsymbol{\theta}_t + \nu_t, & \nu_t &\sim N[0, V], \\ \text{Equação de Evolução:} \quad \boldsymbol{\theta}_t &= \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, & \boldsymbol{\omega}_t &\sim N[\mathbf{0}, V\mathbf{W}_t^*], \\ \text{Informação Inicial:} \quad \boldsymbol{\theta}_0 | y_0, \phi &\sim N[\mathbf{m}_0, V\mathbf{C}_0^*], \\ \phi | y_0 &\sim \text{Gama} \left[\frac{n_0}{2}, \frac{d_0}{2} \right], \end{aligned}$$

para alguma média e variância a priori \mathbf{m}_0 e \mathbf{C}_0 , e para alguns valores iniciais n_0 e d_0 . Assume-se que as sequências de erros observacionais ν_t e de evolução $\boldsymbol{\omega}_t$ são independentes ao longo do tempo e entre si, e independentes da priori $\boldsymbol{\theta}_0 | y_0, \phi$.

As hipóteses usuais de independência permanecem, porém agora condicionais a V . A priori da precisão observacional, ϕ , tem média $E[\phi | y_{0:t}] = n_0/d_0 = 1/S_0$, onde S_0 é uma estimativa a priori da variância observacional V .

Teorema 3 Com o modelo especificado como na Definição 6, obtemos as seguintes distribuições para cada tempo t .

a) Condicional a V :

$$\begin{aligned} \boldsymbol{\theta}_{t-1} | y_{0:t-1}, V &\sim N[\mathbf{m}_{t-1}, V\mathbf{C}_{t-1}^*], \\ \boldsymbol{\theta}_t | y_{0:t-1}, V &\sim N[\mathbf{a}_t, V\mathbf{R}_t^*], \\ Y_t | y_{0:t-1}, V &\sim N[f_t, VQ_t^*], \\ \boldsymbol{\theta}_t | y_{0:t}, V &\sim N[\mathbf{m}_t, V\mathbf{C}_t^*], \end{aligned}$$

com

$$\begin{aligned} \mathbf{a}_t &= \mathbf{G}_t \mathbf{m}_{t-1}, & \mathbf{R}_t^* &= \mathbf{G}_t \mathbf{C}_{t-1}^* \mathbf{G}_t^T + \mathbf{W}_t^* \\ f_t &= \mathbf{F}_t^T \mathbf{a}_t, & Q_t^* &= \mathbf{F}_t^T \mathbf{R}_t^* \mathbf{F}_t + 1 \\ \mathbf{m}_t &= \mathbf{a}_t + \mathbf{A}_t e_t, & \mathbf{C}_t^* &= \mathbf{R}_t^* - \mathbf{A}_t Q_t^* \mathbf{A}_t^T \\ \mathbf{A}_t &= \mathbf{R}_t^* \mathbf{F}_t Q_t^{*-1}, & e_t &= Y_t - f_t. \end{aligned}$$

b) Para a precisão $\phi = V^{-1}$:

$$\begin{aligned} \phi | y_{0:t-1} &\sim \text{Gama} \left[\frac{n_{t-1}}{2}, \frac{d_{t-1}}{2} \right], \\ \phi | y_{0:t} &\sim \text{Gama} \left[\frac{n_t}{2}, \frac{d_t}{2} \right], \end{aligned}$$

onde

$$n_t = n_{t-1} + 1 \quad e \quad d_t = d_{t-1} + e_t^2 Q_t^{*-1}.$$

c) Incondicional a V :

$$\begin{aligned}\boldsymbol{\theta}_{t-1}|y_{0:t-1} &\sim t_{n_{t-1}}[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}], \\ \boldsymbol{\theta}_t|y_{0:t-1} &\sim t_{n_{t-1}}[\mathbf{a}_t, \mathbf{R}_t], \\ Y_t|y_{0:t-1} &\sim t_{n_{t-1}}[f_t, Q_t], \\ \boldsymbol{\theta}_t|y_{0:t} &\sim t_{n_t}[\mathbf{m}_t, \mathbf{C}_t].\end{aligned}$$

onde

$$\begin{aligned}\mathbf{C}_{t-1} &= S_{t-1}\mathbf{C}_{t-1}^*, \quad \mathbf{R}_t = S_{t-1}\mathbf{R}_t^*, \quad Q_t = S_{t-1}Q_t^* \\ \mathbf{C}_t &= S_t\mathbf{C}_t^*, \quad S_{t-1} = d_{t-1}/n_{t-1} \quad e \quad S_t = d_t/n_t.\end{aligned}$$

Do Teorema 3 podemos obter as seguintes relações:

$$\begin{aligned}\mathbf{m}_t &= \mathbf{a}_t + \mathbf{A}_t e_t, & \mathbf{C}_t &= S_t/S_{t-1}[\mathbf{R}_t - \mathbf{A}_t Q_t \mathbf{A}_t^T], \\ n_t &= n_{t-1} + 1, & d_t &= d_{t-1} + S_{t-1}e_t^2 Q_t^{-1}, \\ Q_t &= S_{t-1} + \mathbf{F}_t^T \mathbf{R}_t \mathbf{F}_t, & \mathbf{A}_t &= \mathbf{R}_t \mathbf{F}_t Q_t^{-1}.\end{aligned}$$

Para a demonstração vide Apêndice B e West & Harrison [25].

2.2.7 Fatores de Desconto

A especificação da magnitude de \mathbf{W}_t é importante para o sucesso da modelagem e da previsão. Seus valores controlam a extensão da variação estocástica na evolução do modelo e, assim, determinam a estabilidade ao longo do tempo. Na equação de evolução, \mathbf{W}_t leva a um aumento da incerteza, ou equivalentemente a uma perda de informação, sobre o vetor de estado entre os tempos $t-1$ e t . Considerando um MLD como em (2.4) e (2.5) a especificação de \mathbf{W}_t pode ser feita indiretamente através do uso de descontos. Considere o filtro do MLD, onde $\mathbf{R}_t = \text{Var}(\boldsymbol{\theta}_t|Y_{0:t-1}) = \mathbf{P}_t + \mathbf{W}_t$ e $\mathbf{P}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t^T$. Daí, $\mathbf{W}_t = \mathbf{R}_t - \mathbf{P}_t$, \mathbf{P}_t seria a variância caso não houvesse incerteza sobre a evolução dos estados, isto é, $\mathbf{W}_t = \mathbf{0}$. Se definirmos δ de tal forma que $\mathbf{R}_t = \mathbf{P}_t/\delta$ podemos interpretar δ como a percentagem de informação que passa de $t-1$ para t e

$$\mathbf{W}_t = \frac{1-\delta}{\delta} \mathbf{P}_t. \quad (2.29)$$

Neste caso, δ é denominado fator de desconto e $\delta \in (0, 1]$. Se $\delta = 1$ então $\mathbf{W}_t = \mathbf{0}$ e não há nenhuma perda de informação na evolução do estado $\boldsymbol{\theta}_{t-1}$ a $\boldsymbol{\theta}_t$, de forma que $\text{Var}(\boldsymbol{\theta}_t|Y_{0:t-1}) = \text{Var}(\mathbf{G}_t \boldsymbol{\theta}_{t-1}|Y_{0:t-1}) = \mathbf{P}_t$. Quando $\delta < 1$, por exemplo, $\delta = 0,4$, $\text{Var}(\boldsymbol{\theta}_t|Y_{0:t-1}) = (1/\delta)\mathbf{P}_t = (1/0,4)\text{Var}(\mathbf{G}_t \boldsymbol{\theta}_{t-1}|Y_{0:t-1}) = 2,5\mathbf{P}_t$, apresentando muita incerteza. Na prática, o valor do fator de desconto é frequentemente fixado, variando entre 0,8 a 0,99.

2.2.8 Superposição de Modelos

A estrutura linear dos modelos permite que sejam combinadas várias componentes em um único modelo. O exemplo mais comum é a superposição de uma tendência polinomial com uma parte sazonal.

Teorema 4 *Para um inteiro $h \geq 2$, considere h séries temporais Y_{it} geradas pelos MLD $M_i : \{\mathbf{F}_i, \mathbf{G}_i, V_i, \mathbf{W}_i\}_t$, com vetores de estado $\boldsymbol{\theta}_{it}$ de dimensão n_i , para $i = 1, 2, \dots, h$. Denote os erros de observação e evolução em M_i por ν_{it} e $\boldsymbol{\omega}_{it}$, respectivamente. Assuma que, $\forall i \neq j, (1 \leq i, j \leq h)$, ν_{it} e $\boldsymbol{\omega}_{it}$ são mutuamente independentes das séries ν_{jt} e $\boldsymbol{\omega}_{jt}$. Então a série definida por :*

$$Y_t = \sum_{i=1}^h Y_{it},$$

segue um MLD $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}_t$ com vetor de estado dado por $\boldsymbol{\theta}_t^T = (\boldsymbol{\theta}_{1t}^T, \dots, \boldsymbol{\theta}_{ht}^T)$, de dimensão $n = n_1 + n_2 + \dots + n_h$ e quádrupla definida por:

$$\begin{aligned} \mathbf{F}_t^T &= (\mathbf{F}_{1t}^T, \dots, \mathbf{F}_{ht}^T); \\ \mathbf{G}_t &= \text{diag}[\mathbf{G}_{1t}, \dots, \mathbf{G}_{ht}]; \\ V_t &= \sum_{i=1}^h V_{it}; \\ \mathbf{W}_t &= \text{diag}[\mathbf{W}_{1t}, \dots, \mathbf{W}_{ht}]. \end{aligned}$$

Para demonstração vide Apêndice C e West & Harrison [25].

2.3 Modelo Linear Generalizado Dinâmico

West et al. [26] formalizaram uma extensão dos MLD's baseado no Modelo Linear Generalizado de Nelder & Wedderburn [17], fazendo possível a utilização destes modelos para distribuições na família exponencial.

A classe dos modelos lineares generalizados dinâmicos (MLGD's) são extensões dos MLD's apresentados na Subseção 2.2, mas sem a suposição de normalidade no vetor de observações e sem a suposição de normalidade na evolução dos parâmetros de estado. Assumimos que a série temporal Y_t segue uma distribuição pertencente à família exponencial (FE) com densidade, ou função de probabilidade, escrita como

$$f(Y_t | \eta_t) = \exp \{ V_t^{-1} [Y_t \eta_t - a(\eta_t)] \} b(Y_t, V_t) \quad (2.30)$$

onde $a(\cdot)$ e $b(\cdot, \cdot)$ são funções conhecidas, η_t é o parâmetro natural da distribuição,

satisfazendo

$$\mu_t = E(Y_t|\eta_t) = a'(\eta_t), \quad (2.31)$$

e $V_t > 0$ é um parâmetro de escala, conhecido para todo t . O parâmetro de precisão da distribuição é definido como $\phi = V_t^{-1}$, satisfazendo

$$\text{Var}(Y_t|\eta_t) = \frac{a''(\eta_t)}{\phi_t}, \quad (2.32)$$

em que $a'(\cdot)$ e $a''(\cdot)$ são a primeira e segunda derivadas da função $a(\cdot)$, respectivamente, e μ_t é conhecido como função média.

A notação $X \sim [\alpha, \beta]$ será usada para denotar que X possui distribuição não especificada com média e variância α e β , respectivamente.

Definição 7 *Definindo as seguintes quantidades no tempo t :*

- $\boldsymbol{\theta}_t$ é um vetor p -dimensional denominado parâmetro de estado;
- \mathbf{F}_t é um vetor p -dimensional de constantes conhecidas;
- \mathbf{G}_t é a matriz de evolução;
- $\boldsymbol{\omega}_t$ é o vetor n dimensional do erro de evolução que tem média zero e matriz de variância conhecida \mathbf{W}_t , denotado por $\boldsymbol{\omega} \sim [\mathbf{0}, \mathbf{W}_t]$;
- λ_t é o preditor linear dos parâmetros de estado;
- $g(\eta_t)$ é uma função contínua e monótona, conhecida, tendo o conjunto dos números reais como imagem.

Então, para cada tempo t , o Modelo Linear Generalizado Dinâmico é definido por:

$$\text{Equação de Observação:} \quad f(Y_t|\eta_t) \in FE, \quad g(\eta_t) = \lambda_t = \mathbf{F}_t^T \boldsymbol{\theta}_t, \quad (2.33)$$

$$\text{Equação de Evolução:} \quad \boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim [\mathbf{0}, \mathbf{W}_t], \quad (2.34)$$

$$\text{Informação Inicial:} \quad \boldsymbol{\theta}_0|y_0 \sim [\mathbf{m}_0, \mathbf{C}_0], \quad (2.35)$$

para os momentos \mathbf{m}_0 e \mathbf{C}_0 . Assume-se que as sequências de erros de evolução $\boldsymbol{\omega}_t$ são independentes entre si e independentes da priori $\boldsymbol{\theta}_0|y_0$.

Note que a equação da observação, descrita anteriormente em (2.4) para um MLD, foi substituída pelo par de equações em (2.33), em que $g(\cdot)$ é uma função de

ligação conhecida que relaciona o parâmetro natural da distribuição η_t ao parâmetro de estado $\boldsymbol{\theta}_t$. Note também que a equação de evolução dos estados (2.34) permanece igual à equação de evolução dos estados de um MLD, descrita em (2.5), a não ser pelo fato de que agora os erros de evolução $\boldsymbol{\omega}_t$ não seguem necessariamente uma distribuição normal.

A distribuição *a priori* para os parâmetros de estado no MLGD é especificada de forma semelhante ao MLD, porém definindo apenas o primeiro e segundo momentos de uma distribuição que agora não é necessariamente normal, como na Equação (2.35), onde y_0 é a informação inicial no instante $t = 0$.

2.3.1 Filtro para o MLGD

O procedimento de inferência sobre os parâmetros de estado nesta classe de modelos, devido a complexidade, pode somente ser feito com aproximações da forma analítica ou computacionalmente (Marco & Gamerman [15]). Se todas as quantidades envolvidas no modelo, com exceção dos parâmetros de estado, são conhecidas, algoritmos sequenciais baseados em estimadores *Linear Bayes* (Hartigan [9]) podem ser utilizados na estimação destes parâmetros.

A sequência de procedimentos utilizada na estimação do modelo é sumarizada a seguir:

Pressuposto: *Posteriori* no tempo $t - 1$.

Assumimos que

$$\boldsymbol{\theta}_t | \boldsymbol{\theta}_{t-1} \sim [\mathbf{m}_{t-1}, \mathbf{C}_{t-1}], \quad (2.36)$$

a partir da equação de evolução, os momentos *a priori* de $\boldsymbol{\theta}_t$, são:

$$\boldsymbol{\theta}_t | y_{0:t-1} \sim [\mathbf{a}_t, \mathbf{R}_t], \quad (2.37)$$

onde,

$$\mathbf{a}_t = \mathbf{G}_t \mathbf{m}_{t-1}, \quad (2.38)$$

$$\mathbf{R}_t = \mathbf{G}_t \mathbf{C}_{t-1} \mathbf{G}_t^T + \mathbf{W}_t. \quad (2.39)$$

Passo 1: *Priori* para λ_t .

$\lambda_t = \mathbf{F}_t^T \boldsymbol{\theta}_t$ é uma função linear de $\boldsymbol{\theta}_t$. Assim, sob (2.37) λ_t e $\boldsymbol{\theta}_t$ têm uma

distribuição *priori* conjunta parcialmente especificada por

$$\lambda_t, \boldsymbol{\theta}_t | y_{0:t-1} \sim \left[\left(\begin{array}{c} f_t \\ \mathbf{a}_t \end{array} \right), \left(\begin{array}{cc} Q_t & \mathbf{F}_t^T \mathbf{R}_t \\ \mathbf{R}_t \mathbf{F}_t & \mathbf{R}_t \end{array} \right) \right], \quad (2.40)$$

onde,

$$\begin{aligned} f_t &= E[\lambda_t | y_{0:t-1}] \\ &= E[\mathbf{F}_t^T \boldsymbol{\theta}_t | y_{0:t-1}] \\ &= \mathbf{F}_t^T \mathbf{a}_t, \end{aligned}$$

e

$$\begin{aligned} Q_t &= \text{Var}[\lambda_t | y_{0:t-1}] \\ &= \text{Var}[\mathbf{F}_t^T \boldsymbol{\theta}_t | y_{0:t-1}] \\ &= \mathbf{F}_t^T \text{Var}[\boldsymbol{\theta}_t | y_{0:t-1}] \mathbf{F}_t \\ &= \mathbf{F}_t^T \mathbf{R}_t \mathbf{F}_t. \end{aligned}$$

Passo 2: Previsão um passo a frente.

Assumindo que $Y_t | \eta_t$ pertence à FE e que a *priori* de η_t é conjugada, a previsão um passo a frente é dada por

$$f(y_t | y_{0:t-1}) = \frac{c(r_t, s_t) b(y_t, V_t)}{c(r_t + \phi_t y_t, s_t + \phi_t)}. \quad (2.41)$$

Os hiperparâmetros s_t e r_t são elicitados para serem consistentes com os momentos de λ_t , satisfazendo as equações

$$E[g(n_t) | y_{0:t-1}] = f_t, \quad (2.42)$$

$$\text{Var}[g(n_t) | y_{0:t-1}] = Q_t. \quad (2.43)$$

Passo 3: Atualização para η_t .

Observando y_t

$$f(\eta_t | y_{0:t}) = c(r_t + \phi_t y_t, s_t + \phi_t) \exp[(r_t + \phi_t y_t) \eta_t - (s_t + \phi_t) a(\eta_t)], \quad (2.44)$$

com média e variância a *posteriori* de $\lambda_t = g(\eta_t)$ sendo

$$f_t^* = E[g(\eta) | y_{0:t}], \quad (2.45)$$

$$Q_t^* = \text{Var}[g(\eta) | y_{0:t}]. \quad (2.46)$$

Passo 4a: Estrutura condicional para $\boldsymbol{\theta}_t | \eta_t, y_{0:t-1}$.

O objetivo da atualização é calcular a média e variância a *posteriori* para $\boldsymbol{\theta}_t$. Estes podem ser derivados da *posteriori* conjunta de λ_t e $\boldsymbol{\theta}_t$. A densidade conjunta, pelo Teorema de Bayes, é

$$f(\lambda_t, \boldsymbol{\theta}_t | y_{0:t}) \propto f(\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}) f(\lambda_t | y_{0:t}). \quad (2.47)$$

Assim, pela Equação 2.47, dado λ_t e $y_{0:t-1}$, $\boldsymbol{\theta}_t$ é condicionalmente independente de y_t , e segue que

$$f(\boldsymbol{\theta}_t | y_{0:t}) = \int f(\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}) f(\lambda_t | y_{0:t}) d\lambda_t, \quad (2.48)$$

onde, $f(\lambda_t | y_{0:t})$ é obtida pela atualização e $f(\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1})$ não é especificada completamente.

Para completar o filtro precisamos somente de $\mathbf{m}_t = E(\boldsymbol{\theta}_t | y_{0:t})$ e $\mathbf{C}_t = Var(\boldsymbol{\theta}_t | y_{0:t})$, mas, para isso, necessitamos da média e matriz de variância a *priori* de $\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}$. Porém, devido a especificação incompleta da *priori* conjunta, esses momentos condicionais são desconhecidos. No entanto, dado a especificações dos momentos a *priori* conjunta, eles podem ser estimados usando técnicas bayesianas padrões, como segue.

Passo 4b: Estimação dos momentos de $\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}$.

Os momentos associados a $\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}$ são estimados via *Linear Bayes* (LB) (West & Harrison [25]) :

$$\begin{aligned} \hat{E}[\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}] &= \mathbf{a}_t + \mathbf{R}_t \mathbf{F}_t (\lambda_t - f_t) / Q_t, \\ \hat{Var}[\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}] &= \mathbf{R}_t - \mathbf{R}_t \mathbf{F}_t \mathbf{F}_t^T \mathbf{R}_t / Q_t. \end{aligned}$$

Passo 5: Atualização para $\boldsymbol{\theta}_t$.

Os momentos,

$$E[\boldsymbol{\theta}_t | y_{0:t}] = E[E\{\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}\} | y_{0:t}],$$

e

$$Var[\boldsymbol{\theta}_t | y_{0:t}] = Var[E\{\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}\} | y_{0:t}] + E[Var[\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}] | y_{0:t}],$$

são estimados substituindo os estimadores LB dos momentos de $\boldsymbol{\theta}_t | \lambda_t, y_{0:t-1}$. Então,

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{R}_t \mathbf{F}_t (f_t^* - f_t) / Q_t, \quad (2.49)$$

$$\mathbf{C}_t = \mathbf{R}_t - \mathbf{R}_t \mathbf{F}_t \mathbf{F}_t^T \mathbf{R}_t (1 - Q_t^*/Q_t) / Q_t. \quad (2.50)$$

As demonstrações dos passos do filtro para o MLDG está bem detalhada em West & Harrison [25]).

2.3.2 Exemplo: Modelo Dinâmico Bernoulli (MDB)

Seja Y_t uma série temporal com distribuição de Bernoulli com parâmetro p_t , $Y_t \sim Ber(p_t)$, então

$$\begin{aligned} f(y_t|p_t) &= p_t^{y_t} (1 - p_t)^{1-y_t} \\ &= \left(\frac{p_t}{1 - p_t} \right)^{y_t} (1 - p_t) \\ &= \exp\{y_t \log \left(\frac{p_t}{1 - p_t} \right) + (1 - p_t)\}, \end{aligned}$$

fazendo,

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

temos que,

$$\log(1 - p_t) = -\log(1 + \exp(\eta_t)), \quad (2.51)$$

desse modo, podemos reescrever $f(y_t|p_t)$ em relação a η_t

$$f(y_t|\eta_t) = \exp\{y_t \eta_t - \log[1 + \exp(\eta_t)]\},$$

logo, a distribuição Bernoulli pertence à FE, com $\eta_t = \log[p_t/(1-p_t)]$, $V_t^{-1} = \phi_t = 1$, $b(y_t, V_t) = 1$ e $a(\eta) = \log[1 + \exp(\eta_t)]$.

A função média é

$$\begin{aligned} \mu_t &= E(Y_t|\eta_t) \\ &= a'(\eta_t) \\ &= \log'[1 + \exp(\eta_t)] \\ &= p_t, \end{aligned}$$

logo,

$$\eta_t = \log\left(\frac{p_t}{1-p_t}\right) \text{ e } p_t = [1 + \exp(-\eta_t)]^{-1}.$$

A *priori* conjugada para a distribuição Bernoulli é da forma

$$f(\eta_t|y_{0:t-1}) = c(r_t, s_t) \exp\{r_t\eta_t - s_t \log[1 + \exp(\eta_t)]\},$$

usando $\eta_t = \log[p_t/(1-p_t)]$, então a *priori* em função de p_t é da forma,

$$\begin{aligned} f(p_t|y_{0:t-1}) &= f_{\eta_t|y_{0:t-1}}(p_t) \left| \log' \left(\frac{p_t}{1-p_t} \right) \right| \\ &\propto \exp\{r_t \log(p_t) + s_t \log(1-p_t)\} [p_t(1-p_t)]^{-1} \\ &\propto \exp\{\log(p_t)^{r_t} + \log(1-p_t)^{s_t}\} [p_t(1-p_t)]^{-1} \\ &\propto p_t^{r_t} (1-p_t)^{s_t} [p_t(1-p_t)]^{-1} \\ &\propto p_t^{r_t-1} (1-p_t)^{s_t-1}, \end{aligned}$$

que é o núcleo da distribuição Beta com parâmetros r_t e s_t , com constante normalizadora $c(r_t, s_t) = \frac{\Gamma(r_t+s_t)}{\Gamma(r_t)\Gamma(s_t)}$. Portanto

$$p_t|y_{0:t-1} \sim \text{Beta}(r_t, s_t). \quad (2.52)$$

Neste caso, $g(\cdot)$ é a função identidade,

$$\lambda_t = g(\eta_t) = \eta_t = \log\left(\frac{p_t}{1-p_t}\right). \quad (2.53)$$

Desta forma, temos que

$$\begin{aligned} f_t &= E(\eta_t|y_{0:t-1}) = E\left[\log\left(\frac{p_t}{1-p_t}\right) \middle| y_{0:t-1}\right] \\ &= E[\log(p_t|y_{0:t-1})] - E[\log(1-p_t|y_{0:t-1})] \\ &= \psi(r_t) - \psi(s_t), \end{aligned} \quad (2.54)$$

e

$$\begin{aligned} Q_t &= \text{Var}(\eta_t|y_{0:t-1}) = \text{Var}\left[\log\left(\frac{p_t}{1-p_t}\right) \middle| y_{0:t-1}\right] \\ &= \text{Var}[\log(p_t|y_{0:t-1})] + \text{Var}[\log(1-p_t|y_{0:t-1})] \\ &= \psi'(r_t) + \psi'(s_t), \end{aligned} \quad (2.55)$$

onde, $\psi(a) = \frac{\Gamma'(a)}{\Gamma(a)}$, é a função digama e $\psi'(a)$ a função trigama. Com isso, dados f_t e Q_t , definimos os parâmetros da *priori* conjugada $\text{Beta}(r_t, s_t)$.

Para valores elevados de r_t e s_t , $\psi(a) \approx \log(a)$, enquanto $\psi'(a) \approx a^{-1}$ (West & Harrison [25]). Desta forma,

$$f_t \approx \log\left(\frac{r_t}{s_t}\right) \quad \text{e} \quad Q_t \approx \frac{1}{r_t} + \frac{1}{s_t}.$$

Então,

$$r_t \approx Q_t^{-1}[1 + \exp(f_t)], \quad (2.56)$$

$$s_t \approx Q_t^{-1}[1 + \exp(-f_t)]. \quad (2.57)$$

A posteriori para p_t é calculada através do Teorema de Bayes

$$\begin{aligned} f(p_t|y_{0:t}) &\propto f(p_t|y_{0:t-1})f(y_t|p_t) \\ &\propto p_t^{r_t-1}(1-p_t)^{s_t-1}p_t^{y_t}(1-p_t)^{1-y_t} \\ &\propto p_t^{(r_t+y_t)-1}(1-p_t)^{(s_t-y_t+1)-1}, \end{aligned}$$

logo,

$$p_t|y_{0:t} \sim \text{Beta}(r_t + y_t, s_t - y_t + 1). \quad (2.58)$$

Então,

$$\begin{aligned} f_t^* &= E(\eta_t|y_{0:t}) = E\left[\log\left(\frac{p_t}{1-p_t}\right)\middle|y_{0:t}\right] \\ &= E[\log(p_t|y_{0:t})] - E[\log(1-p_t|y_{0:t})] \\ &= \psi(r_t + y_t) - \psi(s_t - y_t + 1), \end{aligned} \quad (2.59)$$

e

$$\begin{aligned} Q_t^* &= \text{Var}(\eta_t|y_{0:t}) = \text{Var}\left[\log\left(\frac{p_t}{1-p_t}\right)\middle|y_{0:t}\right] \\ &= \text{Var}[\log(p_t|y_{0:t})] + \text{Var}[\log(1-p_t|y_{0:t})] \\ &= \psi'(r_t + y_t) + \psi'(s_t - y_t + 1), \end{aligned} \quad (2.60)$$

portanto, finalizando o filtro, temos que

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{R}_t \mathbf{F}_t (f_t^* - f_t) / Q_t,$$

e

$$\mathbf{C}_t = \mathbf{R}_t - \mathbf{R}_t \mathbf{F}_t \mathbf{F}_t^T \mathbf{R}_t (1 - Q_t^*/Q_t) / Q_t.$$

MLGD's para outras distribuições que pertencem a família exponencial podem ser vistos em Triantafyllopoulos [24], Correia [5] e Silva et al. [21].

2.4 Multiprocessos

A técnica de multiprocessos (Harrison & Stevens [8] e West & Harrison [25]) pode ser usada para escolher um Modelo Dinâmico (MD) ou combinações de classes de MD's que tenham a melhor performance em uma determinada série temporal. Modelos multiprocessos são combinações de modelos através de uma classe de MD's que proporcionam um supermodelo geral para a série. Genericamente, temos um MD definido pela quádrupla habitual em cada momento t , que denotaremos por $\mathbf{M}_t : \{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}_t$, condicionada a y_0 , informação a *priori* que representa todas as quantidades conhecidas. Suponha que todos os parâmetros que definem o modelo são indicados por $\boldsymbol{\alpha}$ e possivelmente são sujeitas a incertezas como fatores de descontos e as variâncias, V_t e \mathbf{W}_t . Representamos a dependência do modelo sobre estas incertezas por $\mathbf{M}_t = \mathbf{M}_t(\boldsymbol{\alpha})$, ($t=1,2,\dots$). Logo $\mathbf{M}_t(\boldsymbol{\alpha})$ é um MD para o tempo t e um dado valor de $\boldsymbol{\alpha}$.

Definimos \mathbf{A} como sendo o conjunto dos possíveis valores de $\boldsymbol{\alpha}$, onde assumimos que \mathbf{A} é discreto e finito. A classe de MD's no tempo t é dada por:

$$\{\mathbf{M}_t(\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in \mathbf{A}\}.$$

Considere as seguintes classes:

(Classe I) para algum $\boldsymbol{\alpha}_0 \in \mathbf{A}$, $\mathbf{M}_t(\boldsymbol{\alpha}_0)$ é válido para todo t .

(Classe II) para alguma sequência de valores $\boldsymbol{\alpha}_t \in \mathbf{A}$, ($t = 1, 2, 3, \dots$), $\mathbf{M}_t(\boldsymbol{\alpha}_t)$ é válido para o tempo t .

Na Classe I um único MD é adequado para todo tempo, mas há incerteza quanto aos verdadeiros valores que definem o vetor de parâmetros $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0$. Na Classe II consideramos que não existe um único MD aceito como adequado para todo tempo.

Neste trabalho, abordaremos apenas o modelo multiprocesso classe I (mais sobre o classe II pode ser encontrado em West & Harrison [25]).

Definição 8 *Suponha que para todo t , $\mathbf{M}_t(\boldsymbol{\alpha})$ é válido para algum $\boldsymbol{\alpha} \in \mathbf{A}$, com o conjunto dos parâmetros finito e discreto, $\mathbf{A} = \{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_N\}$, para algum inteiro $N \geq 1$, então a série $(Y_t)_{t \geq 1}$ segue um modelo multiprocesso classe I.*

Dado qualquer valor particular $\boldsymbol{\alpha} \in \mathbf{A}$, o MD $\mathbf{M}_t(\boldsymbol{\alpha})$ pode ser analisado usualmente produzindo sequências de *prioris*, *posteriors* e distribuições de previsões

que são sequencialmente atualizadas ao longo do tempo. Definimos \mathbf{X}_t como sendo qualquer vetor de quantidades aleatórias no tempo t , por exemplo, o vetor de estados $\boldsymbol{\theta}_t$ ou uma observação futura Y_{t+k} ($k > 0$). A inferência sobre \mathbf{X}_t em $\mathbf{M}_t(\boldsymbol{\alpha})$ é baseada na densidade $f(\mathbf{x}_t|\boldsymbol{\alpha}, \mathbf{y}_{0:t})$.

Iniciando com uma probabilidade a *priori* $f(\boldsymbol{\alpha}|y_0)$ para $\boldsymbol{\alpha}$, a informação é sequencialmente atualizada para proporcionar inferências sobre $\boldsymbol{\alpha}$ via *posteriori* no tempo t , $f(\boldsymbol{\alpha}|y_{0:t})$. Para a atualização usamos o Teorema de Bayes

$$f(\boldsymbol{\alpha}|y_{0:t}) \propto f(\boldsymbol{\alpha}|y_{0:t-1})f(y_{0:t}|\boldsymbol{\alpha}, y_{0:t-1}), \quad (2.61)$$

onde $f(y_{0:t}|\boldsymbol{\alpha}, y_{0:t-1})$ é a densidade de previsão um passo à frente de $\mathbf{M}_t(\boldsymbol{\alpha})$. A *posteriori* $f(\boldsymbol{\alpha}|y_{0:t})$ dá informações sobre $\boldsymbol{\alpha}$, identificando os valores mais prováveis com base nas *prioris* e nos dados observados.

Com a densidade marginal

$$f(\mathbf{x}_t|y_{0:t}) = \int_{\mathbf{A}} f(\mathbf{x}_t|\boldsymbol{\alpha}, y_{0:t})f(\boldsymbol{\alpha}|y_{0:t})d\boldsymbol{\alpha}, \quad (2.62)$$

fazemos inferências sobre \mathbf{X}_t . O processo de inferência sobre $\boldsymbol{\alpha}$ (vide Silva [20]) é composto pelo seguintes passos:

1. inicie o processo atribuindo *prioris* iniciais para cada $\boldsymbol{\alpha} \in \mathbf{A}$. Geralmente fazemos $p_0(j) = f(\boldsymbol{\alpha}_j|y_0) = 1/N$, ($j = 1, \dots, N$);
2. calcule a verossimilhança $l_t(j) = f(y_t|\boldsymbol{\alpha}_j, y_{0:t-1})$ no tempo t ;
3. calcule a *posteriori* para o tempo t ,

$$f(\boldsymbol{\alpha}_j|y_{0:t}) = \frac{1}{c_t}f(\boldsymbol{\alpha}_j|y_{0:t-1})f(y_{0:t}|\boldsymbol{\alpha}_j, y_{0:t-1})$$

com,

$$c_t = \sum_{j=1}^N f(\boldsymbol{\alpha}_j|y_{0:t-1})f(y_{0:t}|\boldsymbol{\alpha}_j, y_{0:t-1})$$

como assumimos que $\boldsymbol{\alpha}$ é discreto, as densidades marginais a *posteriori* são

$$f(\mathbf{x}_t|y_{0:t}) = \sum_{j=1}^N f(\mathbf{x}_t|\boldsymbol{\alpha}_j, y_{0:t})f(\boldsymbol{\alpha}_j|y_{0:t}). \quad (2.63)$$

No caso do MLD, se a variância observacional for conhecida então, por (2.63), todas as distribuições a *posteriores* de $\boldsymbol{\theta}_t$ e densidades preditivas de Y_{t+k} serão misturas de distribuições normais. Caso contrário, V_t desconhecido e constante, serão misturas de distribuições t-Student, pela Definição 6.

Os momentos das densidades de misturas são calculados pela seguinte expressão:

$$E[\mathbf{H}_t|y_{0:t}] = \sum_{j=1}^N E[\mathbf{H}_t|\boldsymbol{\alpha}_j, y_{0:t}]f(\boldsymbol{\alpha}_j|y_{0:t}), \quad (2.64)$$

onde \mathbf{H}_t pode representar qualquer expressão de variáveis de interesse. As densidades marginais para os dois primeiros momentos são:

$$E[\mathbf{X}_t|y_{0:t}] = \sum_{j=1}^N E[\mathbf{X}_t|\boldsymbol{\alpha}_j, y_{0:t}]f(\boldsymbol{\alpha}_j|y_{0:t}), \quad (2.65)$$

$$\begin{aligned} Var[\mathbf{X}_t|y_{0:t}] = \sum_{j=1}^N \{ & Var[\mathbf{X}_t|\boldsymbol{\alpha}_j, y_{0:t}] + [E[\mathbf{X}_t|\boldsymbol{\alpha}_j, y_{0:t}] - E[\mathbf{X}_t|y_{0:t}]] \times \\ & \times [E[\mathbf{X}_t|\boldsymbol{\alpha}_j, y_{0:t}] - E[\mathbf{X}_t|y_{0:t}]]^T \} f(\boldsymbol{\alpha}_j|y_{0:t}). \end{aligned} \quad (2.66)$$

Nas aplicações a escolha do melhor MD, para uma determinada série temporal, será baseada nas probabilidades a *posteriori*, $p_t(j) = f(\boldsymbol{\alpha}_j|y_{0:t})$. Probabilidades a *posterioris* definidas sobre valores de $\boldsymbol{\alpha}$ com pouco peso, indica que o modelo $\mathbf{M}_t(\boldsymbol{\alpha})$ pouco contribui para o modelo multiprocesso, indicando que este modelo não é adequado para a série, no tempo t .

2.5 Métricas de Performance de Previsão

Existem algumas medidas que nos permitem quantificar a qualidade dos MD's. Devido à natureza dos dados, contínuo com massa em zero, trabalharemos com as seguintes medidas

- **MAD** (*Mean Absolute Deviation*) Erro Absoluto Médio, definida por:

$$MAD = \frac{1}{k-h} \sum_{t=h+1}^k |y_t - f_t| \quad (2.67)$$

- **MSE** (*Mean Square Error*) Erro quadrático Médio, definida por:

$$MSE = \frac{1}{k-h} \sum_{t=h+1}^k (y_t - f_t)^2, \quad (2.68)$$

onde k denota o horizonte de previsão, h o corte, y_t a observação e f_t o valor da previsão no tempo, com $t \geq 1$. O corte é um recurso que retira as h primeiras previsões do cálculo das métricas de performance.

Para mais detalhes sobre essas métricas de desempenho vide Flores [6], Hyndman & Koehler [10], Silva [20] e Souza [22].

Capítulo 3

Modelo Dinâmico para Séries Temporais Contínuas com Massa em Zero

Neste capítulo apresentamos um novo modelo para séries temporais com dados contínuos e massa em zero, $Y_t \geq 0$, $t = 1, 2, \dots$. Decompomos a série em duas componentes, denotadas por U_t e Z_t . A primeira se refere à ocorrência de zeros e será modelada segundo um modelo dinâmico Bernoulli (ver Subseção 2.3.2). A segunda se refere a parte contínua da série e será modelada por um modelo dinâmico lognormal.

3.1 Representação Estocástica

Seja uma série temporal tal que $Y_t \geq 0$, $t = 1, 2, \dots$. Suponha que

$$Y_t = Z_t U_t, \tag{3.1}$$

onde, $Z_t \sim \log N[\mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t]$ e $U_t \sim \text{Ber}[p_t]$, com $\log N$ denotando a distribuição Log-Normal. Então, a função de distribuição de Y_t , é dada por

$$\begin{aligned}
F(y_t|\boldsymbol{\theta}_{1,t}, p_t) &= P(Y_t \leq y_t) = P(Z_t U_t \leq y_t) \\
&= \sum_{u=0}^1 P(Z_t U_t \leq y_t | U_t = u) P(U_t = u) \\
&= P(Z_t U_t \leq y | U_t = 0) P(U_t = 0) + P(Z_t U_t \leq y_t | U_t = 1) P(U_t = 1) \\
&= P(0 \leq y) P(U_t = 0) + P(Z_t \leq y_t) P(U_t = 1) \\
&= P(U_t = 0) + P(Z_t \leq y_t) P(U_t = 1) \\
&= p_t + (1 - p_t) L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t), \tag{3.2}
\end{aligned}$$

onde $L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t)$ é a função de distribuição da Lognormal com parâmetros $\mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}$ e V_t .

3.2 Modelo para séries com dados contínuos e massa em zero

De (3.2) e de Agresti & Min [1], podemos representar a verossimilhança do modelo, no tempo t , como:

$$\ell(\boldsymbol{\theta}_{1,t}, p_t) \propto p_t^{I(y_t=0)} \{(1 - p_t) L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t)\}^{1-I(y_t=0)} \tag{3.3}$$

com $I(\cdot)$ denotando a função indicadora.

Para complementar o modelo, a *priori* para p_t é:

$$p_t | y_{0:t-1} \sim Ber[r_t, s_t]. \tag{3.4}$$

Os hiperparâmetros r_t e s_t , são escolhidos para serem consistentes com os momentos de $\lambda_{2,t}$, satisfazendo as equações

$$E[\lambda_{2,t} | y_{0:t-1}] = f_{2,t}, \tag{3.5}$$

$$Var[\lambda_{2,t} | y_{0:t-1}] = Q_{2,t}. \tag{3.6}$$

com $\lambda_{2,t}$ sendo o preditor linear que, nesse caso, é denotado por

$$\lambda_{2,t} = g(p_t) = \log\left(\frac{p_t}{1 - p_t}\right) = \mathbf{F}_{2,t}^T \boldsymbol{\theta}_{2,t}, \tag{3.7}$$

pelo MDB, Equação (2.56) e (2.57), temos que

$$r_t \approx [1 + \exp(f_{2,t})]/Q_{2,t}^{-1}, \quad (3.8)$$

$$s_t \approx [1 + \exp(-f_{2,t})]/Q_{2,t}^{-1}. \quad (3.9)$$

Para uma série m -dimensional podemos caracterizar nosso modelo de forma semelhante ao MLD e MLGD como

Equações das Observações:

$$F(y_t | \boldsymbol{\theta}_{1,t}, p_t) = L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t)(1 - p_t) + p_t; \quad (3.10)$$

$$\lambda_{2,t} = g(p_t) = \log \left(\frac{p_t}{1 - p_t} \right) = \mathbf{F}_{2,t}^T \boldsymbol{\theta}_{2,t}; \quad (3.11)$$

$$p_t | y_{0:t-1} \sim \text{Beta}[r_t, s_t]. \quad (3.12)$$

Equações de evolução:

$$\boldsymbol{\theta}_{1,t} = \mathbf{G}_{1,t} \boldsymbol{\theta}_{1,t-1} + \boldsymbol{\omega}_{1,t}, \quad \boldsymbol{\omega}_{1,t} \sim N[\mathbf{0}, \mathbf{W}_{1,t}]; \quad (3.13)$$

$$\boldsymbol{\theta}_{2,t} = \mathbf{G}_{2,t} \boldsymbol{\theta}_{2,t-1} + \boldsymbol{\omega}_{2,t}, \quad \boldsymbol{\omega}_{2,t} \sim [\mathbf{0}, \mathbf{W}_{2,t}]. \quad (3.14)$$

em que, para $t = 1, 2, \dots, m$ e $i = 1, 2$, temos que:

- $\boldsymbol{\theta}_{i,t}$ é um vetor q_i -dimensional denominado parâmetro de estado;
- $\mathbf{F}_{i,t}$ é um vetor q_i -dimensional de constantes conhecidas;
- $\mathbf{G}_{i,t}$ é a matriz de evolução $q_i \times q_i$;
- V_t é o segundo parâmetro associado ao modelo Lognormal;
- $\mathbf{W}_{i,t}$ é uma matriz de covariância $q_i \times q_i$ conhecida associada ao erro de evolução $\boldsymbol{\omega}_{i,t}$.

Note que os parâmetros de estados, $\boldsymbol{\theta}_{1,t}$ e $\boldsymbol{\theta}_{2,t}$, foram especificados de acordo com o MLD e MLGD, respectivamente, e portanto, pela Equação (2.8), temos que

$$\boldsymbol{\theta}_{1,t} | \boldsymbol{\theta}_{1,t-1} \sim N[\mathbf{G}_{1,t} \boldsymbol{\theta}_{1,t-1}, \mathbf{W}_{1,t}], \quad (3.15)$$

e por pressuposto do MLGD, o modelo é completado assumindo as informações a *prioris*

$$\boldsymbol{\theta}_{1,0} | y_0 \sim N[\mathbf{m}_{1,0}, \mathbf{C}_{1,0}]; \quad (3.16)$$

$$\boldsymbol{\theta}_{2,0} | y_0 \sim [\mathbf{m}_{2,0}, \mathbf{C}_{2,0}], \quad (3.17)$$

com y_0 sendo a informação inicial no tempo $t = 0$.

O procedimento para estimação dos parâmetros da parte contínua do modelo é feito pelo Filtro de Kalman.

Proposição 1 *Para todo $t > 0$,*

$$F(y_t | \boldsymbol{\theta}_{1,t}, y_{0:t-1}) = L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t) \frac{s_t}{r_t + s_t} + \frac{r_t}{r_t + s_t}. \quad (3.18)$$

Demonstração:

$$\begin{aligned} F(y_t | \boldsymbol{\theta}_{1,t}, y_{0:t-1}) &= \int_0^1 F(y_t | \boldsymbol{\theta}_{1,t}, p_t) f(p_t | y_{0:t-1}) dp_t \\ &= \int_0^1 [L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t)(1 - p_t) + p_t] \left[\frac{1}{B(r_t, s_t)} p_t^{r_t-1} (1 - p_t)^{s_t-1} \right] dp_t \\ &= \int_0^1 L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t)(1 - p_t) \frac{p_t^{r_t-1} (1 - p_t)^{s_t-1}}{B(r_t, s_t)} dp_t \\ &\quad + \int_0^1 \frac{p_t^{r_t}}{B(r_t, s_t)} (1 - p_t)^{s_t-1} dp_t \\ &= \frac{L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t)}{B(r_t, s_t)} \int_0^1 p_t^{r_t-1} (1 - p_t)^{s_t} dp_t \\ &\quad + \frac{1}{B(r_t, s_t)} \int_0^1 p_t^{r_t} (1 - p_t)^{s_t-1} dp_t \\ &= \frac{L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t)}{B(r_t, s_t)} B(r_t, s_t + 1) + \frac{B(r_t + 1, s_t)}{B(r_t, s_t)} \\ &= L(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V_t) \frac{s_t}{r_t + s_t} + \frac{r_t}{r_t + s_t}. \end{aligned}$$

O filtro para a parte contínua do modelo proposto é dado no Teorema 5.

Teorema 5 *Parte contínua: a Previsão um passo a frente e as distribuições a posteriori, para todo t , são dadas por:*

a) *Posteriori marginal no tempo $t-1$:*

Para alguma média $\mathbf{m}_{1,t-1}$ e matriz de variância $\mathbf{C}_{1,t-1}$

$$\boldsymbol{\theta}_{1,t-1} | y_{0:t-1} \sim N[\mathbf{m}_{1,t-1}, \mathbf{C}_{1,t-1}].$$

b) *Priori marginal no tempo t :*

$$\boldsymbol{\theta}_{1,t} | y_{0:t-1} \sim N[\mathbf{a}_{1,t}, \mathbf{R}_{1,t}],$$

onde,

$$\mathbf{a}_{1,t} = \mathbf{G}_{1,t} \mathbf{m}_{1,t-1} \text{ e } \mathbf{R}_{1,t} = \mathbf{G}_{1,t} \mathbf{C}_{1,t-1} \mathbf{G}_{1,t}^T + \mathbf{W}_{1,t}.$$

c) *Previsão um passo à frente:*

$$F(y_t|y_{0:t-1}) = L(y_t|f_{1,t}, Q_{1,t}) \frac{s_t}{r_t + s_t} + \frac{r_t}{r_t + s_t},$$

onde,

$$f_{1,t} = \mathbf{F}_{1,t}^T \mathbf{a}_{1,t} \text{ e } Q_{1,t} = \mathbf{F}_{1,t}^T \mathbf{R}_{1,t} \mathbf{F}_{1,t} + V_{1,t}.$$

d) *Posteriori marginal no tempo t:*

$$\boldsymbol{\theta}_{1,t}|y_{0:t} \sim N[\mathbf{m}_{1,t}, \mathbf{C}_{1,t}],$$

com,

$$\mathbf{m}_{1,t} = \begin{cases} \mathbf{a}_{1,t}, & \text{se } y_t = 0 \\ \mathbf{a}_{1,t} + \mathbf{A}_{1,t}(\log(y_t) - f_{1,t}), & \text{se } y_t > 0 \end{cases}$$

e

$$\mathbf{C}_{1,t} = \begin{cases} \mathbf{R}_{1,t}, & \text{se } y_t = 0 \\ \mathbf{R}_{1,t} - \mathbf{A}_{1,t} Q_{1,t} \mathbf{A}_{1,t}^T, & \text{se } y_t > 0 \end{cases}$$

com,

$$\mathbf{A}_{1,t} = \mathbf{R}_{1,t} \mathbf{F}_{1,t} Q_{1,t}^{-1}.$$

Demonstração:

Seja a notação $n(X|a, b)$ a densidade da variável normal, X, com parâmetros a e b , e $l(X|a, b)$ a fdp da variável lognormal, X, com parâmetros a e b .

Pela informação *a priori*, $\boldsymbol{\theta}_{1,0}|y_0 \sim N(\mathbf{m}_{1,0}, \mathbf{C}_{1,0})$. Supondo *a)* verdadeiro, *b)* segue diretamente do filtro de Kalman do MLD.

c) Previsão um passo a frente:

$$\begin{aligned}
F(y_t|y_{0:t}) &= \int F(y_t|y_{0:t-1}, \boldsymbol{\theta}_{1,t})f(\boldsymbol{\theta}_{1,t}|y_{0:t-1})d\boldsymbol{\theta}_{1,t} \\
&= \int \left[\frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t}L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t) \right] n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \\
&= \frac{r_t}{r_t + s_t} \int n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \\
&\quad + \frac{s_t}{r_t + s_t} \int L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t)n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \\
&= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int \left[\int_0^{y_t} l(u_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t)du_t \right] n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \\
&= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int_0^{y_t} \left[\int l(u_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t)n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \right] du_t \\
&= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int_0^{y_t} l(u_t|f_{1,t}, Q_{1,t})du_t \\
&= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t}L(y_t|f_{1,t}, Q_{1,t})
\end{aligned}$$

d) *Posteriori* no tempo t :

$$\begin{aligned}
f(\boldsymbol{\theta}_{1,t}|y_t \leq u, y_{0:t-1}) &= \frac{P(y_t \leq u|\boldsymbol{\theta}_{1,t}, y_{0:t-1})f(\boldsymbol{\theta}_{1,t}|y_{0:t-1})}{P(y_t \leq u|y_{0:t-1})} \\
&= \frac{\left[\frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t}L(u|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t) \right] n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})}{\frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t}L(u|f_{1,t}, Q_{1,t})}
\end{aligned}$$

se $u = 0$, temos que $y_t = 0$ e

$$\begin{aligned}
f(\boldsymbol{\theta}_{1,t}|y_{0:t}) &= \frac{r_t}{r_t + s_t}n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})\frac{(r_t + s_t)}{r_t} \\
&= n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})
\end{aligned}$$

Se $u > 0$, temos que $y_t > 0$. Seja $B = (u, u + \epsilon)$

$$\begin{aligned}
f(\boldsymbol{\theta}_{1,t}|y_t \in B, y_{0:t-1}) &= \frac{[F(u + \epsilon|\boldsymbol{\theta}_{1,t}, y_{0:t-1}) - F(u|\boldsymbol{\theta}_{1,t}, y_{0:t-1})] n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})}{F(u + \epsilon|y_{0:t-1}) - F(u|y_{0:t-1})} \\
&= \frac{F(u + \epsilon|\boldsymbol{\theta}_{1,t}, y_{0:t-1}) - F(u|\boldsymbol{\theta}_{1,t}, y_{0:t-1})}{\epsilon} \frac{n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})}{F(u + \epsilon|y_{0:t-1}) - F(u|y_{0:t-1})} \\
&= \frac{\epsilon}{\epsilon}
\end{aligned}$$

aplicando o limite,

$$\begin{aligned} f(\boldsymbol{\theta}_{1,t}|y_{0:t}) &= \frac{l(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t)n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})}{l(y_t|f_{1,t}, Q_{1,t})} \\ &= n(\boldsymbol{\theta}_{1,t}|\mathbf{m}_{1,t}, \mathbf{C}_{1,t}). \end{aligned}$$

Colorário 2 *Momentos a posteriori no tempo t para a parte contínua.*

$$\mathbf{m}_{1,t} = \mathbf{a}_{1,t} + \mathbf{A}_{1,t}(\log(y_t) - f_{1,t})I(y_t > 0), \quad (3.19)$$

e

$$\mathbf{C}_{1,t} = \mathbf{R}_{1,t} - \mathbf{A}_{1,t}Q_{1,t}\mathbf{A}_{1,t}^T I(y_t > 0). \quad (3.20)$$

Proposição 2 *Para todo $t > 0$*

$$F(y_t|p_t, y_{0:t-1}) = p_t + (1 - p_t)L(y_t|f_{1,t}, Q_{1,t}). \quad (3.21)$$

Demonstração:

$$\begin{aligned} F(y_t|p_t, y_{0:t-1}) &= \int F(y_t|\boldsymbol{\theta}_{1,t}, p_t)f(\boldsymbol{\theta}_{1,t}|y_{0:t-1})d\boldsymbol{\theta}_{1,t} \\ &= \int [p_t + (1 - p_t)L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t)] n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \\ &= p_t \int n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \\ &\quad + (1 - p_t) \int L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V_t)n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, \mathbf{R}_{1,t})d\boldsymbol{\theta}_{1,t} \\ &= p_t + (1 - p_t)L(y_t|f_{1,t}, Q_{1,t}). \end{aligned}$$

Proposição 3 *posteriori para p_t , para todo t .*

$$f(p_t|y_{0:t}) \sim \text{Beta}[r_t + I(y_t = 0), s_t + I(y_t > 0)] \quad (3.22)$$

Demonstração: se $y_t = 0$

$$\begin{aligned}
f(p_t|y_t = 0, y_{0:t-1}) &= \frac{P(y_t = 0|p_t, y_{0:t-1})f(p_t|y_{0:t-1})}{P(y_t = 0|y_{0:t-1})} \\
&= \frac{p_t^{r_t-1}(1-p_t)^{s_t-1}}{B(r_t, s_t)} \\
&= \frac{r_t}{r_t + s_t} \\
&= \frac{p_t^{r_t}(1-p_t)^{s_t-1}(r_t + s_t)}{B(r_t, s_t)r_t} \\
&= \frac{\Gamma(r_t + s_t)p_t^{r_t}(1-p_t)^{s_t-1}(r_t + s_t)}{\Gamma(r_t)\Gamma(s_t)r_t} \\
&= \frac{\Gamma(r_t + s_t + 1)p_t^{r_t}(1-p_t)^{s_t-1}}{\Gamma(r_t + 1)\Gamma(s_t)}.
\end{aligned}$$

$\therefore p_t|y_t = 0, y_{0:t-1} \sim \text{Beta}[r_t + 1, s_t]$.

Se $y_t > 0$

$$\begin{aligned}
f(p_t|y_{0:t}) &= \frac{\left[\frac{dF(y_t|p_t, y_{0:t-1})}{dy_t} \right] f(p_t|y_{0:t-1})}{\frac{dF(y_t|y_{0:t-1})}{dy_t}} \\
&= \frac{(1-p_t)l(y_t|f_{1,t}, Q_{1,t})f(p_t|y_{0:t-1})}{\frac{s_t}{r_t + s_t}l(y_t|f_{1,t}, Q_{1,t})} \\
&= \frac{(1-p_t)f(p_t|y_{0:t-1})(r_t + s_t)}{s_t} \\
&= \frac{(1-p_t)p_t^{r_t-1}(1-p_t)^{s_t-1}(r_t + s_t)}{s_t B(r_t, s_t)} \\
&= \frac{(1-p_t)^{s_t}p_t^{r_t-1}(r_t + s_t)\Gamma(r_t + s_t)}{s_t\Gamma(s_t)\Gamma(r_t)} \\
&= \frac{\Gamma(r_t + s_t + 1)}{\Gamma(r_t)\Gamma(s_t + 1)}p_t^{r_t-1}(1-p_t)^{s_t}
\end{aligned}$$

$\therefore p_t|y_{0:t} \sim \text{Beta}[r_t, s_t + 1] \quad \forall y_t > 0$.

Teorema 6 *Parte discreta: a previsão um passo à frente e as distribuições a posteriori, para todo t , são dadas por:*

a) *Momentos da posteriori no tempo $t-1$:*

Para alguma média $\mathbf{m}_{2,t-1}$ e matriz de variância $\mathbf{C}_{2,t-1}$

$$\boldsymbol{\theta}_{2,t-1}|y_{0:t-1} \sim [\mathbf{m}_{2,t-1}, \mathbf{C}_{2,t-1}].$$

b) Momentos a priori no tempo t :

$$\boldsymbol{\theta}_{2,t}|y_{0:t-1} \sim [\mathbf{a}_{2,t}, \mathbf{R}_{2,t}],$$

onde,

$$\mathbf{a}_{2,t} = \mathbf{G}_{2,t}\mathbf{m}_{2,t-1} \text{ e } \mathbf{R}_{2,t} = \mathbf{G}_{2,t}\mathbf{C}_{2,t-1}\mathbf{G}_{2,t}^T + \mathbf{W}_{2,t}.$$

c) Momentos para previsão um passo a frente:

$$f_{2,t} = E[g(p_t)|y_{0:t-1}],$$

$$Q_{2,t} = V[g(p_t)|y_{0:t-1}],$$

onde,

$$f_{2,t} = \mathbf{F}_{2,t}^T \mathbf{a}_{2,t} \text{ e } Q_{2,t} = \mathbf{F}_{2,t}^T \mathbf{R}_{2,t} \mathbf{F}_{2,t}.$$

d) Momentos a posteriori no tempo t :

$$\boldsymbol{\theta}_{2,t}|y_{0:t} \sim [\mathbf{m}_{2,t}, \mathbf{C}_{2,t}],$$

onde,

$$\mathbf{m}_{2,t} = \mathbf{a}_{2,t} + \mathbf{A}_{2,t}(f_{2,t}^* - f_{2,t}),$$

e

$$\mathbf{C}_{2,t} = \mathbf{R}_{2,t} - \mathbf{A}_{2,t}Q_{2,t}\mathbf{A}_{2,t}^T(1 - Q_{2,t}^*/Q_{2,t}),$$

com,

$$\mathbf{A}_{2,t} = \mathbf{R}_{2,t}\mathbf{F}_{2,t}Q_{2,t}^{-1};$$

$$f_{2,t}^* = \psi(r_t + I(y_t = 0)) - \psi(s_t + I(y_t > 0));$$

$$Q_{2,t}^* = \psi'(r_t + I(y_t = 0)) + \psi'(s_t + I(y_t > 0)).$$

Demonstração: A demonstração é baseada no MDB (ver Seção 2.3.2).

Sabemos, pela informação a priori, que $\boldsymbol{\theta}_{2,0}|y_0 \sim [\mathbf{m}_{2,0}, \mathbf{C}_{2,0}]$. Supondo a) verdadeiro, segue que

b) Momentos a *priori* no tempo t :

$$\begin{aligned}
E[\boldsymbol{\theta}_{2,t}|y_{0:t-1}] &= E[E(\boldsymbol{\theta}_{2,t}|\boldsymbol{\theta}_{2,t-1}, y_{0:t-1})] \\
&= E[(\mathbf{G}_{2,t}\boldsymbol{\theta}_{2,t-1}|y_{0:t-1})] \\
&= \mathbf{G}_{2,t}E(\boldsymbol{\theta}_{2,t-1}|y_{0:t-1}) \\
&= \mathbf{G}_{2,t}\mathbf{m}_{2,t-1};
\end{aligned}$$

$$\begin{aligned}
Var[\boldsymbol{\theta}_{2,t}|y_{0:t-1}] &= Var[E(\boldsymbol{\theta}_{2,t}|\boldsymbol{\theta}_{2,t-1}, y_{0:t-1})] + E[Var(\boldsymbol{\theta}_{2,t}|\boldsymbol{\theta}_{2,t-1}, y_{0:t-1})] \\
&= Var[\mathbf{G}_{2,t}\boldsymbol{\theta}_{2,t-1}|y_{0:t-1}] + E[\mathbf{W}_{2,t}] \\
&= \mathbf{G}_{2,t}\mathbf{C}_{2,t-1}\mathbf{G}_{2,t}^T + \mathbf{W}_{2,t}.
\end{aligned}$$

c) Momentos para a previsão um passo a frente:

$$\begin{aligned}
E[g(p_t)|y_{0:t-1}] &= E[\mathbf{F}_{2,t}^T\boldsymbol{\theta}_{2,t}|y_{0:t-1}] \\
&= \mathbf{F}_{2,t}^TE[\boldsymbol{\theta}_{2,t}|y_{0:t-1}] \\
&= \mathbf{F}_{2,t}^T\mathbf{a}_{2,t};
\end{aligned}$$

$$\begin{aligned}
Var[g(p_t)|y_{0:t-1}] &= Var[\mathbf{F}_{2,t}^T\boldsymbol{\theta}_{2,t}|y_{0:t-1}] \\
&= \mathbf{F}_{2,t}^TVar[\boldsymbol{\theta}_{2,t}|y_{0:t-1}]\mathbf{F}_{2,t} \\
&= \mathbf{F}_{2,t}^T\mathbf{R}_{2,t}\mathbf{F}_{2,t}.
\end{aligned}$$

d) Momentos a *posteriri* no tempo t :

Pela Proposição 3, sabemos que

$$p_t|y_{0:t} \sim Beta[r_t + I(y_t = 0), s_t + I(y_t > 0)],$$

disto temos que

$$\begin{aligned}
f_{2,t}^* &= E[g(p_t)|y_{0:t}] \\
&= \psi(r_t + I(y_t = 0)) - \psi(s_t + I(y_t > 0)),
\end{aligned}$$

e

$$\begin{aligned}
Q_{2,t}^* &= Var[g(p_t)|y_{0:t}] \\
&= \psi'(r_t + I(y_t = 0)) + \psi'(s_t + I(y_t > 0)).
\end{aligned}$$

Assim,

$$\mathbf{m}_{2,t} = \mathbf{a}_{2,t} + \mathbf{A}_{2,t}(f_{2,t}^* - f_{2,t}),$$

e

$$\mathbf{C}_{2,t} = \mathbf{R}_{2,t} - \mathbf{A}_{2,t}Q_{2,t}\mathbf{A}_{2,t}^T(1 - Q_{2,t}^*/Q_{2,t}).$$

3.2.1 Modelo com V_t Desconhecido

Assumindo que V_t é desconhecido e considerando $V_t = V = 1/\phi$. Então, para cada tempo t , o modelo proposto é especificado como

Equações das Observações:

$$F(y_t|\boldsymbol{\theta}_{1,t}, p_t, V) = L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V)(1 - p_t) + p_t; \quad (3.23)$$

$$\lambda_{2,t} = g(p_t) = \log\left(\frac{p_t}{1 - p_t}\right) = \mathbf{F}_{2,t}^T\boldsymbol{\theta}_{2,t}; \quad (3.24)$$

$$p_t|y_{0:t-1}, V \sim \text{Beta}[r_t, s_t]. \quad (3.25)$$

Equações de Evolução:

$$\boldsymbol{\theta}_{1,t}|\boldsymbol{\theta}_{1,t-1}, V \sim N[\mathbf{G}_{1,t}\boldsymbol{\theta}_{1,t-1}, V\mathbf{W}_{1,t}^*]; \quad (3.26)$$

$$\boldsymbol{\theta}_{2,t}|\boldsymbol{\theta}_{2,t-1}, V \sim [\mathbf{G}_{2,t}\boldsymbol{\theta}_{2,t-1}, \mathbf{W}_{2,t}]. \quad (3.27)$$

Informações Iniciais:

$$\boldsymbol{\theta}_{1,0}|y_0, V \sim N[\mathbf{m}_{1,0}, V\mathbf{C}_{1,0}^*]; \quad (3.28)$$

$$\boldsymbol{\theta}_{2,0}|y_0, V \sim [\mathbf{m}_{2,0}, \mathbf{C}_{2,0}]; \quad (3.29)$$

$$V|y_0 \sim \text{GamaInversa}[n_0/2, d_0/2] \implies \phi|y_0 \sim \text{Gama}[n_0/2, d_0/2]. \quad (3.30)$$

$E[\phi|y_0] = \frac{n_0/2}{d_0/2} = \frac{n_0}{d_0} = \frac{1}{S_0}$, S_0 é a estimativa inicial de V .

O processo de inferência para a parte contínua do modelo proposto com V desconhecido é dada no Teorema 7.

Teorema 7 *Parte Contínua:* Com o modelo especificado, obtemos as seguintes distribuições para cada tempo $t \geq 1$.

a) *Distribuição a priori:*

$$F(y_t|\boldsymbol{\theta}_{1,t}, y_{0:t-1}, V) = L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V)\frac{s_t}{r_t + s_t} + \frac{r_t}{r_t + s_t}.$$

b) *Condicional a V:*

$$\begin{aligned}\boldsymbol{\theta}_{1,t-1}|y_{0:t-1}, V &\sim N[\mathbf{m}_{1,t-1}, V\mathbf{C}_{1,t-1}^*]; \\ \boldsymbol{\theta}_{1,t}|y_{0:t-1}, V &\sim N[\mathbf{a}_{1,t}, V\mathbf{R}_{1,t}^*]; \\ F(y_t|y_{0:t-1}, V) &= L(y_t|f_{1,t}, VQ_{1,t}^*)\frac{s_t}{r_t + s_t} + \frac{r_t}{r_t + s_t}; \\ \boldsymbol{\theta}_{1,t}|y_{0:t}, V &\sim N[\mathbf{m}_{1,t}, V\mathbf{C}_{1,t}^*].\end{aligned}$$

com,

$$\begin{aligned}\mathbf{a}_{1,t} &= \mathbf{G}_{1,t}\mathbf{m}_{1,t-1}; & \mathbf{R}_{1,t}^* &= \mathbf{G}_{1,t}\mathbf{C}_{1,t-1}^*\mathbf{G}_{1,t}^T + \mathbf{W}_{1,t}^*; \\ f_{1,t} &= \mathbf{F}_{1,t}^T\mathbf{a}_{1,t}; & Q_{1,t}^* &= \mathbf{F}_{1,t}^T\mathbf{R}_{1,t}^*\mathbf{F}_{1,t} + 1; \\ \mathbf{m}_{1,t} &= \mathbf{a}_{1,t} + \mathbf{A}_{1,t}(\log y_t - f_{1,t})I(y_t > 0); & \mathbf{C}_{1,t}^* &= \mathbf{R}_{1,t}^* - \mathbf{A}_{1,t}Q_{1,t}^*\mathbf{A}_{1,t}^T I(y_t > 0); \\ \mathbf{A}_{1,t} &= \mathbf{R}_{1,t}^*\mathbf{F}_{1,t}Q_{1,t}^{*-1}.\end{aligned}$$

c) *Para $\phi = V^{-1}$:*

$$\begin{aligned}\phi|y_{0:t-1} &\sim \text{Gama}[n_{t-1}/2, d_{t-1}/2]; \\ \phi|y_{0:t} &\sim \text{Gama}[n_t/2, d_t/2];\end{aligned}$$

onde,

$$n_t = n_{t-1} + I(y_t > 0) \quad e \quad d_t = d_{t-1} + \frac{(\log y_t - f_{1,t})^2}{Q_{1,t}^*} I(y_t > 0).$$

d) *Incondicional a V:*

$$\begin{aligned}\boldsymbol{\theta}_{1,t-1}|y_{0:t-1} &\sim t_{n_{t-1}}[\mathbf{m}_{1,t-1}, \mathbf{C}_{1,t-1}]; \\ \boldsymbol{\theta}_{1,t}|y_{0:t-1} &\sim t_{n_{t-1}}[\mathbf{a}_{1,t}, \mathbf{R}_{1,t}]; \\ F(y_t|y_{0:t-1}) &= \text{LogT}(y_t|n_{t-1}, f_{1,t}, Q_{1,t})\frac{s_t}{r_t + s_t} + \frac{r_t}{r_t + s_t}; \\ \boldsymbol{\theta}_{1,t}|y_{0:t} &\sim t_{n_t}[\mathbf{m}_{1,t}, \mathbf{C}_{1,t}].\end{aligned}$$

onde, $\text{LogT}(y_t|n_{t-1}, f_{1,t}, Q_{1,t})$ é a função de distribuição da Log - t com parâmetros n_{t-1} , $f_{1,t}$ e $Q_{1,t}$. E

$$\begin{aligned}\mathbf{C}_{1,t-1} &= S_{t-1}\mathbf{C}_{1,t-1}^*, \quad \mathbf{R}_{1,t} = S_{t-1}\mathbf{R}_{1,t}^*, \quad Q_{1,t} = S_{t-1}Q_{1,t}^*, \\ \mathbf{C}_{1,t} &= S_t\mathbf{C}_{1,t}^*, \quad S_{t-1} = d_{t-1}/n_{t-1} \quad e \quad S_t = d_t/n_t.\end{aligned}$$

Demonstração:

a) Distribuição a *priori* :

Segue diretamente da demonstração da Proposição 1, porém com V desconhecido.

b) Condicional a V :

Pela informação a *priori* e supondo a *priori* no tempo t verdadeira, então pelos mesmos resultados do Filtro, temos

$$\boldsymbol{\theta}_{1,t}|y_{0:t-1}, V \sim N[\mathbf{a}_{1,t}, V\mathbf{R}_{1,t}^*],$$

com,

$$\mathbf{a}_{1,t} = \mathbf{G}_{1,t}\mathbf{m}_{1,t},$$

e

$$\begin{aligned} V\mathbf{R}_{1,t}^* &= \mathbf{G}_{1,t}\mathbf{C}_{1,t-1}^*V\mathbf{G}_{1,t}^T + \mathbf{W}_{1,t}^*V \\ &= V(\mathbf{G}_{1,t}\mathbf{C}_{1,t-1}^*\mathbf{G}_{1,t}^T + \mathbf{W}_{1,t}^*). \end{aligned}$$

Para a previsão um passo à frente, a demonstração é análoga ao modelo com V conhecido

$$\begin{aligned} F(y_t|y_{0:t-1}, V) &= \int F(y_t|y_{0:t-1}, \boldsymbol{\theta}_{1,t}, V)f(\boldsymbol{\theta}_{1,t}|y_{0:t-1}, V)d\boldsymbol{\theta}_{1,t} \\ &= \int \left[\frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t}L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V) \right] n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, V\mathbf{R}_{1,t}^*)d\boldsymbol{\theta}_{1,t} \\ &= \frac{r_t}{r_t + s_t} \int n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, V\mathbf{R}_{1,t}^*)d\boldsymbol{\theta}_{1,t} \\ &\quad + \frac{s_t}{r_t + s_t} \int L(y_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V)n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, V\mathbf{R}_{1,t}^*)d\boldsymbol{\theta}_{1,t} \\ &= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int \left[\int_0^{y_t} l(u_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V)du_t \right] n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, V\mathbf{R}_{1,t}^*)d\boldsymbol{\theta}_{1,t} \\ &= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int_0^{y_t} \left[\int n(\boldsymbol{\theta}_{1,t}|\mathbf{a}_{1,t}, V\mathbf{R}_{1,t}^*)d\boldsymbol{\theta}_{1,t} \right] l(u_t|\mathbf{F}_{1,t}^T\boldsymbol{\theta}_{1,t}, V)du_t \\ &= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int_0^{y_t} l(u_t|f_{1,t}, VQ_{1,t}^*)du_t \\ &= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t}L(y_t|f_{1,t}, VQ_{1,t}^*), \end{aligned}$$

onde,

$$f_{1,t} = \mathbf{F}_{1,t}^T\mathbf{a}_{1,t},$$

e

$$VQ_{1,t}^* = \mathbf{F}_{1,t}^T V \mathbf{R}_{1,t}^* \mathbf{F}_{1,t} + V$$

Para a *posteriori* marginal no tempo t :

$$\begin{aligned} f(\boldsymbol{\theta}_{1,t} | y_t \leq u, y_{0:t-1}, V) &= \frac{P(y_t \leq u | \boldsymbol{\theta}_{1,t}, y_{0:t-1}, V) f(\boldsymbol{\theta}_{1,t} | y_{0:t-1}, V)}{P(y_t \leq u | y_{0:t-1}, V)} \\ &= \frac{\left[\frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} L(u | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V) \right] n(\boldsymbol{\theta}_{1,t} | \mathbf{a}_{1,t}, V \mathbf{R}_{1,t}^*)}{\frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} L(u | f_{1,t}, V Q_{1,t}^*)}. \end{aligned}$$

Se $u = 0$, temos que $y_t = 0$ e

$$\begin{aligned} f(\boldsymbol{\theta}_{1,t} | y_{0:t}, V) &= \frac{r_t}{r_t + s_t} n(\boldsymbol{\theta}_{1,t} | \mathbf{a}_{1,t}, V \mathbf{R}_{1,t}^*) \frac{r_t + s_t}{r_t} \\ &= n(\boldsymbol{\theta}_{1,t} | \mathbf{a}_{1,t}, V \mathbf{R}_{1,t}^*). \end{aligned}$$

Se $u > 0$, temos que $y_t > 0$. Seja $B = (u, u + \epsilon)$ e $H = \{y_{0:t-1}, V\}$

$$\begin{aligned} f(\boldsymbol{\theta}_{1,t} | y_t \in B, H) &= \frac{[F(u + \epsilon | \boldsymbol{\theta}_{1,t}, H) - F(u | \boldsymbol{\theta}_{1,t}, H)] n(\boldsymbol{\theta}_{1,t} | \mathbf{a}_{1,t}, V \mathbf{R}_{1,t}^*)}{F(u + \epsilon | y_{0:t-1}) - F(u | H)} \\ &= \frac{F(u + \epsilon | \boldsymbol{\theta}_{1,t}, H) - F(u | \boldsymbol{\theta}_{1,t}, H)}{\epsilon} \frac{n(\boldsymbol{\theta}_{1,t} | \mathbf{a}_{1,t}, V \mathbf{R}_{1,t}^*)}{\frac{F(u + \epsilon | H) - F(u | H)}{\epsilon}}, \end{aligned}$$

aplicando o limite ($\epsilon \rightarrow 0$),

$$\begin{aligned} f(\boldsymbol{\theta}_{1,t} | y_{0:t}, V) &= \frac{l(y_t | \mathbf{F}_{1,t}^T \boldsymbol{\theta}_{1,t}, V) n(\boldsymbol{\theta}_{1,t} | \mathbf{a}_{1,t}, V \mathbf{R}_{1,t}^*)}{l(y_t | f_{1,t}, V Q_{1,t}^*)} \\ &= n(\boldsymbol{\theta}_{1,t} | \mathbf{m}_{1,t}, V \mathbf{C}_{1,t}^*). \end{aligned}$$

$$\therefore \boldsymbol{\theta}_{1,t} | y_{0:t}, V \sim N(\mathbf{m}_{1,t}, V \mathbf{C}_{1,t}^*),$$

onde,

$$\mathbf{m}_{1,t} = \mathbf{a}_{1,t} + \mathbf{A}_{1,t} (\log(y_t) - f_{1,t}) I(y_t > 0),$$

e

$$\begin{aligned} V \mathbf{C}_{1,t}^* &= V \mathbf{R}_{1,t}^* - \mathbf{A}_{1,t} V Q_{1,t}^* \mathbf{A}_{1,t}^T I(y_t > 0) \\ &= V (\mathbf{R}_{1,t}^* - \mathbf{A}_{1,t} Q_{1,t}^* \mathbf{A}_{1,t}^T) I(y_t > 0), \end{aligned}$$

com

$$\mathbf{A}_{1,t} = \mathbf{R}_{1,t}^* \mathbf{F}_{1,t} \mathbf{Q}_{1,t}^{*-1}.$$

c) Para ϕ :

Pela informação a priori $\phi|y_0 \sim \text{Gama}[n_0/2, d_0/2]$, supondo verdadeiro $\phi|y_{0:t-1} \sim \text{Gama}[n_{t-1}/2, d_{t-1}/2]$, então:

$$\begin{aligned} f(\phi|y_t \leq u, y_{0:t-1}) &= \frac{P(y_t \leq u|\phi, y_{0:t-1})f(\phi|y_{0:t-1})}{P(y_t \leq u|y_{0:t-1})} \\ &= \frac{\left[\frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} L(u|f_{1,t}, VQ_{1,t}^*) \right] f(\phi|y_{0:t-1})}{P(y_t \leq u|y_{0:t-1})}. \end{aligned}$$

Se $u = 0$, temos que $y_t = 0$ e

$$f(\phi|y_{0:t}) = f(\phi|y_{0:t-1}).$$

Se $u > 0$, temos que $y_t > 0$. Seja $B = (u, u + \epsilon)$, então

$$f(\phi|y_t \in B, y_{0:t-1}) = \frac{\frac{F(u + \epsilon|\phi, y_{0:t-1}) - F(u|\phi, y_{0:t-1})}{\epsilon} f(\phi|y_{0:t-1})}{\frac{F(u + \epsilon|y_{0:t-1}) - F(u|y_{0:t-1})}{\epsilon}},$$

aplicando o limite ($\epsilon \rightarrow 0$),

$$\begin{aligned} f(\phi|y_{0:t}) &= \frac{l(y_t|f_{1,t}, \phi^{-1}Q_{1,t}^*)f(\phi|y_{0:t-1})}{f(y_t|y_{0:t-1})} \\ &\propto \phi^{1/2} \exp \left\{ -\frac{\phi}{2Q_{1,t}^*} (\log(y_t) - f_{1,t})^2 \right\} \phi^{\frac{n_{t-1}}{2}-1} \exp \left\{ -\phi \frac{d_{t-1}}{2} \right\} \\ &\propto \phi^{\frac{n_{t-1}+1}{2}-1} \exp \left\{ -\left(d_{t-1} + \frac{(\log(y_t) - f_{1,t})^2}{Q_{1,t}^*} \right) \frac{\phi}{2} \right\} \end{aligned}$$

que é o núcleo de uma distribuição Gama com parâmetros $n_{t-1} + 1$ e $d_{t-1} + (\log(y_t) - f_{1,t})^2/Q_{1,t}^*$. Portanto, para qualquer y_t

$$\phi|y_{0:t} \sim \text{Gama}[n_t/2, d_t/2],$$

onde,

$$n_t = n_{t-1} + I(y_t > 0) \quad e \quad d_t = d_{t-1} + \frac{(\log(y_t) - f_{1,t})^2}{Q_{1,t}^*} I(y_t > 0).$$

d) Incondicional a V:

A demonstração da *priori* no tempo t , a *posteriori* no tempo $t - 1$ e para a *posteriori* no tempo t , segue diretamente do MLD com variância desconhecida (ver subseção 2.2.6), partindo da teoria da Normal Gama.

Previsão um passo a frente:

$$\begin{aligned}
F(y_t|y_{0:t}) &= \int_0^\infty F(y_t|y_{0:t-1}, \phi) f(\phi|y_{0:t-1}) d\phi \\
&= \frac{r_t}{r_t + s_t} \int_0^\infty f(\phi|y_{0:t-1}) d\phi + \frac{s_t}{r_t + s_t} \int_0^\infty L(y_t|f_{1,t}, VQ_{1,t}^*) f(\phi|y_{0:t-1}) d\phi \\
&= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int_0^{y_t} U du_t,
\end{aligned}$$

onde,

$$\begin{aligned}
U &= \int_0^\infty \frac{\sqrt{\phi}}{u_t \sqrt{2\pi Q_{1,t}^*}} \exp \left\{ -\frac{\phi}{2Q_{1,t}^*} (\log(u_t) - f_{1,t})^2 \right\} \times \\
&\quad \times \frac{(d_{t-1}/2)^{\frac{n_{t-1}}{2}}}{\Gamma(n_{t-1}/2)} \phi^{\frac{n_{t-1}}{2}-1} \exp \left\{ -\frac{d_{t-1}}{2} \phi \right\} d\phi \\
&= \frac{(d_{t-1}/2)^{\frac{n_{t-1}}{2}}}{u_t \sqrt{2\pi Q_{1,t}^*} \Gamma(n_{t-1}/2)} \times \\
&\quad \times \left[\int_0^\infty \phi^{\frac{n_{t-1}+1}{2}-1} \exp \left\{ -\left(d_{t-1} + \frac{(\log(u_t) - f_{1,t})^2}{Q_{1,t}^*} \right) \frac{\phi}{2} d\phi \right\} \right] \\
&= \frac{(d_{t-1}/2)^{\frac{n_{t-1}}{2}}}{u_t \Gamma(n_{t-1}/2) \sqrt{2\pi Q_{1,t}^*}} \Gamma\left(\frac{n_{t-1}+1}{2}\right) \left\{ \frac{1}{2} \left[d_{t-1} + \frac{(\log(u_t) - f_{1,t})^2}{Q_{1,t}^*} \right] \right\}^{-\frac{n_{t-1}+1}{2}} \\
&= \frac{\Gamma\left(\frac{n_{t-1}+1}{2}\right)}{u_t \Gamma(n_{t-1}/2) \sqrt{\pi Q_{1,t}^*} d_{t-1}} \left[1 + \frac{(\log(u_t) - f_{1,t})^2}{Q_{1,t}^* d_{t-1}} \right]^{-\frac{n_{t-1}+1}{2}} \\
&= \frac{\Gamma\left(\frac{n_{t-1}+1}{2}\right)}{u_t \Gamma(n_{t-1}/2) \sqrt{\pi Q_{1,t}^*} n_{t-1}} \left[1 + \frac{1}{n_{t-1}} \frac{(\log(u_t) - f_{1,t})^2}{Q_{1,t}^*} \right]^{-\frac{n_{t-1}+1}{2}},
\end{aligned}$$

que é a fdp de uma Log-t com parâmetros n_{t-1} , $f_{1,t}$ e $Q_{1,t}$, que denotaremos

por $\text{logt}(u_t|n_{t-1}, f_{1,t}, Q_{1,t})$, logo

$$\begin{aligned} F(y_t|y_{0:t}) &= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \int_0^{y_t} \text{logt}(u_t|n_{t-1}, f_{1,t}, Q_{1,t}) du_t \\ &= \frac{r_t}{r_t + s_t} + \frac{s_t}{r_t + s_t} \text{LogT}(y_t|n_{t-1}, f_{1,t}, Q_{1,t}). \end{aligned}$$

A distribuição Log-t não possui média finita (Marchenko [16]), devido a isso usamos a mediana da distribuição um passo à frente como estatística de previsão.

O filtro para a parte discreta é o mesmo para o modelo com V conhecido.

Capítulo 4

Análise de Desempenho da Nossa Proposta

Neste capítulo estudamos nossa proposta através de séries temporais geradas de modo que seja possível visualizar tendências e sazonalidades. Não se trata de um estudo de simulação, onde o objetivo é criar cenários que se assemelham com a realidade, mas gerar séries com comportamentos claros e verificar se o modelo proposto consegue identificá-los.

Para cada série foram ajustados os modelos de tendência constante, linear, tendência sazonal e superposição de modelos. Esses modelos foram avaliados via multiprocesso classe I (Seção 2.4) e através das medidas de desempenho.

Para fins de abreviação, identificamos os modelos usando as seguintes notações:

- $\text{Mod.Pol}_{cont}(i)$: Modelo polinomial de ordem i , $i = 1, 2$, na parte contínua;
- $\text{Mod.Pol}_{disc}(i)$: Modelo polinomial de ordem i , $i = 1, 2$, na parte discreta;
- $\text{Mod.Trig}_{cont}(p,A)$: Modelo com representação trigonométrica da sazonalidade na parte contínua, com período p e harmônico(s) A ;
- $\text{Mod.Trig}_{disc}(p,A)$: Modelo com representação trigonométrica da sazonalidade na parte discreta, com período p e harmônico(s) A ;
- "+": superposição de modelos.

Para o estudo utilizamos os seguintes modelos:

- **Modelo 1:** $\text{Mod.Pol}_{cont}(1)$ e $\text{Mod.Pol}_{disc}(1)$;
- **Modelo 2:** $\text{Mod.Pol}_{cont}(1)$ e $\text{Mod.Pol}_{disc}(2)$;
- **Modelo 3:** $\text{Mod.Pol}_{cont}(2)$ e $\text{Mod.Pol}_{disc}(1)$;

- **Modelo 4:** Mod.Pol_{cont}(2) e Mod.Pol_{disc}(2);
- **Modelo 5:** Mod.Pol_{cont}(1) + Mod.Trig_{cont}(10,1) e Mod.Pol_{disc}(1);
- **Modelo 6:** Mod.Pol_{cont}(1) e Mod.Trig_{disc}(10,1);
- **Modelo 7:** Mod.Pol_{cont}(2) e Mod.Trig_{disc}(10,1).

Cada subseção desse capítulo é composta pelo estudo de uma série temporal que apresenta o comportamento de um desses modelos. Como as séries são geradas sem a adição de ruídos e algumas vezes sem componentes aleatórios em uma das partes, usaremos o conjunto de parâmetros iniciais, que denotaremos por ρ_0 , igual para todos os modelos: $\rho_0 = \{m_{1,0} = m_{2,0} = \mathbf{0}_{p_i \times 1}, C_{1,0} = C_{2,0} = \mathbf{1}_{p_i \times 1}, S_0 = n_0 = 1, p_0 = 0, 5\}$, onde $\mathbf{0}_{p_i \times 1}$ e $\mathbf{1}_{p_i \times 1}$ são vetores p_i -dimensionais de zeros e uns, respectivamente, com $i=1,2$.

4.1 Modelo 1

Considere a série temporal Y_t , $t = 1, \dots, 100$, com tendência constante, tanto na parte contínua quanto na discreta, fornecida pela equação:

$$Y_t = \exp(\pi)z_t,$$

com, z_t representando valores gerados de uma distribuição binomial com $p = 0,8$, para todo t .

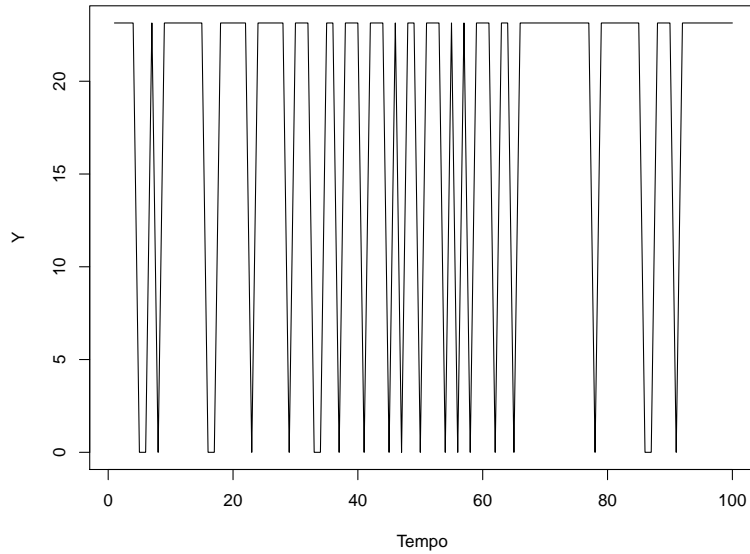


Figura 4.1: Série com tendências constantes.

Na Figura 4.1, observamos que o nível na parte contínua e a ocorrência de zeros é constante, caracterizando um modelo polinomial de primeira ordem para as partes discreta e contínua. Para a modelagem dessa série usamos os modelos 1,2,3 e 4.

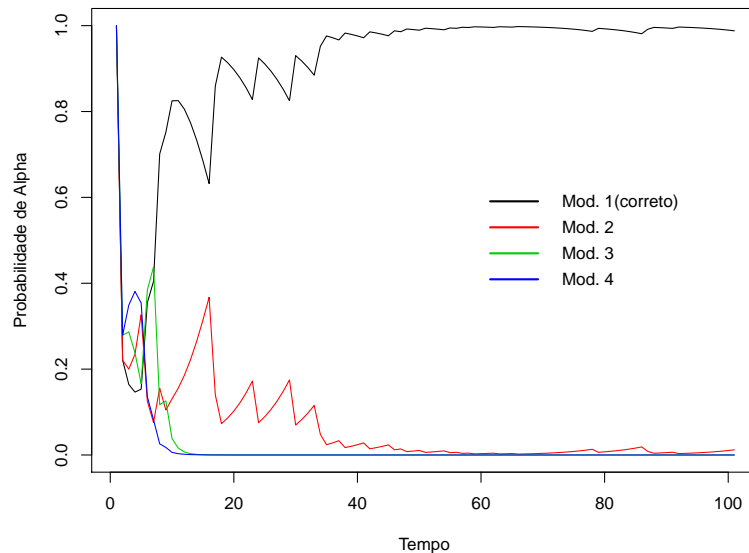


Figura 4.2: Multiprocesso Classe I: Série com tendências constantes.

Na Figura 4.2, temos o gráfico da probabilidade *a posteriori* de α_j , $f(\alpha_j|Y_{0:t})$, ($j = 1, 2, 3, 4$), em cada tempo t , via multiprocesso classe I. O Modelo 1 obteve o melhor desempenho indicando que nossa proposta conseguiu identificar o comportamento desta série.

Tabela 4.1: Medidas de desempenho: Série com tendências constantes

| | MAD | MSE |
|----------|--------|---------|
| Modelo 1 | 5,444 | 52,671 |
| Modelo 2 | 2,842 | 19,896 |
| Modelo 3 | 10,254 | 143,039 |
| Modelo 4 | 5,890 | 53,739 |

Na Tabela 4.1, temos as medidas de desempenho para os quatro primeiros modelos. Observamos que o Modelo 2 apresentou os menores valores do MAD, 2,842, e do MSE, 19,896, indicando que esse obteve o melhor desempenho para a série. O Modelo 1 ficou como o segundo melhor modelo para a série.

4.2 Modelo 2

Uma série contínua com massa em zero com tendência constante na parte contínua e tendência linear na ocorrência de zeros foi gerada a partir da equação:

$$Y_t = \exp(\pi)z_t, \quad t = 1, 2, \dots, 100,$$

com z_t sendo valores gerados de uma distribuição binomial com $p = \frac{\exp(u_t)}{1+\exp u_t}$, onde $u_1 = -1$ e $u_t = u_{t-1} + 0,05$.

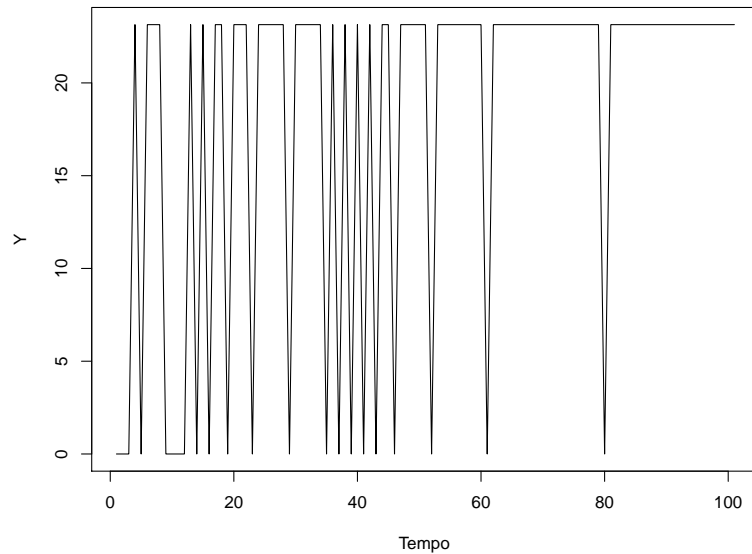


Figura 4.3: Série com tendência constante na parte contínua e linear na parte discreta.

Na Figura 4.3, encontra-se o gráfico da série gerada pela equação acima. Nesta série observamos que a parte contínua é constante, $\exp(\pi) = 23,14069$, enquanto a probabilidade da ocorrência de zeros diminui com o passar do tempo.

A Figura 4.4 mostra que o multiprocesso identificou corretamente a série.

Tabela 4.2: Medidas de desempenho: Série com tendência constante(contínua) e linear (discreta)

| | MAD | EMS |
|----------|--------|---------|
| Modelo 1 | 7,326 | 107,226 |
| Modelo 2 | 4,913 | 73,476 |
| Modelo 3 | 11,047 | 178,497 |
| Modelo 4 | 7,014 | 101,715 |

Observamos pela Tabela 4.2 das medidas de desempenho, que as melhores medidas de ajuste foram a do Modelo 2, indicando que este modelo é o mais adequado

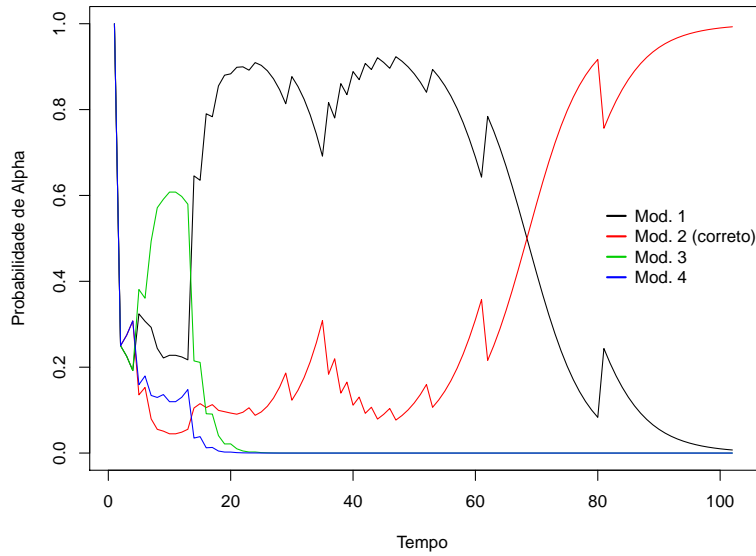


Figura 4.4: Multiprocesso Classe I: Série com tendência constante(contínua) e linear (discreta).

para a série.

4.3 Modelo 3

A partir da equação:

$$Y_t = \exp(t/100)z_t, \quad t = 1, 2, \dots, 100,$$

com z_t sendo valores simulados a partir de uma distribuição binomial com $p = 0,8$, geramos uma série com tendência linear na parte contínua e tendência constante na parte discreta como mostra a Figura 4.5.

Pela Figura 4.6 observamos que o Modelo 3 tem maior probabilidade *a posteriori*.

Tabela 4.3: Medidas de desempenho: Série com tendência linear (contínua) e constante (discreta).

| | MAD | MSE |
|----------|-------|-------|
| Modelo 1 | 0,213 | 0,075 |
| Modelo 2 | 0,131 | 0,034 |
| Modelo 3 | 0,168 | 0,052 |
| Modelo 4 | 0,101 | 0,024 |

Pelas medidas de desempenho, Tabela 4.3, o modelo com as menores medidas é o Modelo 4, mostrando que este modelo é superior aos demais, mas contrariando a

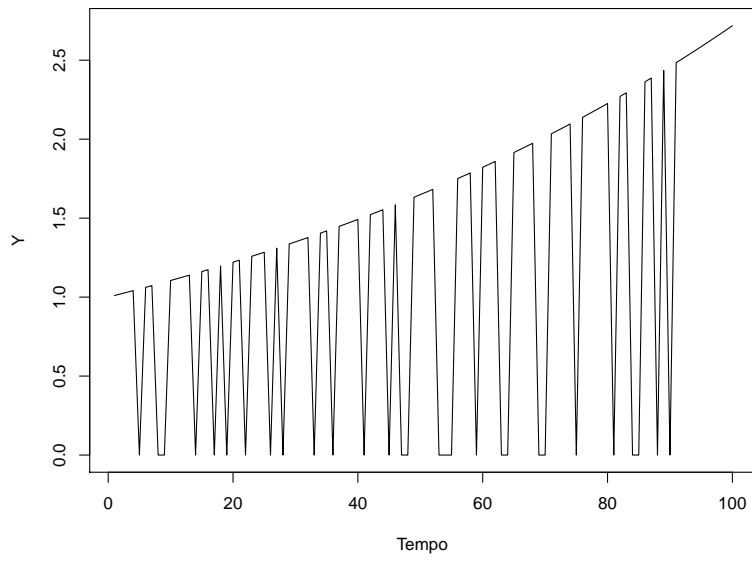


Figura 4.5: Série com tendência linear na parte contínua e constante na parte discreta.

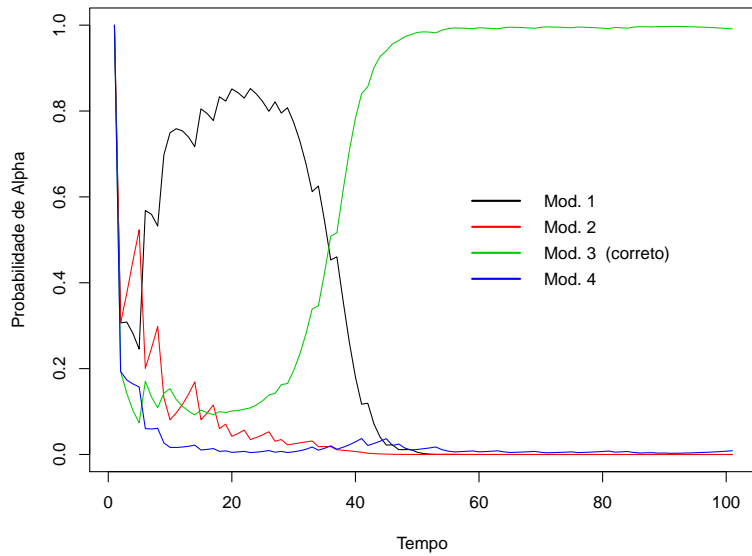


Figura 4.6: Multiprocesso Classe I: Série com tendência linear (contínua) e constante (discreta).

análise via multiprocesso que classificou o Modelo 3.

4.4 Modelo 4

Para gerar uma série que tenha tendência linear na parte contínua e na discreta, usamos a seguinte equação:

$$Y_t = \exp(t/100)z_t, \quad t = 1, 2, \dots, 100,$$

com z_t sendo valor gerados de uma distribuição binomial com $p = \frac{\exp(u_t)}{1+\exp u_t}$, onde $u_1 = -1$ e $u_t = u_{t-1} + 0.05$.

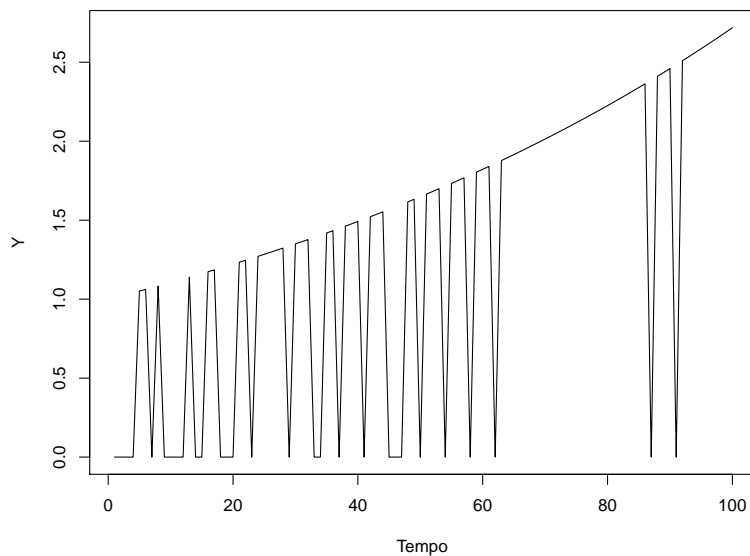


Figura 4.7: Modelo com tendências lineares

Observamos na Figura 4.7, a série gerada pela equação acima, apresentando tendência linear na parte contínua e tendência linear na ocorrência de zeros.

O Modelo 4 expressa as tendências da série na Figura 4.7. Analisando a Figura 4.8, observamos que o multiprocessado julgou o Modelo 4 como o mais adequado.

Tabela 4.4: Medidas de desempenho: Série com tendências lineares.

| | MAD | MSE |
|----------|-------|-------|
| Modelo 1 | 0,314 | 0,212 |
| Modelo 2 | 0,195 | 0,145 |
| Modelo 3 | 0,272 | 0,197 |
| Modelo 4 | 0,182 | 0,144 |

Através do MAD e MSE, o Modelo 4 mostrou-se superior aos demais, na Tabela 4.4 encontramos esses valores para os modelos.

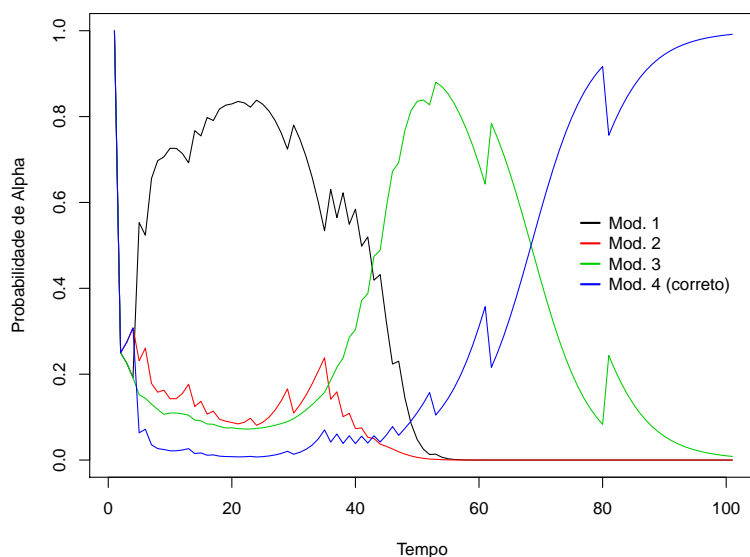


Figura 4.8: Multiprocesso Classe I: Série com tendências lineares.

4.5 Modelo 5

O Modelo 5, é semelhante ao Modelo 1, porém com um ciclo sazonal de período 10. Para gerar uma série com essas características, utilizamos a seguinte equação:

$$Y_t = \exp\{2 \sin(2t\pi/10)\} z_t, \quad t = 1, 2, \dots, 100,$$

com z_t sendo valores simulados de uma distribuição binomial com $p = 0,6$. A Figura 4.9 mostra a série gerada. A partir das probabilidades geradas pelo multiprocesso (Figura 4.10), o modelo com melhor performance é o Modelo 5.

Tabela 4.5: Medidas de desempenho: Série com tendência constante "+" sazonal na parte contínua e constante na discreta.

| | MAD | EMS |
|----------|-------|-------|
| Modelo 1 | 1,140 | 4,337 |
| Modelo 2 | 1,001 | 3,479 |
| Modelo 3 | 1,161 | 4,489 |
| Modelo 4 | 1,023 | 3,609 |
| Modelo 5 | 1,104 | 4,091 |

Na Tabela 4.5, temos as medidas desempenho para os cinco primeiros modelos. Observamos que através das medidas o modelo que se mostrou superior aos demais foi o Modelo 2. Apesar destas medidas serem próximas, o Modelo 5 ficou como sendo o terceiro melhor.

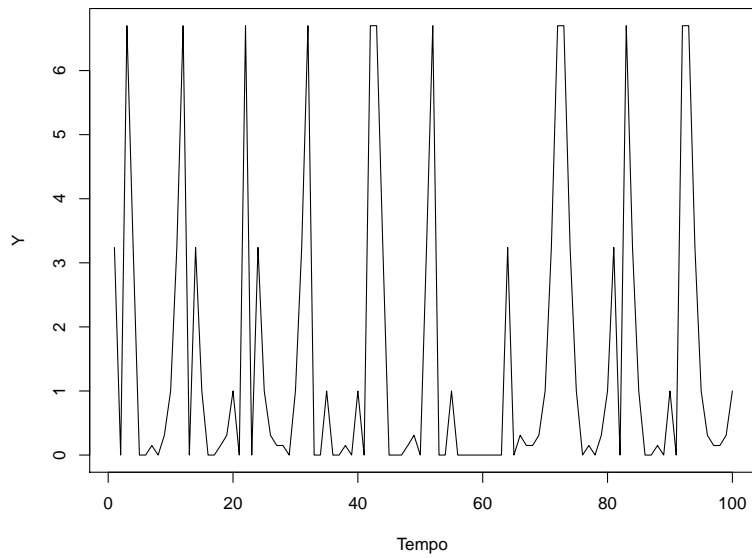


Figura 4.9: Série com tendências constante e sazonal na parte contínua e tendência constante na parte discreta

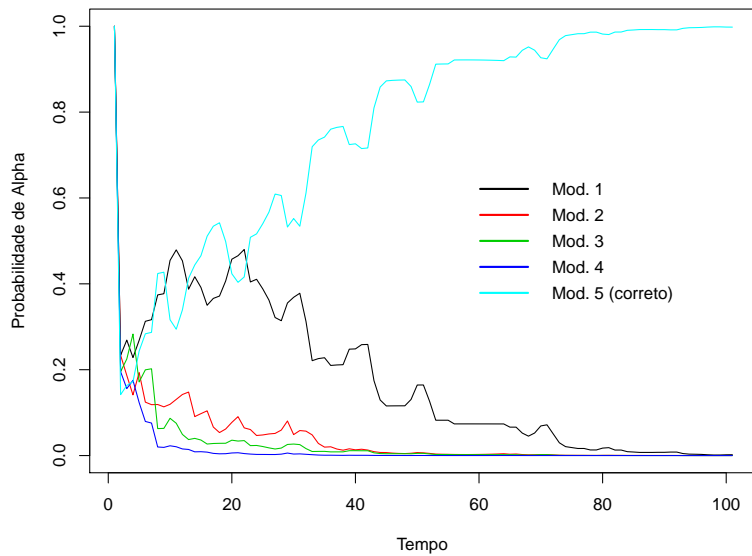


Figura 4.10: Multiprocesso Classe I: Série com tendência constante "+" sazonal na parte contínua e constante na discreta.

4.6 Modelo 6

Uma série com dados contínuos com massa em zero, $Y_t, t = 1, 2, \dots, 100$, com tendência constante na parte contínua e sazonalidade com período dez para a parte

discreta, pode ser gerada através da equação:

$$Y_t = \exp(\pi)z_t,$$

com z_t sendo valores simulados de uma distribuição binomial com $p = \frac{\exp(u_t)}{1+\exp u_t}$, onde $u_t = 3 \sin(2t\pi/10)$.

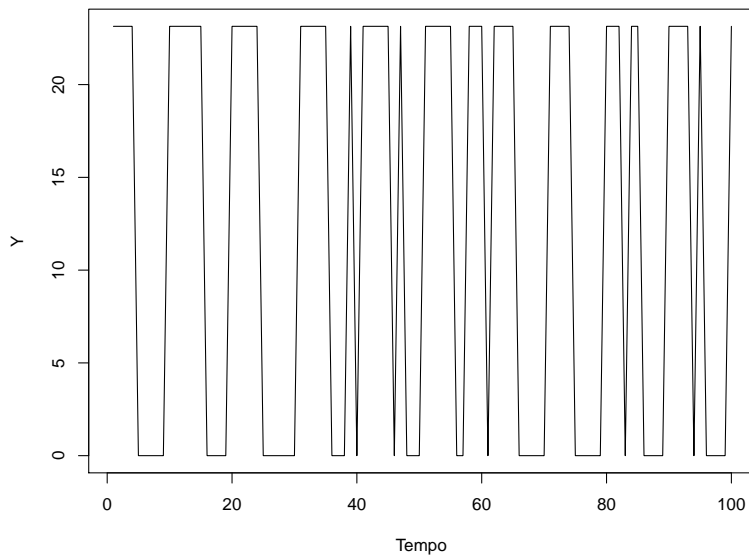


Figura 4.11: Série com tendência constante na parte contínua e sazonal na parte discreta

Na Figura 4.11, observamos claramente valores constante na parte contínua da série, $\exp(\pi) = 23,14069$, e ciclo sazonal na ocorrência de zeros. Para a análise dessa série foi usado os seis primeiros modelos, onde a Figura 4.12 apresenta as probabilidades a *posterioris*, via multiprocesso, para os modelo 1 ao 6.

Obsevamos, através da Figura 4.12, que o modelo com melhor desempenho é o Modelo 6, o qual é composto por um modelo polinomial de ordem 1 na parte contínua e modelo trigonométrico com $p = 10$ e harmônico $A=1$.

Tabela 4.6: Medidas de desempenho: Série com tendência constante (contínua) e sazonal (discreta).

| | MAD | MSE |
|----------|--------|---------|
| Modelo 1 | 7,775 | 119,985 |
| Modelo 2 | 7,021 | 106,405 |
| Modelo 3 | 11,103 | 229,555 |
| Modelo 4 | 10,174 | 198,543 |
| Modelo 5 | 9,215 | 161,510 |
| Modelo 6 | 4,834 | 70,112 |

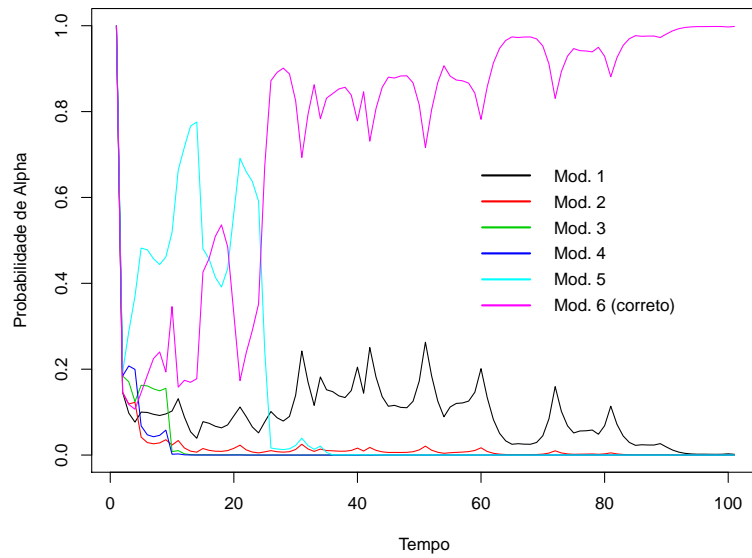


Figura 4.12: Multiprocesso Classe I: Série com tendência constante (contínua) e sazonal (discreta).

Observamos pela tabela das medidas de desempenho (Tabela 4.6), que as melhores medidas foram a do Modelo 6, logo, esse modelo mostrou-se superior aos demais.

4.7 Modelo 7

Uma série com dados contínuos com massa em zero, $Y_t, t = 1, 2, \dots, 100$, com tendência linear na parte contínua e sazonalidade com período 10 na parte discreta, pode ser gerada através da equação:

$$Y_t = \exp(0.03t)z_t,$$

com z_t sendo valores gerados de uma distribuição binomial com $p = \frac{\exp(u_t)}{1 + \exp(u_t)}$, onde $u_t = 3 \sin(2t\pi/10)$.

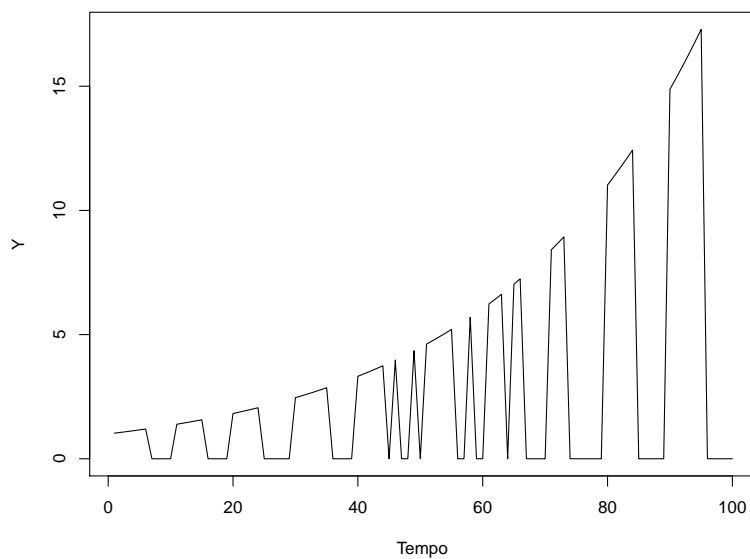


Figura 4.13: Série com tendência crescente na parte contínua e sazonal na discreta

Observamos, pela Figura 4.13, as tendências expressas no Modelo 7. Para o ajuste da série foram utilizados todos os modelos.

Pela Figura 4.14, observamos que a análise multiprocesso classe I apontou o Modelo 7 como o que teve a melhor performance na análise dessa série.

Através do MAD e MSE, o Modelo 3 mostrou-se superior aos demais, (Tabela 4.7).

4.8 Discussão

Nas seções anteriores, criamos séries temporais nas quais as componentes de tendência e sazonalidade, tanto na parte discreta quanto na contínua, podiam ser verificadas por inspeção visual.

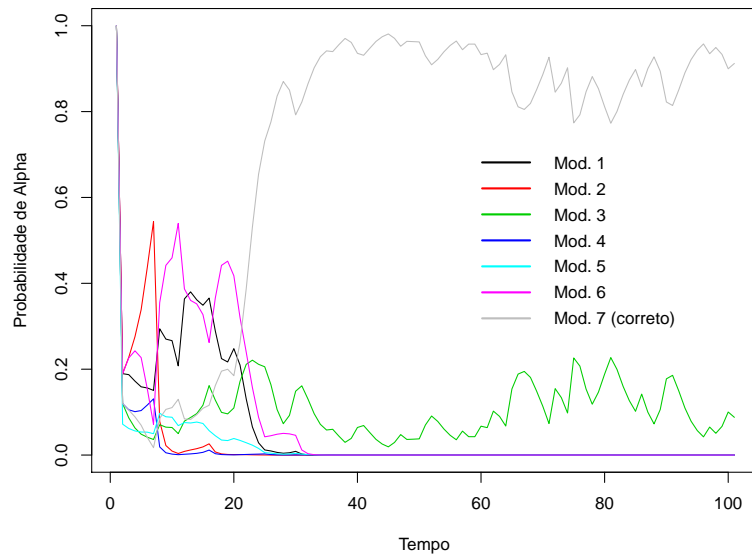


Figura 4.14: Multiprocesso Classe I: Série com tendência crescente (contínua) e sazonal (discreta)

Tabela 4.7: Medidas de desempenho: Série com tendência crescente (contínua) e sazonal (discreta)

| | MAD | MSE |
|----------|-------|--------|
| Modelo 1 | 1,669 | 11,362 |
| Modelo 2 | 1,580 | 12,774 |
| Modelo 3 | 0,662 | 1,269 |
| Modelo 4 | 0,751 | 4,438 |
| Modelo 5 | 1,622 | 10,238 |
| Modelo 6 | 1,315 | 7,648 |
| Modelo 7 | 0,889 | 5,693 |

Em seguida, aplicamos as seguintes ferramentas para discriminar os modelos: MAD, MSE e multiprocessos classe I.

Dos resultados expostos, levantamos evidências de que, dentre estas três, os multiprocessos são os mais qualificados para a identificação correta do modelo.

Capítulo 5

Aplicação em Séries Reais

Neste capítulo são apresentadas aplicações do modelo proposto a conjuntos de dados reais de precipitação pluviométrica diários proveniente dos municípios de Manaus/Am, Tarauacá/AC e São Gonçalo/PB. Estas séries apresentam a parte discreta, quando não chove em um dia, e a parte contínua, a quantidade que chove em um dia. Os dados foram coletados pelo Instituto Nacional de Meteorologia (INMET) e estão disponíveis no Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP).

Para escolhermos o modelo com melhor desempenho usaremos somente a análise multiprocesso classe I, pois as métricas de performances mostraram-se inapropriadas para a avaliação dos modelos.

5.1 Manaus

Serão analisados dois conjuntos de dados do município de Manaus. O primeiro possui 89 observações coletadas entre os dias 21 de fevereiro de 2015 à 21 de maio do mesmo ano o qual denominaremos de Manaus 1. O segundo são dados coletados durante o período de 3 anos, iniciando em janeiro de 2012, denominaremos por Manaus 2.

Na Figura 5.1 apresentamos a série pluviométrica de Manaus 1. São dados de um período chuvoso o que explica a ocorrência de 30,34% de dias não chuvosos.

Para essa série foram usados 4 modelos:

- **Modelo 1:** Mod.Pol_{cont}(1) e Mod.Pol_{disc}(1);
- **Modelo 2:** Mod.Pol_{cont}(1) e Mod.Pol_{disc}(2);
- **Modelo 3:** Mod.Pol_{cont}(2) e Mod.Pol_{disc}(1);
- **Modelo 4:** Mod.Pol_{cont}(2) e Mod.Pol_{disc}(2)

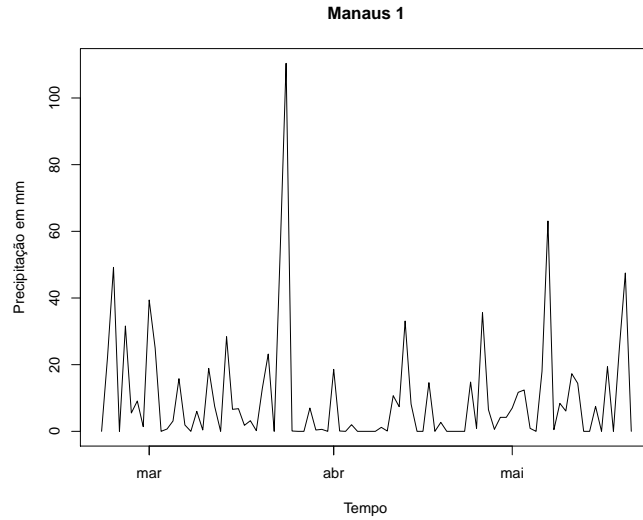


Figura 5.1: Série da Precipação de Manaus 1

Os parâmetros iniciais, para os modelos acima, foram: $\mathbf{m}_{1,0} = \mathbf{m}_{2,0} = 0$, $\mathbf{C}_{1,0} = \mathbf{C}_{2,0} = n_0 = 1 = S_0 = 1$ e $p_0 = 0,5$. Os fatores de descontos utilizados são dados na Tabela 5.1.

Tabela 5.1: Fatores de descontos para a Série de Manaus 1.

| | δ_{cont} | δ_{disc} |
|----------|-----------------|-----------------|
| Modelo 1 | 0,999 | 0,8 |
| Modelo 2 | 0,99 | 0,8 |
| Modelo 3 | 0,999 | 0,86 |
| Modelo 4 | 0,997 | 0,86 |

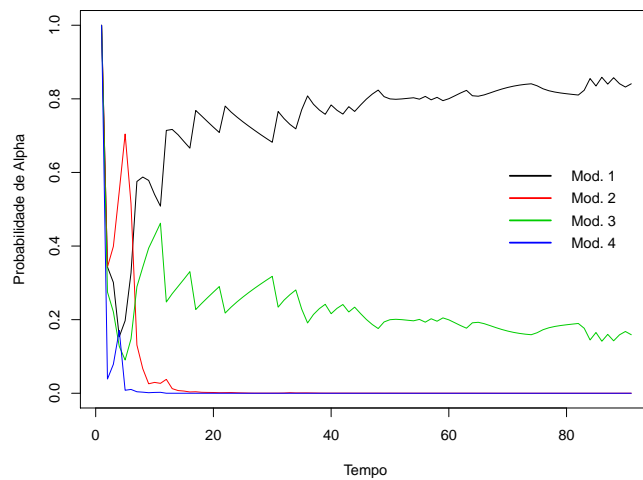


Figura 5.2: Multiprocesso para os quatro modelos ajustados em Manaus 1

Pelo multiprocesso (ver Figura 5.2), temos que o melhor modelo para esta série foi o Modelo 1, composto pelos modelos constantes em ambas as partes.

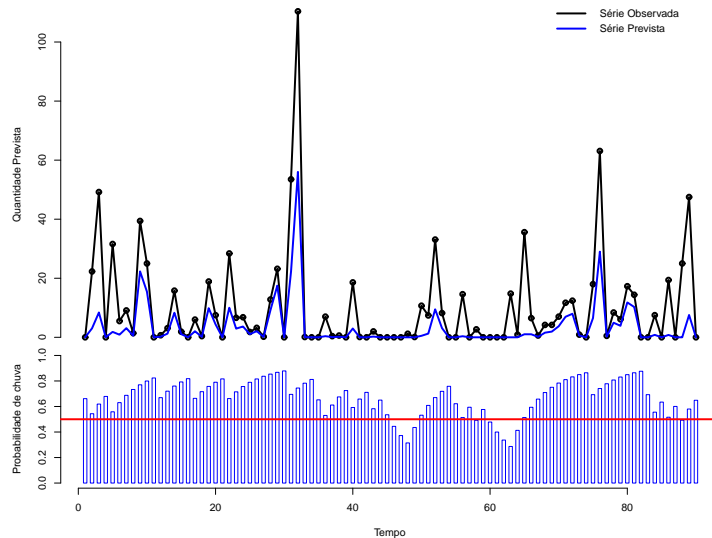


Figura 5.3: Modelagem da série usando o modelo proposto com $V_t = V$ desconhecido.

Na Figura 5.3, apresentamos os gráficos do modelo proposto com $V_t = V$ desconhecido aplicado na série de Manaus 1. No gráfico superior temos o ajuste do modelo proposto à série observada. Observamos que o modelo consegue identificar, em boa parte da série, os dias que ocorrem chuva ou não. Porém em dia que ocorre chuva o modelo subestima a quantidade de chuva. Isso pode ser explicado pelo fato de estarmos usando a mediana do modelo para prever a quantidade de chuva. No gráfico inferior temos a probabilidade *a priori* de chuva, no tempo t .

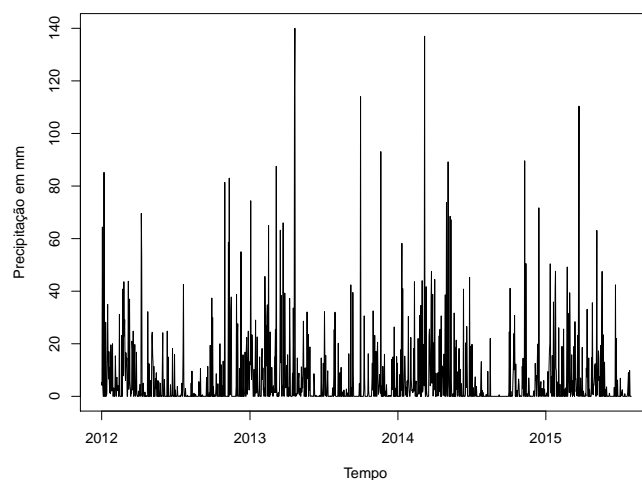


Figura 5.4: Precipitação no município de Manaus no período de 2012 a 2015.

Na Figura 5.4, apresenta-se a série temporal da precipitação pluviométrica de Manaus 2. Na figura podemos identificar períodos sazonais que é explicada pelas estações climáticas de Manaus: inverno e verão. O período que menos chove em Manaus, verão, se estende do início do mês de junho até final de novembro, enquanto o período chuvoso inicia em dezembro e se estende até maio (Carlo & Lamberts [3]).

Para a série de Manaus 2 foram ajustados 7 modelos:

- **Modelo 1:** Mod.Pol_{cont}(1) e Mod.Pol_{disc}(1);
- **Modelo 2:** Mod.Pol_{cont}(2) e Mod.Pol_{disc}(1);
- **Modelo 3:** Mod.Pol_{cont}(1) e Mod.Pol_{disc}(2);
- **Modelo 4:** Mod.Pol_{cont}(1)+Mod.Trig_{cont}(365,1) e Mod.Pol_{disc}(1);
- **Modelo 5:** Mod.Pol_{cont}(1) e Mod.Pol_{disc}(1)+Mod.Trig_{disc}(365,1);
- **Modelo 6:** Mod.Pol_{cont}(1) + Mod.Trig_{cont}(365,1) e Mod.Pol_{disc}(1) + Mod.Trig_{disc}(365,1);
- **Modelo 7:** Mod.Pol_{cont}(1) e Mod.Pol_{disc}(2)+Mod.Trig_{disc}(365,1).

Os parâmetros iniciais usados para modelar essa série foram os mesmos da série anterior. Os fatores de descontos seguem na Tabela 5.2.

Tabela 5.2: Fatores de descontos para a série de Manaus 2.

| | δ_{cont} | δ_{disc} |
|----------|-----------------|-----------------|
| Modelo 1 | 0,999 | 0,9 |
| Modelo 2 | 0,999 | 0,9 |
| Modelo 3 | 0,999 | 0,94 |
| Modelo 4 | 0,999;0,997 | 0,9 |
| Modelo 5 | 0,999 | 0,994;0,999 |
| Modelo 6 | 0,999;0,999 | 0,999;0,97 |
| Modelo 7 | 0,999 | 0,999;0,98 |

Pela Figura 5.5, observamos que o melhor modelo ajustado foi o Modelo 5. Apresentamos o gráfico do ajuste na Figura 5.6.

Na Figura 5.6, observa-se que o modelo se ajustou bem aos dados nos períodos mais chuvosos, mas nos períodos secos o modelo não obteve um bom desempenho, prevendo muitos dias secos quando na verdade há, pelo menos, pouca ocorrência de dias chuvosos. Como estamos trabalhando com a mediana, nosso modelo prevê zero se a probabilidade de chover for menor que meio, mas nos dias secos essas probabilidades ficam em torno de 0,3, pelo segundo gráfico da Figura 5.6, o que explica o modelo prever somente zeros nesse período. Portanto, para série que a

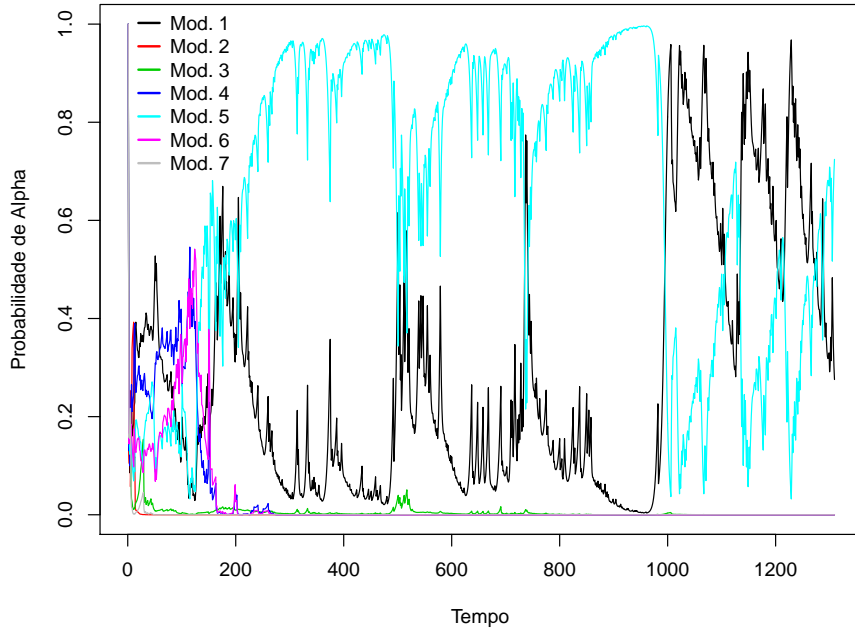


Figura 5.5: Multiprocesso para os modelos da série de Manaus 2.

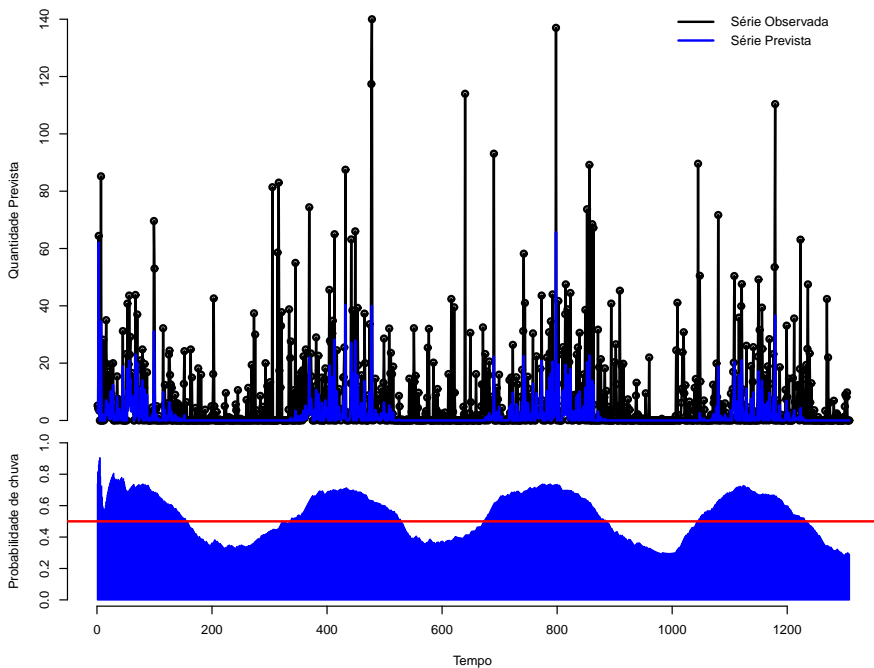


Figura 5.6: Ajuste do modelo à série de Manaus no período de 2012 a 2015.

proporção de zeros for maior que 50% a previsão pela mediana não é uma boa estratégia. Também, observamos que nosso modelo consegue identificar e modelar a sazonalidade, tanto na parte contínua, quanto na discreta pela probabilidade de

chuva, p_t .

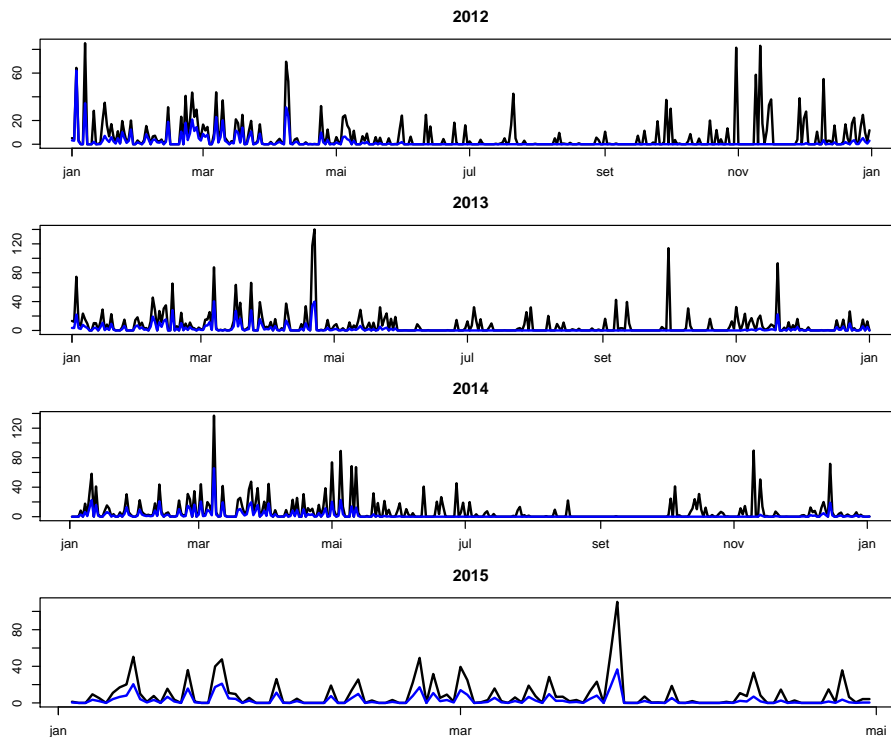


Figura 5.7: Ajuste do modelo à série de Manaus por ano.

5.2 Tarauacá

Tarauacá é um município do estado do Acre que faz fronteira com o estado do Amazonas. Seu clima é parecido com o do município de Manaus, pois ambos pertencem ao clima equatorial úmido. Na Figura 5.8, apresentamos a série pluviométrica deste município referente aos anos de 2012 à abril de 2015.

Esta série é caracterizada por grande ocorrência de chuvas no inverno e período com menor ocorrência de chuvas.

Para modelar esta série utilizamos os mesmos modelos usados para modelar a série de Manaus 2. Os parâmetros iniciais, foram: $\mathbf{m}_{1,0} = \mathbf{m}_{2,0} = 0$, $\mathbf{C}_{1,0} = \mathbf{C}_{2,0} = n_0 = 1 = S_0 = 1$ e $p_0 = 0,5$. Os fatores de descontos seguem na Tabela 5.2

Pelo multiprocessamento, Figura 5.9, o melhor modelo foi o Modelo 5, modelo composto por uma parte constante, polinomial de 1^a ordem, na parte contínua e na parte discreta pela superposição dos modelos constantes e trigonométrico.

O modelo proposto, com V desconhecido, ajustou-se bem à série de Tarauacá. Pela Figura 5.10, observa-se que o modelo consegue, no inverno, identificar os dias que chove e, diferente da série de Manaus, consegue identificar o quanto chove. Pela Figura 5.11 consegue-se ter uma melhor visualização da série estimada, no último

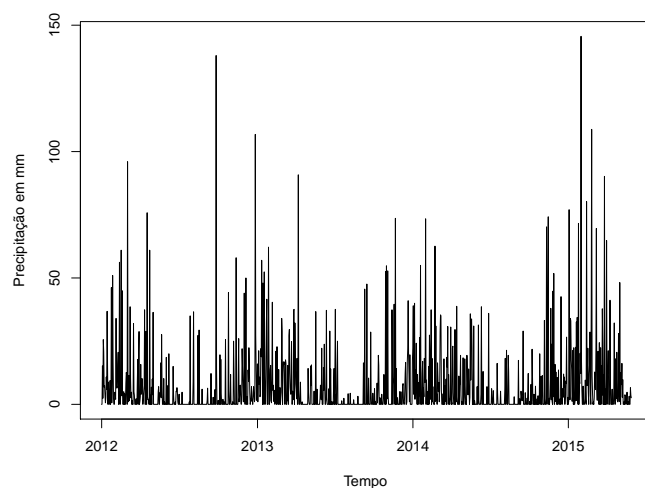


Figura 5.8: Série da precipitação pluviométrica de Tarauacá.

Tabela 5.3: Fatores de descontos para a Série de Tarauacá.

| | δ_{cont} | δ_{disc} |
|----------|-----------------|-----------------|
| Modelo 1 | 0,999 | 0,95 |
| Modelo 2 | 0,999 | 0,95 |
| Modelo 3 | 0,999 | 0,95 |
| Modelo 4 | 0,995;0.995 | 0,95 |
| Modelo 5 | 0,998 | 0,999;0,99 |
| Modelo 6 | 0,996;0.99 | 0,999;0,999 |
| Modelo 7 | 0,999 | 0,999;0,992 |

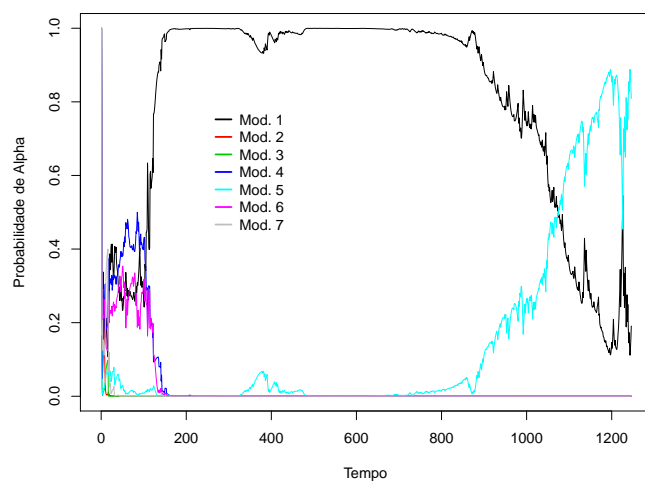


Figura 5.9: Multiprocesso - Tarauacá

gráfico temos a série observada, linha preta, e a série prevista, linha azul, para os meses de janeiro a março do ano de 2015.

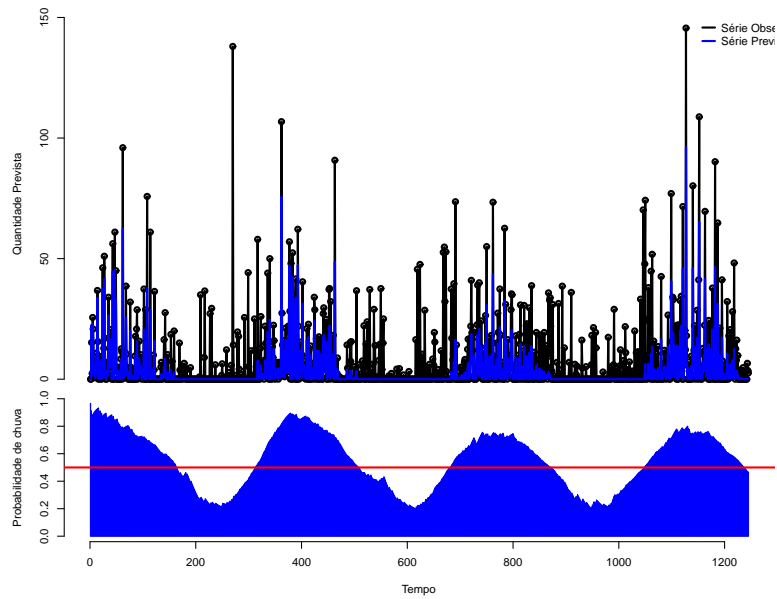


Figura 5.10: Ajuste do modelo à série de Tarauacá no período de 2012 a meado de 2015.

Porém, no período com menos ocorrência de chuva o modelo, também, não obteve bom desempenho. Para os valores de p_t a baixo da linha vermelha, segundo gráfico da Figura 5.10, o modelo prevê zero, para a observação correspondente ao tempo, como nos períodos de verão estes valores são baixo, então, para períodos onde $p_t < 0,5$ o nosso modelo terá problemas para modelar dias que chovem.

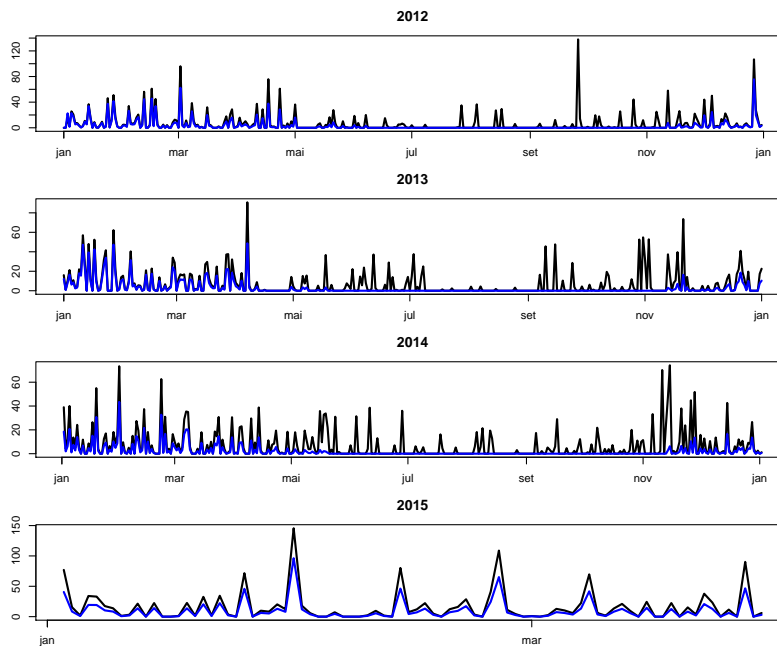


Figura 5.11: Ajuste do modelo à série de Tarauacá por ano.

5.3 São Gonçalo

São Gonçalo é um distrito do município de Sousa/PB. O clima de São Gonçalo é classificado como semiárido quente, caracterizado pela escassez e irregularidade de chuvas. Na Figura 5.13, apresenta-se a série pluviométrica desse distrito.

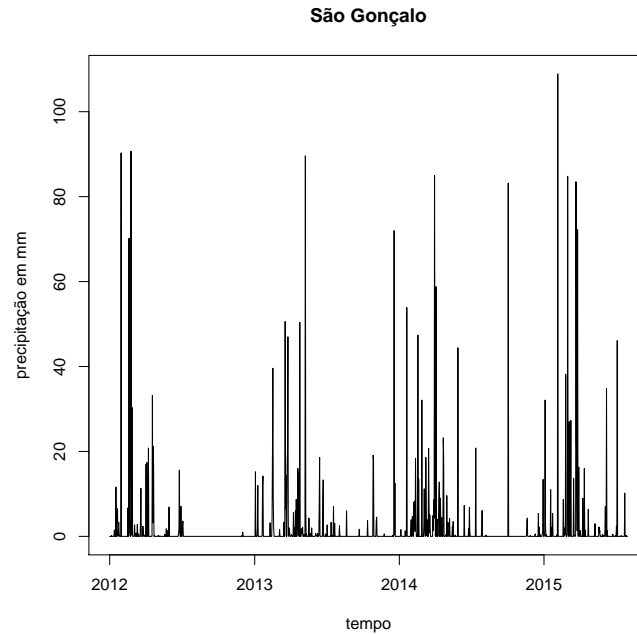


Figura 5.12: Série da precipitação pluviométrica de São Gonçalo.

Para essa série de São Gonçalo foram ajustados os mesmos modelos usados na série de Manaus 2, com parâmetros iniciais: $\mathbf{m}_{1,0} = \mathbf{m}_{2,0} = 0$, $\mathbf{C}_{1,0} = \mathbf{C}_{2,0} = n_0 = 1 = S_0 = 1$ e $p_0 = 0,5$, e fatores de descontos apresentados na Tabela 5.3.

Tabela 5.4: Fatores de descontos para a Série de São Gonçalo.

| | δ_{cont} | δ_{disc} |
|----------|-----------------|-----------------|
| Modelo 1 | 0,999 | 0,91 |
| Modelo 2 | 0,999 | 0,9 |
| Modelo 3 | 0,999 | 0,95 |
| Modelo 4 | 0,999;0,997 | 0,9 |
| Modelo 5 | 0,999 | 0,996;0,998 |
| Modelo 6 | 0,999;0,999 | 0,99;0,98 |
| Modelo 7 | 0,999 | 0,999;0,999 |

Pela Figura 5.13, o modelo que obteve o melhor desempenho, na série de São Gonçalo, foi o Modelo 4, pois obteve as melhores probabilidades a *posteriori*, via multiprocesso. Apresentamos o gráfico do ajuste na Figura 5.14.

Pela Figura 5.14, observa-se que o modelo não consegue identificar os dias que ocorrem chuvas, com exceção de alguns pontos. Isso ocorre devido as estimativas

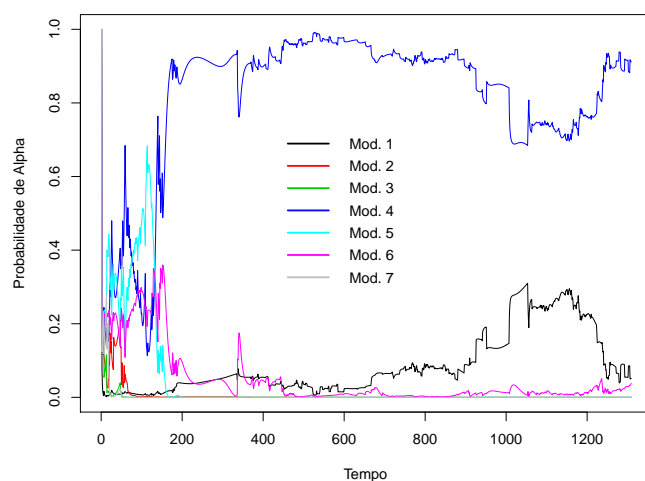


Figura 5.13: Multiprocesso para os modelos de São Gonçalo.

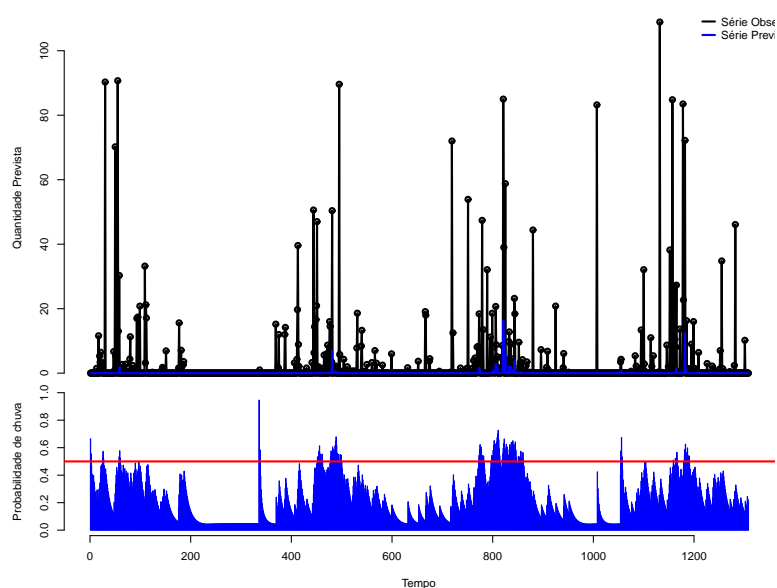


Figura 5.14: Ajuste do modelo à série de São Gonçalo no período de 2012 a meado de 2015

de p_t serem baixas, portanto, em séries em que a ocorrência de zeros for grande, maior que 50%, o modelo não terá um bom desempenho.

Capítulo 6

Considerações Finais e Trabalhos Futuros

Conseguimos através deste trabalho propor um modelo que é capaz de compreender e descrever o comportamento de séries temporais contínuas cujas observações são não negativas com massa em zero. O modelo é baseado nas metodologias dos MLD's e MLGD's o que nos permitiu fazer as inferências dos parâmetros de forma recursiva e com auxílio do filtro de Kalman. O modelo possui solução analítica o que descarta a necessidade de utilizar métodos baseados em simulações para obter aproximações dos momentos das densidades a *posterioris* dos estados, tal como o Método de Monte Carlo via Cadeias de Markov (MCMC).

Através do estudo das séries geradas observamos que o modelo proposto descreve eficientemente as tendências, tais como a constante, linear e sazonal, além de descrever as superposições. Através das séries simuladas a escolha do melhor modelo pode ser feita através dos modelos multiprocessos classe I.

A função de distribuição um passo a frente para o modelo com V desconhecido possui densidade Log-t que por sua vez não tem valor esperado. Devido a isso usamos a mediana como estatística de previsão um passo à frente.

Na aplicação aos dados reais o modelo proposto mostrou-se eficiente nos períodos chuvosos, conseguindo identificar os dias que não chovem e os dias que chovem. Nos dias que chovem identificou quanto chove, no caso do município de Tarauáca. Porém, nos períodos mais secos, o modelo mostrou uma restrição, não conseguindo ser eficiente nas previsões. Isso se deve ao fato de estarmos usando a mediana como estatística de previsão um passo à frente, pois se a probabilidade de chuva em um determinado tempo t for menor que meio o modelo considera a previsão relacionado a esse tempo como zero, assim, o modelo não terá bom desempenho em séries em que a proporção de zeros for maior que meio.

Para trabalhos futuros propomos a generalização da nova metodologia à séries temporais multivariadas contínuas com massa em zero. E também propomos a

utilizações de outros percentis, em vez da mediana, como estatística de previsão.

Referências Bibliográficas

- [1] Agresti, A., Min, Y., Modeling Nonnegative Data with Clumping at Zero: A Survey. *Journal of the Iranian Statistical Society*, 1(1 and 2): pp. 7-33, 2002.
- [2] Armstrong, J., Collopy, F., Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, (8):pp. 690-80, 1992.
- [3] Carlo, J., Lamberts, R., Estudos de Estratégias Bioclimáticas para a Cidade de Manaus. IX Encontro Nacional de tecnologia do Ambiente Construído, 2002.
- [4] Cassela, G., *Statistical Inference*. Duxbury Advanced Series. Duxbury Thomson Learning, 2002.
- [5] Correia, L. T., Modelos Dinâmicos para Dados Agregados. Master's Thesis, Universidade de Brasília, 2010.
- [6] Flores, B., A Pragmatic View of Accuracy Measurement in Forecasting. *International Journal Management Science*, 14(2):pp. 93-98, 1986.
- [7] Gomes, C. H. A., Estimação Bayesiana com Priori de Jeffreys em Modelos Espaço de Estado. Master's Thesis, Universidade Federal do Rio de Janeiro, 2006.
- [8] Harrison, P. J., Stevens, C. F., Bayesian Forecasting. *Journal of the Royal Statistical Society*, 38(3):pp. 205-247, 1976.
- [9] Hartigan, J.A., Linear Bayesian Methods. *Journal of Royal Statistical Society. Series B*, 31(3): pp. 446-454, 1969.
- [10] Hyndman, R., Koehler, A., Another Look at Measures of Forecast Accuracy. *International Journal of Forecasting*, (22):pp. 6790-688, 2006.
- [11] James, B. R., *Probabilidade: Um Curso em Nível Intermediário*. Projeto Euclides. Instituto de Matemática Pura e Aplicada, 2008.

- [12] Kalman, R. E., A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering, Série D*, 82: pp. 35-45, 1960.
- [13] Liu, J., West, M. Combined parameter and state estimation in simulation based filtering, in A. Doucet, N. De Freitas and N. Gordon (eds), *Sequential Monte Carlo methods in practice*, Springer, pp. 197-223, 2001.
- [14] Makridakis, S., Hibon, M., The M3 Competition: Results, Conclusions and Implications. *International Journal of Forecasting*, (16):pp. 451-476, 2000.
- [15] Marco, F. R. A, Gamerman, D., Dynamic Generalized Linear Model. *Biostatistics - Basel*, (5): pp. 57-72, 2000.
- [16] Marchenko, Y. V., Multivariate Skew-t distributions in Econometrics and Environmetrics. doctorate degree Dissertation, Texas A&M University, 2010.
- [17] Nelder, J.A, Wedderburn, R. W. M., Generalized linear models. *Journal of Royal Statistical Society. Series A*, 135(3): pp. 370-384, 1972.
- [18] Petris, G., Petrone, S., Campagnoli, P., *Dynamic Linear Model with R*. Springer, 2009.
- [19] Petris, G., An r package for Dynamic linear Models. *Journal of Statistical Software*, 36(12):pp. 1-16, 2010.
- [20] Silva, G.S., Um Modelo Linear com Abafador de tendência Dinâmico para Previsão Bayesiana para Séries Temporais. Master's Thesis, Universidade Federal do Amazonas, 2014.
- [21] Silva, C. Queiroz, Rodrigues, G. S., Migon, H. S., Bayesian Dynamic Dirichlet Models, *Communications in Statistics - Simulation and Computation*, 44(3): pp. 787-818, 2015.
- [22] Souza, P. O., Modelos Dinâmicos para a Distribuição Poisson Generalizada. Master's Thesis, universidade de Brasília, 2014.
- [23] Triantafyllopoulos, K., *Dynamic Generalized Linear Models for Non-Gaussian time series Forecasting*. Cornell university Library, 2008.
- [24] Triantafyllopoulos, K., Inference of Dynamic Generalized Linear Models: On-Line Computation and Appraisal. *International Statistical Institute*, 77(3): pp. 430-450, 2009.
- [25] West, M., Harrison, J., *Bayesian Forecasting and Dynamic Models*. 2nd Edition. Spriger Series in Statistics, 1997.

- [26] West, M., Harrison, P. J., Migon, H. S., Dynamic Generalized Linear Models and Bayesian Forecasting. *Journal American Statistical Association*, 80(389):73-83, 1985.

Apêndice A

Propriedades da Distribuição Normal

A.1 Transformações lineares:

Se $X \sim N_p(\mu, \Sigma)$, A é uma matriz $r \times p$ e b um vetor de dimensão $r \times 1$, então:

$$y = Ax + b \sim N(A\mu, A\Sigma A^T)$$

A.2 Distribuições Marginais

Se o vetor x é dividido em dois blocos, com x_1 contendo os primeiros r componentes e x_2 contendo os $p - r$ demais. procedendo as mesmas partições nos parâmetros μ e Σ , tem-se:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{e} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

As marginais são dadas por: $x_i \sim N(\mu_i, \Sigma_{ii})$, $i = 1, 2$.

A.3 Distribuições Condicionais

Considerando, ainda, as partições x_1 e x_2 de x , μ e Σ , segue que:

$$x_1|x_2 \sim N(\mu_{1.2}, \Sigma_{11.2})$$

onde

$$\mu_{1.2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2) \quad \text{e} \quad \Sigma_{11.2} = \Sigma_{11} + \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Para esses resultados, é obviamente necessário que as submatrizes Σ_{22} e Σ_{11} respectivamente tenham posto máximo, caso contrário suas inversas não existem.

A.4 Reconstrução da Conjunta

Se $x_1|x_2 \sim N(\mu_1 + B_1(x_2 - \mu_2), B_2)$ e $x_2 \sim N(\mu_2, \Sigma_{22})$, então:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N(\mu, \Sigma)$$

com

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad \text{e} \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

onde

$$\Sigma_{11} = B_2 + B_1 \Sigma_{22} B_1^T \quad \text{e} \quad \Sigma_{21}^T = \Sigma_{12} = B_1 \Sigma_{22}$$

Apêndice B

Modelo com Variância Desconhecida

Demonstração do Teorema 7:

a) Condicional a V :

É demonstrado pelos mesmos resultados do filtro de Kalman considerando a *posteriori* no tempo $t - 1$, $\boldsymbol{\theta}|Y_{0:t-1}, V \sim N[m_{t-1}, V\mathbf{C}_{t-1}^*]$, verdadeira.

b) Para a precisão:

Também é por indução, assumindo a priori $\phi|Y_{0:t-1} \sim G[n_{t-1}/2, d_{t-1}/2]$ verdadeira. Pelo teorema de Bayes $f(\phi|Y_{0:t}) \propto f(Y_t|\phi, Y_{0:t-1})$ onde $Y_t|Y_{0:t}, \phi$ tem distribuição Normal pela definição do modelo e, por hipótese, $\phi|Y_{0:t-1}$ tem distribuição Gama, logo

$$\begin{aligned} f(\phi|Y_{0:t}) &\propto \phi^{1/2} \exp\left\{-\frac{\phi e_t^2}{2Q_t^*}\right\} \phi^{\frac{n_{t-1}}{2}-1} \exp\left\{-\frac{\phi d_{t-1}}{2}\right\} \\ &\propto \phi^{\frac{n_{t-1}+1}{2}-1} \exp\left\{-\left(d_{t-1} + \frac{e_t^2}{Q_t^*}\right) \frac{\phi}{2}\right\} \\ &\propto \phi^{n_t/2} \exp\{-d_t \phi/2\}, \end{aligned}$$

é o núcleo de uma distribuição Gama com parâmetros $n_t/2$ e $d_t/2$, onde $n_t = n_{t-1} + 1$ e $d_t = d_{t-1} + e_t^2 Q_t^{*-1}$.

c) Incondicional a V :

Segue diretamente da teoria da Normal Gama onde para um vetor aleatório $\boldsymbol{\theta}$ ($n \times 1$) com $\boldsymbol{\theta}|\phi \sim N[\mathbf{m}, \phi^{-1}\mathbf{C}^*]$ e $\phi \sim G[n/2, d/2]$, então $\boldsymbol{\theta} \sim t_n[\mathbf{m}, S\mathbf{C}^*]$ onde $S = d/n$.

Apêndice C

Superposição de Modelos

Demonstração do Teorema 4:

Para o vetor de observação $Y_t = \mathbf{F}_t^T \boldsymbol{\theta}_t + \nu_t$, onde $\mathbf{F}_t^T = (\mathbf{F}_{1t}^T, \dots, \mathbf{F}_{ht}^T)$, $\boldsymbol{\theta}_t^T = (\boldsymbol{\theta}_{1t}^T, \dots, \boldsymbol{\theta}_{ht}^T)$ e $\nu_t = \sum_{i=1}^h \nu_{it}$ onde $\nu_{it} \sim N(\mathbf{0}, V_i)$ e independentes entre si. Temos, pelas propriedades da normal, que $\nu_t \sim N(\mathbf{0}, V_t)$ com $V_t = \sum_{i=1}^h V_i$. Para o vetor de estado $\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t$, onde $\boldsymbol{\omega}_t^T = (\boldsymbol{\omega}_{1t}^T, \dots, \boldsymbol{\omega}_{ht}^T)$ com $\boldsymbol{\omega}_{it} \sim N(\mathbf{0}, \mathbf{W}_{it})$ e independentes entre si, temos que $\boldsymbol{\omega}_t \sim N(\mathbf{0}, \mathbf{W}_t)$ e independente de ν_t . Assim, temos definido o MLD.