

Luiza Barbosa Amorim Ferreira

Um estudo comparativo para modelos de
séries temporais de contagem

Belo Horizonte
2015

Luiza Barbosa Amorim Ferreira

Um estudo comparativo para modelos de
séries temporais de contagem

Dissertação apresentada ao Instituto
de Ciências Exatas da Universidade
Federal de Minas Gerais, para a ob-
tenção de Título de Mestre em Esta-
tística, na Área de Séries Temporais.

Orientadora: Glaura da Conceição
Franco

Co-orientador: Frank Magalhães de
Pinho

Belo Horizonte
2015

Luiza Barbosa Amorim Ferreira.

Um estudo comparativo para modelos de séries temporais de contagem

85 páginas

Dissertação (Mestrado) - Instituto de Ciências Exatas da Universidade Federal de Minas Gerais. Departamento de Estatística.

1. Processos Autoregressivos
2. Processos Médias Móveis
3. Modelo Linear Generalizado

I. Universidade Federal de Minas Gerais. Instituto de Ciências Exatas. Departamento de Estatística.

Comissão Julgadora:

Prof. Dr.

Marcos Oliveira Prates

Prof. Dr.

Luis Mauricio Castro Cepero

Prof. Dr.

Frank Magalhães de Pinho

Prof. Dra.

Glaura da Conceição Franco

Agradecimentos

À Deus por mais esta vitória em minha vida.

À Nossa Senhora Aparecida pela poderosa intercessão nessa jornada.

Aos meus pais, Sérgio e Suely, e minha irmã Mariana por acreditarem e me apoiarem para tornar esse projeto possível.

Ao meu marido, amigo e companheiro José Carlos, que sempre esteve ao meu lado pacientemente dando força e incentivo.

À toda minha família, pelas orações e torcida.

Aos meus amigos da pós graduação, em especial Rachel, Renata, Larissa, Marcela e Maurício, vocês foram imprescindíveis para a conclusão deste trabalho.

A minha orientadora professora Glaura e co-orientador Frank pela orientação, dedicação e conhecimentos compartilhados.

Resumo

Neste trabalho, duas metodologias de séries temporais de dados de contagem são avaliadas, o Modelo Autoregressivo Média Móvel Generalizado (GARMA) e o Modelo Autoregressivo Média Móvel Linear Generalizado (GLARMA). O objetivo principal é analisar a qualidade do ajuste dos modelos em questão, através de algoritmos implementados em linguagem R. Um outro objetivo é comparar estes modelos ao Modelo Linear Generalizado (MLG), que permite o ajuste de dados não-Gaussianos, mas não leva em consideração a dependência temporal existente neste tipo de observações. Um estudo de simulação é realizado afim de verificar o comportamento das estimativas. Duas aplicações a séries reais são realizadas, o número de empresas que decretaram falência nos Estados Unidos nos anos de 1985 a 2012, e número de casos mensais de poliomelite em um hospital. Os modelos utilizados descrevem bem o comportamento das séries.

Palavras-chave: Processos Autoregressivos, Processos Médias Móveis, Modelo Linear Generalizado, Distribuição Poisson.

Abstract

In this work, two methods for time series of counts are evaluated, the Autoregressive Moving Average Generalized Model (GARMA) and the Autoregressive Moving Average Model Generalized Linear Model (GLARMA). The main objective is to analyze the quality of fit of the above models, using algorithms implemented in the R language. Another objective is to compare these models to the Generalized Linear Model (GLM), which allows the fit of non Gaussian observations, but does not take into account the time dependence of such data. A simulation study is conducted in order to verify the behavior of the estimates. Two applications to real series are performed, the number of companies that went bankrupt in the United States in the years 1985-2012, and monthly number of cases of polio in a hospital. The models describe satisfactorily the behavior of the series.

Keywords: Autoregressive processes, Moving Average processes, Generalized Linear Model, Poisson Distribution

Lista de Figuras

1.1	Série temporal do número de casos de poliomelite	2
1.2	Autocorrelação e Autocorrelação parcial dos resíduos do modelo ajustado à série de número de casos de poliomelite	3
4.1	Boxplot das estimativas para séries geradas pelo GARMA AR(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro ϕ	25
4.2	Boxplot das estimativas para séries geradas pelo GARMA AR(1) com covariável tendência. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 e os da terceira linha para ϕ	27
4.3	Boxplot para séries geradas pelo GARMA MA(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro θ	29

4.4	Boxplot para séries geradas pelo GARMA MA(1) com covariável	30
4.5	Boxplot para séries geradas pelo GARMA MA(1) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro β_1 e os da terceira linha para o parâmetro θ	30
4.6	Boxplot para séries geradas pelo GARMA AR(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro ϕ_1 e os da terceira para ϕ_2	32
4.7	Boxplot para séries geradas pelo GARMA AR(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira para parâmetro ϕ_1 e os da quarta para o parâmetro ϕ_2	33
4.8	Boxplot para séries geradas pelo GARMA MA(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro θ_1 e os da terceira para o parâmetro θ_2	35

4.9	Boxplot para séries geradas pelo GARMA MA(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira para o parâmetro θ_1 e os da quarta linha para o parâmetro θ_2 . . .	36
4.10	Boxplot para séries geradas pelo GLARMA AR(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro ϕ	38
4.11	Boxplot para séries geradas pelo GLARMA AR(1) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 e os da terceira para o parâmetro ϕ	39
4.12	Boxplot para séries geradas pelo GLARMA MA(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro θ	40
4.13	Boxplot para séries geradas pelo GLARMA MA(1) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 e os da terceira linha para θ	42

4.14	Boxplot para séries geradas pelo GLARMA AR(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro ϕ_1 e os da terceira para o parâmetro ϕ_2	44
4.15	Boxplot para séries geradas pelo GLARMA AR(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira para o parâmetro ϕ_1 e os da quarta linha para o parâmetro ϕ_2	45
4.16	Boxplot para séries geradas pelo GLARMA MA(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro θ_1 e os da terceira para o parâmetro θ_2	47
4.17	Boxplot para séries geradas pelo GLARMA MA(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira linha para o parâmetro θ_1 e os da quarta linha para o parâmetro θ_2	48
5.1	Autocorrelação e autocorrelação parcial da série dos casos de poliomelite	51
5.2	Análise dos resíduos para o modelo GLARMA AR(1) ajustado à série de número mensal de casos de poliomelite	53

5.3	Autocorrelograma e autocorrelograma parcial dos resíduos do ajuste à série de número mensal de casos de poliomelite	54
5.4	Série temporal do número mensal de casos de poliomelite. A linha contínua representa a série real e a linha tracejada o modelo GLARMA AR(1) ajustado à série.	55
5.5	Série temporal do número de casos de falências de empresas Norte Americanas	58
5.6	Função de autocorrelação e autocorrelação parcial do número de empresas que decretaram falência.	59
5.7	Resíduos GLARMA AR(3)	61
5.8	FAC e FACP GLARMA AR(3)	62
5.9	Série temporal do número de casos de falência. A linha contínua representa a série real e a linha tracejada o modelo GLARMA AR(3) ajustado à série.	63
7.1	Resíduos do ajuste MLG dos casos de poliomelite	67

Lista de Tabelas

1.1	Ajuste MLG para os dados de Poliomelite	3
5.1	AIC dos modelos ajustados para os dados de casos de poliomielite	52
5.2	Ajuste GLARMA AR(1) para os dados de Poliomelite	52
5.3	Previsão para dados de poliomelite utilizando os modelos GLARMA AR(1), GARMA e MLG	56
5.4	AIC dos modelos ajustados à série do número de empresas Norte Ameri- canas que decretaram falência	59
5.5	Ajuste GLARMA AR(3) para o número de casos de falências de empresas Norte Americanas	60
5.6	Previsão para dados de falência utilizando os modelos GLARMA AR(3) e GARMA AR(3)	64
7.1	Estimativas para séries geradas pelo GARMA AR(1) sem covariável	68
7.2	Estimativas para séries geradas pelo GARMA AR(1) com covariável ten- dência	69
7.3	Estimativas para séries geradas pelo modelo GARMA MA(1) sem covariável	70
7.4	Estimativas para séries geradas pelo modelo GARMA MA(1) com cova- riável tendência	71
7.5	Estimativas para séries geradas pelo modelo GARMA AR(2) sem covariável	72

7.6	Monte Carlo para GARMA AR(2) com covariável tendência	73
7.7	Estimativas para séries geradas pelo modelo GARMA MA(2) sem covariável	74
7.8	Estimativas para séries geradas pelo modelo GARMA MA(2) com covariável tendência	75
7.9	Estimativas para séries geradas pelo modelo GLARMA AR(1) sem covariável	76
7.10	Estimativas para séries geradas pelo modelo GLARMA AR(1) com covariável tendência	77
7.11	Estimativas para séries geradas pelo modelo GLARMA MA(1) sem covariável	78
7.12	Estimativas para séries geradas pelo modelo GLARMA MA(1) covariável tendência	79
7.13	Estimativas para séries geradas pelo modelo GLARMA AR(2) sem covariável	80
7.14	Estimativas para séries geradas pelo modelo para GLARMA AR(2) com covariável tendência	81
7.15	Estimativas para séries geradas pelo GLARMA MA2 sem covariável	82
7.16	Estimativas para séries geradas pelo GLARMA MA2 com covariável tendência	83

Sumário

Lista de Figuras	vi
Lista de Tabelas	xi
Sumário	xiii
1 Introdução	1
2 Revisão de Literatura	6
2.1 Processo Linear	6
2.2 Estacionariedade e Invertibilidade	7
2.3 Modelo Autoregressivo Média Móvel	8
2.3.1 Previsão nos modelos ARMA	10
2.4 Modelos Lineares Generalizados	11
3 Modelos para séries de contagem	15
3.1 Modelo Autoregressivo Média Móvel Generalizado (GARMA)	16
3.1.1 Estimação nos modelos GARMA	17
3.1.2 Propriedades dos modelos GARMA	17
3.1.3 Previsão no modelo GARMA	18
3.2 Modelo Autoregressivo Média Móvel Linear Generalizado (GLARMA) . .	19

3.2.1	Estimação nos modelos GLARMA	20
3.2.2	Propriedades dos modelos GLARMA	21
3.2.3	Previsão no modelo GLARMA	21
4	Estudo de simulação	23
4.1	Séries geradas pelo modelo GARMA	24
4.1.1	GARMA(1,0)	24
4.1.2	GARMA(0,1)	28
4.1.3	GARMA(2,0)	31
4.1.4	GARMA(0,2)	34
4.2	Séries geradas pelo modelo GLARMA	37
4.2.1	GLARMA(1,0)	37
4.2.2	GLARMA(0,1)	40
4.2.3	GLARMA(2,0)	43
4.2.4	GLARMA(0,2)	46
5	Aplicação a dados reais	50
5.1	Poliomelite	50
5.2	Falência de empresas Norte Americanas	57
6	Conclusão e considerações finais	65
7	Anexo	67
7.1	Anexo 1	67
7.2	Anexo 2	68
	Referências Bibliográficas	84

Capítulo 1

Introdução

Séries temporais para dados de contagem são registros da frequência relativa da ocorrência de determinados eventos em sucessivos intervalos de tempo, e tem como característica importante a dependência entre as observações. Elas surgem nas mais variadas áreas de aplicação tais como indústria, medicina, economia e fenômenos meteorológicos.

Na literatura aparecem em diversos trabalhos, tais como no estudo dos casos de poliomelite nos Estados Unidos com Zeger (1988), também com os casos diários de asma em um hospital de Sydney, no estudo de Davis & Wang (2000) e o número de delitos registrados em Vitória com Silva & Reisen (2011).

Se o interesse é ajustar um modelo de regressão para uma série temporal de dados de contagem é necessário utilizar uma distribuição de probabilidade discreta para inteiros não negativos, como a Poisson ou a Binomial Negativa, já que a variável resposta é uma contagem.

O modelo linear generalizado (MLG), introduzidos por Nelder & Wedderburn em 1972, é uma opção para essa modelagem. Este modelo é uma generalização do modelo linear clássico proposto no século XIX por Gauss (1809) e Legendre (1805) e relaciona a média da variável resposta com as covariáveis através de uma função de ligação. O MLG abriu um leque de opções para a variável resposta, permitindo que a mesma seja da família exponencial. Desta forma dados cujas distribuições são dadas pela Poisson, Binomial Negativa, Normal, Gama, entre outras, puderam ser modeladas. Este modelo

tem sido muito utilizado em aplicações práticas com séries de contagens como no estudo de Tao et al (2011) à cerca dos efeitos agudos de mortalidade de monóxido de carbono na China, e também por Veselý et al (2009) sobre a poluição do ar por partículas em suspensão na República Checa.

Para motivar a utilização dos modelos que serão abordados neste trabalho, inicialmente será apresentada a modelagem dos dados de poliomielite, introduzidos no estudo de Zeger (1988), utilizando-se o MLG.

Os dados são referentes ao número mensal de casos de poliomielite nos EUA nos anos de 1970 a 1983 conforme relatado pelo Centro de Controle de Doenças. São 168 dados com uma variação de 0 a 14 casos, ver Figura 1.1. É possível observar que os dados são inteiros não negativos, já que trata-se de uma série de contagem, por isso é necessária a utilização de uma distribuição de probabilidade discreta, como a Poisson.

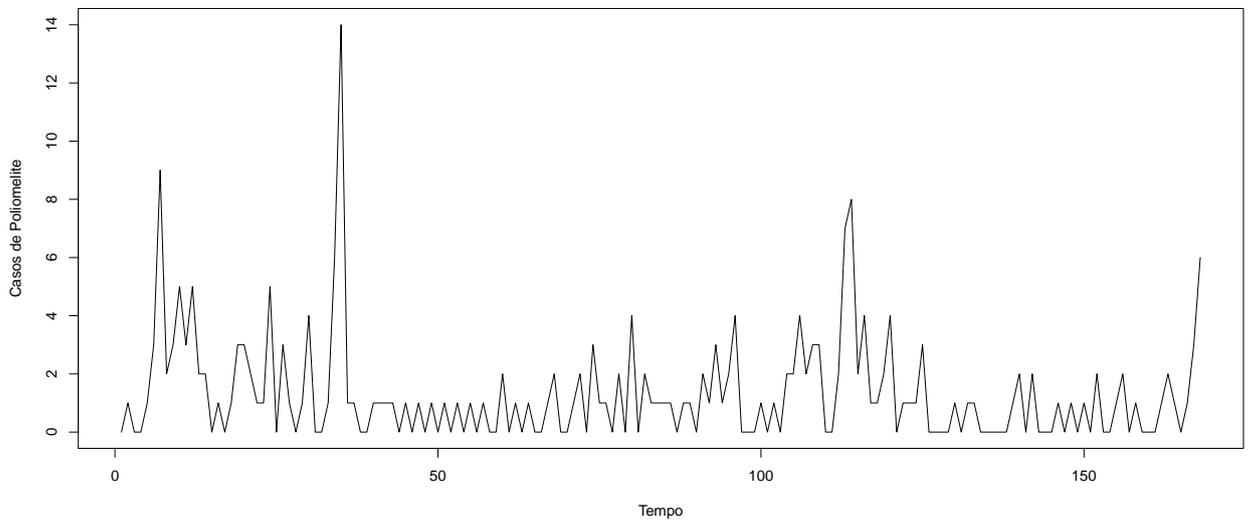


Figura 1.1: Série temporal do número de casos de poliomielite

Seguindo a sugestão do artigo de Zeger (1988), para a modelagem foram utilizadas quatro covariáveis para modelar a sazonalidade (seno e cosseno, anual e semestral) e um componente de tendência linear. O Modelo Linear Generalizado ajustado é apresentado na Tabela 1.1.

Tabela 1.1: Ajuste MLG para os dados de Poliomelite

	Estimativa	Erro padrão	p-valor
Intercepto	0,207	0,075	0,006
Tendência	-4,799	1,402	0,000
Cosseno anual	-0,148	0,097	0,126
Seno anual	-0,531	0,109	0,000
Cosseno semestral	0,169	0,099	0,087
Seno semestral	-0,432	0,100	0,000

O gráfico de resíduos é apresentado no ANEXO 7.1 e é possível observar alguns outliers, mas a suposição de homocedasticidade parece estar satisfeita, já que os resíduos comportam-se aleatoriamente em torno de zero.

Além da suposição de homocedasticidade é necessário investigar também se os resíduos são não autocorrelacionados. Na Figura 1.2 seguem os gráficos de autocorrelação (FAC) e autocorrelação parcial (FACP) dos resíduos do MLG ajustado.

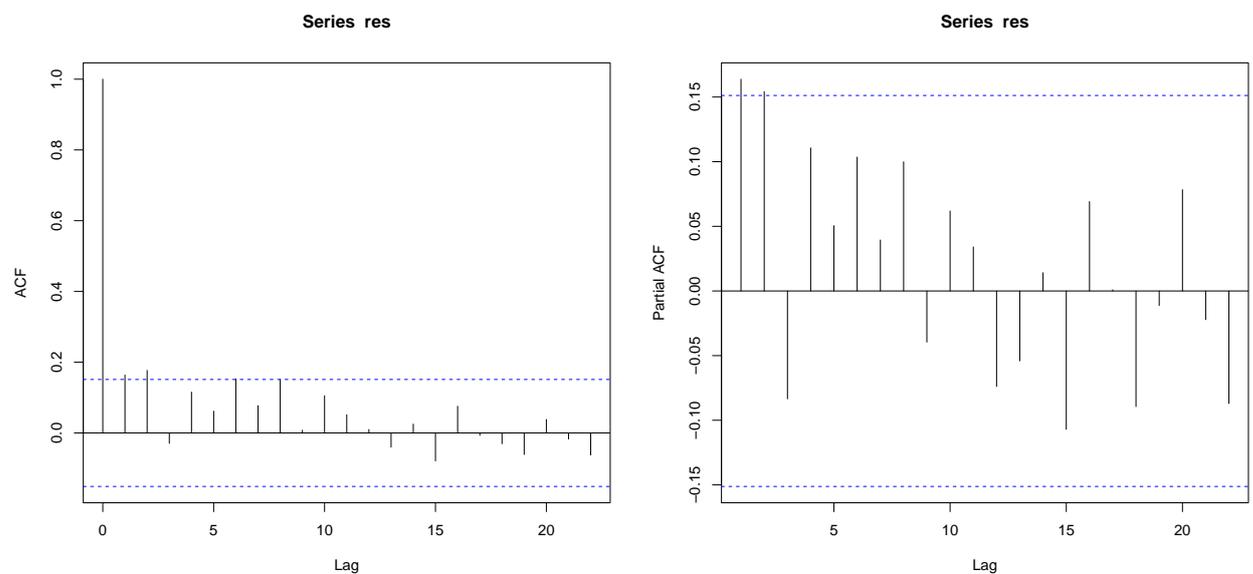


Figura 1.2: Autocorrelação e Autocorrelação parcial dos resíduos do modelo ajustado à série de número de casos de poliomelite

É possível perceber dois picos nos lags 1 e 2 da FAC e da FACP, indicativo de que

este modelo não foi capaz de eliminar a autocorrelação existente nos dados, já que não leva em conta a dependência temporal dos mesmos.

Uma opção para melhorar o ajuste, visto que o MLG não é capaz de capturar a dependência no tempo destes dados, seria o uso de modelos para séries temporais. Um dos modelos mais utilizados é o modelo Autoregressivo Média Móvel (ARMA), proposto por Box e Jenkins (1976). Entretanto, este modelo utiliza a suposição de que a série tem distribuição de probabilidade conjunta Gaussiana e no estudo de séries de contagem, para um bom ajuste dos dados, é necessário uma distribuição de probabilidade discreta.

Mais recentemente, foram propostos novos procedimentos para a modelagem de séries temporais, que relacionam os modelos MLG e ARMA. Nestas propostas é acrescido ao MLG uma estrutura autoregressiva média móvel, afim de modelar dados que tenham distribuição pertencente à família exponencial e apresentem dependência temporal. Davis et al (2003) propuseram o Modelo Autoregressivo Média Móvel Linear Generalizado (GLARMA) e Benjamin et al (2003) propuseram o Modelo Autoregressivo Média Móvel Generalizado (GARMA), que são extensões do MLG.

A vantagem em acrescentar termos autoregressivo e média móvel à estrutura do modelo linear generalizado é poder modelar séries temporais, já que estes componentes capturam a dependência temporal dos dados. Além disto, há o ganho em se modelar séries não Gaussianas, com a estrutura mais flexível do MLG.

Existem, na literatura, outros procedimentos para modelar séries com distribuição Não-Gaussiana, como o modelo INAR, proposto por Al-Osh & Alzaid (1987), onde é um processo de contagem que considera variáveis inteiras não negativas, ou os modelos dinâmicos generalizados, abordados por vários autores, tais como Gamerman et al (2013), Harvey & Fernandes (1989) e West et al (1985).

Um dos objetivos deste trabalho é ampliar o conhecimento dos modelos GARMA e GLARMA, avaliando as estimativas dos parâmetros obtidas por séries temporais geradas pelos modelos GARMA e GLARMA, através de um extenso estudo de simulação. Um segundo objetivo é modelar e fazer previsões para séries reais usando os modelos

GARMA e GLARMA, e comparar estes resultados à modelagem obtida pelo MLG.

As próximas seções estão estruturadas da seguinte forma: na Seção 2 é feita uma revisão dos modelos básicos: ARMA e MLG. Na Seção 3 é feita uma revisão dos modelos GLARMA e GARMA. Na Seção 4 é apresentado um estudo de simulação a fim de verificar o desempenho dos estimadores dos parâmetros, na Seção 5 análise de dados reais são discutidos e na Seção 6 a conclusão e considerações finais.

Capítulo 2

Revisão de Literatura

Neste capítulo serão apresentados alguns conceitos básicos para analisar uma série temporal, assim como uma descrição resumida do modelo Modelo Autoregressivo Média Móvel (ARMA) e Modelo Linear Generalizado (MLG).

2.1 Processo Linear

De acordo com Box & Jenkins (1976) o processo linear é baseado na idéia de que uma série temporal y_t com sucessivos valores altamente dependentes pode ser gerada de uma série de choques independentes u_t . Estes choques têm uma distribuição usualmente Normal e são chamados de ruídos brancos. Segundo Morettin & Tolo (2006)

$$E(u_t) = 0,$$

e

$$\gamma_k = E(u_t, u_{t+k}) = \begin{cases} \sigma^2 & \text{se } k = 0. \\ 0 & \text{se } k \neq 0 \end{cases}$$

O ruído branco u_t é transformado no processo y_t pelo que é chamado filtro linear, $\psi(B)$. A operação simplesmente é uma soma ponderada das observações anteriores, de modo que

$$y_t = \mu + u_t + \psi_1 u_{t-1} + \psi_2 u_{t-2} + \dots = \mu + \psi(B)u_t. \quad (2.1)$$

sendo μ a média do processo e B o operador de retardo que representa uma defasagem de k períodos de tempo atrás, definido por:

$$B^k u_t = u_{t-k}. \quad (2.2)$$

O modelo dado em 2.1 também pode ser escrito como a soma ponderada dos valores passados de y_t adicionando u_t , então

$$y_t = \mu + \pi_1 y_{t-1} + \pi_2 y_{t-2} + \dots + u_t. \quad (2.3)$$

Analogamente, pode-se escrever

$$u_t = y_t - \pi_1 y_{t-1} - \pi_2 y_{t-2} - \dots - \mu = \pi(B)y_t - \mu. \quad (2.4)$$

2.2 Estacionariedade e Invertibilidade

Estacionariedade

Uma das suposições mais frequentes na modelagem de séries temporais é que a série seja estacionária, no qual o processo permanece em equilíbrio sobre uma média constante. De acordo com Harvey (1993) um processo é estacionário se as seguintes condições são satisfeitas para todos os valores de t :

$$E(y_t) = \mu, \quad (2.5)$$

$$E[(y_t - \mu)^2] = \sigma^2, \quad (2.6)$$

$$\gamma_k = \text{cov}[y_t, y_{t+k}] = E[(y_t - \mu)(y_{t+k} - \mu)]. \quad (2.7)$$

Similarmente a autocorrelação de lag k é

$$\rho_k = \frac{E[(y_t - \mu)(y_{t+k} - \mu)]}{\sqrt{E[(y_t - \mu)^2]E[(y_{t+k} - \mu)^2]}} = \frac{\gamma_k}{\gamma_0}, k = 0, \pm 1, \pm 2, \dots \quad (2.8)$$

Um processo estocástico $y_t = \psi(B)u_t$ será estacionário se

$$\psi(B) = \sum_{k=0}^{\infty} \psi_k B^k,$$

converge para $|B| < 1$.

Invertibilidade

De acordo com Box & Jenkins (1976) se existe o interesse em fazer previsões com um modelo é necessário que ele seja invertível.

Um processo estocástico y_t será invertível se

$$\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j,$$

converge para $|B| < 1$.

2.3 Modelo Autoregressivo Média Móvel

O modelo Autoregressivo Média Móvel (ARMA) proposto por Box & Jenkins(1976) é definido por

$$\tilde{y}_t = \phi_1 \tilde{y}_{t-1} + \dots + \phi_p \tilde{y}_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-p}, \quad t = 1, \dots, n, \quad (2.9)$$

ou

$$\phi(B)\tilde{y}_t = \theta(B)u_t, \quad (2.10)$$

onde

$$\tilde{y}_t = y_t - \mu, \phi(B) = (1 - \phi B - \dots - \phi_p B^p) \text{ e } \theta(B) = (1 - \theta B - \dots - \theta_q B^q).$$

De 2.10 tem-se

$$\tilde{y}_t = \psi(B)u_t = \frac{\theta(B)}{\phi(B)}u_t. \quad (2.11)$$

O modelo em 2.9 é conhecido como Autoregressivo Média Móvel de ordem p, q , ou ARMA (p, q) .

A metodologia Box & Jenkins está fundamentada na identificação de um modelo adequado para a série, estimação dos parâmetros, verificação do modelo e previsão de valores futuros da série.

A identificação do modelo ARMA (p, q) se faz através das funções de autocorrelação (FAC), dada em 2.8, e autocorrelação parcial (FACP), que é a correlação entre duas observações seriais, eliminando a dependência dos termos intermediários.

Para identificar os valores de p e q é analisado o comportamento das autocorrelações e das autocorrelações parciais amostrais e comparado com a função de autocorrelação e autocorrelação parcial dos modelos teóricos ARMA (p, q) . Maiores detalhes podem ser vistos em Box & Jenkins (1976).

Após a identificação dos modelos, deve-se estimar os parâmetros. Neste caso os parâmetros a serem estimados são $\phi = (\phi_1, \phi_2, \dots, \phi_p)$, $\theta = (\theta_1, \theta_2, \dots, \theta_q)$ e $\sigma^2 = Var(u_t)$. Um dos métodos mais utilizados para estimar os parâmetros é o método da máxima verossimilhança.

Para o modelo ARMA (p, q) , os resíduos podem ser escritos como

$$u_t = \tilde{y}_t - \phi_1 \tilde{y}_{t-1} - \dots - \phi_p \tilde{y}_{t-p} + \theta_1 u_{t-1} + \dots + \theta_q u_{t-q}.$$

Tem-se que, sob a suposição de normalidade dos u_t , a função de densidade conjunta de u_1, u_2, \dots, u_n é

$$L(\phi, \theta, \sigma^2) = f(u_1, \dots, u_n) = (2\pi)^{-\frac{n}{2}} (\sigma)^{-n} \exp\left(-\sum_{t=1}^n \frac{u_t^2}{2\sigma^2}\right). \quad (2.12)$$

Os estimadores são obtidos maximizando a função verossimilhança acima, o que é feito através de procedimentos numéricos.

A verificação da adequação dos modelos estimados pode ser feita por análise de resíduos. A suposição inicial de $u_t \sim N(0, \sigma^2)$ e que u_t , $t=1, 2, \dots, n$ são independentes deve ser satisfeita.

Também pode-se utilizar o Critério de Informação de Akaike (AIC) para verificar a adequação do ajuste do modelo. O AIC é dado por

$$AIC = -2\ln L + 2m$$

onde m é o número de parâmetros do modelo e L é a função de verossimilhança do modelo.

2.3.1 Previsão nos modelos ARMA

Após identificação do modelo e estimação dos parâmetros é possível utilizar o modelo para fazer previsões para valores futuros da série temporal.

Seja $\hat{y}_t(l)$ o valor previsto na origem t para um horizonte de l períodos de tempo futuros.

O modelo ARMA dado em 2.9 pode ser escrito no tempo $t + l$ como

$$y_{t+l} = \phi_1 y_{t+l-1} + \dots + \phi_p y_{t+l-p} + u_{t+l} - \theta_1 u_{t+l-1} - \dots - \theta_q u_{t+l-q}, \quad (2.13)$$

ou utilizando o operador $\psi(B)$,

$$y_{t+l} = \sum_{j=-\infty}^{t+l} \psi_{t+l-j} u_j = \sum_{j=0}^{\infty} \psi_j u_{t+l-j}, \quad (2.14)$$

onde $\psi_0=1$

A melhor previsão para o valor futuro y_{t+l} , no sentido de apresentar erro quadrático médio (EQM) mínimo, é calculada como a esperança condicional de y_{t+l} dada as observações passadas da série, ou seja

$$\hat{y}_t(l) = E[y_{t+l}|y_t, y_{t-1}, \dots] = E_t[y_{t+l}]. \quad (2.15)$$

Também é possível reescrever a previsão $\hat{y}_t(l)$ como

$$\hat{y}_t(l) = \phi_1 E_t[y_{t+l-1}] + \dots + \phi_p E_t[y_{t+l-p}] - \theta_1 E_t[u_{t+l-1}] - \dots - \theta_q E_t[u_{t+l-q}] + E_t[u_{t+l}], \quad (2.16)$$

Para o cálculo de ?? deve-se considerar

- $E_t[y_{t-j}] = y_{t-j}$
- $E_t[y_{t+j}] = \hat{y}_t(j)$
- $E_t[u_{t-j}] = u_{t-j}$
- $E_t[u_{t+j}] = 0,$

para $j > 0$

2.4 Modelos Lineares Generalizados

Os Modelos lineares generalizados (MLG) propostos por Nelder & Wedderburn(1972) são uma extensão dos modelos lineares clássicos para dados que não possuem distribuição Gaussiana.

O modelo linear generalizado é um modelo em que a variável resposta Y segue uma distribuição da família exponencial, dada pela forma

$$f(y; \vartheta, \varphi) = \exp\{\varphi[y\vartheta - b(\vartheta)] + c(y, \varphi)\}, \quad (2.17)$$

onde φ e ϑ são parâmetros e b e c são funções específicas.

Seja η o preditor linear dado por

$$\eta = \sum_{j=1}^d x_j \beta_j, \quad (2.18)$$

onde, x_j , $j=1,2,\dots,d$, são covariáveis e β_j parâmetros a serem estimados.

A função de ligação g , que deve ser monótona e diferenciável, relaciona o preditor linear η com a esperança, μ , da variável resposta Y ,

$$g(\mu) = \eta. \quad (2.19)$$

Quando a variável resposta é uma contagem, é possível utilizar a distribuição de Poisson. Neste trabalho será utilizada essa distribuição para a modelagem de séries temporais de dados de contagem, já que é uma distribuição que compreende inteiros não negativos.

Se Y tem distribuição Poisson com média $\mu > 0$, então

$$f_Y(y|\mu) = \frac{e^{-\mu} \mu^y}{y!}, \quad (2.20)$$

que na forma da família exponencial pode ser escrito como

$$f_Y(y|\mu) = \exp(y \log \mu - \mu + (-\log y!)). \quad (2.21)$$

A função de ligação será a função log, então tem-se

$$\log \mu = \eta = \sum_{j=1}^d x_j \beta_j = \underline{x}' \beta, \quad (2.22)$$

onde $\underline{x}' = (1, x_1, \dots, x_d)$ e $\beta = (\beta_0, \beta_1, \dots, \beta_d)$.

A estimação dos $d + 1$ parâmetros $(\beta_0, \beta_1, \dots, \beta_d)$ é feita pelo método de máxima verossimilhança (EMV).

No MLG Poisson a função de verossimilhança é dada por

$$L(\beta) = \prod_{i=1}^n \frac{\exp(y_i \mathbf{x}_i' \beta)}{y_i!} \exp(-\exp \mathbf{x}_i' \beta), \quad \mathbf{x}_i = (1, x_{i1}, \dots, x_{id}), \quad (2.23)$$

pois, de acordo com 2.22, tem-se $\mu = \exp(\mathbf{x}'\beta)$.

Como o estimador de máxima verossimilhança não pode ser calculado de forma analítica, é necessário a utilização de algoritmos numéricos para a estimação do vetor de parâmetros $\beta = (\beta_0, \beta_1, \dots, \beta_d)$. O algoritmo de Newton Raphson é um dos mais utilizados e está descrito, resumidamente, a seguir.

O algoritmo utiliza-se da expansão da função escore $U_\beta = \frac{\partial L(\beta)}{\partial \beta}$ em torno de um valor inicial $\beta^{(0)}$, tal que

$$U_\beta \approx U_\beta^{(0)} + U_\beta^{\prime(0)}(\beta - \beta^{(0)}). \quad (2.24)$$

Assim obtém-se o processo iterativo

$$\beta^{(m+1)} = \beta^{(m)} + (-U_\beta^{\prime(m)})^{-1} U_\beta^{(m)}. \quad (2.25)$$

O processo é repetido até que se obtenha a convergência.

Adequação do modelo

Existem algumas medidas para verificar a qualidade dos ajustes dos modelos. Uma destas medidas é denominada Deviance e equivale à diferença de log-verossimilhanças

maximizadas, que serve para medir a distância dos valores ajustados pelo modelo aos dados.

A estatística Deviance é dada por

$$D^*(Y, \hat{\mu}) = 2[L(Y; Y) - L(\hat{\beta}; Y)] \quad (2.26)$$

sendo $L(Y; Y)$ a função de verossimilhança para o modelo saturado, isto é, quando o número de parâmetros é igual a n e $\hat{\beta}$ é a estimativa do parâmetro.

Encontrando um valor pequeno para a função deviance é um indicativo que para menor quantidade de parâmetros, obtém-se um ajuste tão bom quanto para o modelo saturado.

Capítulo 3

Modelos para séries de contagem

Este capítulo tem por objetivo apresentar dois modelos que são extensões do MLG fundamentados nos artigos de Benjamin et al (2003) e Davis et al (2003).

Das metodologias citadas nos capítulos anteriores, os modelos ARMA modelam dados autocorrelacionados, como é o caso das séries temporais, todavia eles restringem-se à suposição de Normalidade. Os modelos lineares generalizados modelam distribuições da família exponencial, mas não consideram as correlações que ocorrem entre as observações em função do tempo. Por isso, é necessário a utilização de outros modelos para a análise de séries temporais de dados de contagem.

Neste trabalho serão utilizados os modelos Autoregressivo Média Móvel Generalizado (GARMA) e o Autoregressivo Média Móvel Linear Generalizado (GLARMA) . Ambos são úteis para modelar variáveis respostas dependentes no tempo e que seguem uma distribuição da família exponencial. Neste trabalho, será dada atenção especial à distribuição de probabilidade Poisson, já que esta é uma das mais utilizadas para a modelagem de séries temporais de dados de contagem.

3.1 Modelo Autoregressivo Média Móvel Generalizado (GARMA)

O GARMA proposto por Benjamin et al (2003) é uma extensão dos modelos MLG para variáveis resposta que, condicionais à informação passada, possuem distribuições pertencentes à família exponencial.

Seja y_t uma série temporal e \mathbf{x}_t um vetor de covariáveis, para $t=1,2,\dots,n$. No modelo GARMA a distribuição condicional de cada observação y_t dado as informações anteriores $H_{t-1} = (x_1, \dots, x_t, y_1, \dots, y_{t-1}, \mu_1, \dots, \mu_{t-1})$ tem distribuição na família exponencial dada em 2.17.

Como no MLG, μ_t está relacionada a η_t , um preditor linear, através de uma função de ligação g

$$g(\mu_t) = \eta_t = \mathbf{x}'_t \beta + \tau_t, \quad (3.1)$$

onde

$$\tau_t = \sum_{j=1}^p \phi_j \{g(y_{t-j}) - \mathbf{x}'_{t-j} \beta\} + \sum_{j=1}^q \theta_j \{g(y_{t-j}) - \eta_{t-j}\}. \quad (3.2)$$

A seguir é descrito um exemplo de modelo GARMA, em que y_t segue uma distribuição Poisson e pode ser utilizada para modelar dados de contagem.

GARMA Poisson

O modelo GARMA Poisson é o modelo GARMA onde a função distribuição de probabilidade de y_t é Poisson.

É possível definir o modelo como

$$f(y_t | H_{t-1}) = \frac{\mu_t^{y_t} e^{-\mu_t}}{y_t!}. \quad (3.3)$$

A função de ligação utilizada é a função logarítmica, pois ela somente admite valores naturais para η_t .

Assim

$$\eta_t = \log \mu_t = \mathbf{x}'_t \beta + \sum_{j=1}^p \phi_j \{\log (y_{t-j}^*) - \mathbf{x}'_{t-j} \beta\} + \sum_{j=1}^q \theta_j \log \frac{y_{t-j}^*}{\mu_{t-j}}, \quad (3.4)$$

onde $y_{t-j}^* = \max(y_{t-j}, c)$ e $0 < c < 1$, uma restrição necessária quando $y_{t-j} = 0$ para que seja possível utilizar a função logarítmica.

3.1.1 Estimação nos modelos GARMA

De acordo com Benjamin et al (2003) o vetor de parâmetros $\gamma' = (\beta', \theta', \phi')$ é estimado pelo método de máxima verossimilhança.

A função log verossimilhança no GARMA Poisson é

$$\log L(\gamma) = \sum_{t=1}^n (y_t \log \mu_t - \mu_t - \log y_t!), \quad (3.5)$$

onde

$$\mu_t = \exp(\eta_t) = \exp(\mathbf{x}'_t \beta + \sum_{j=1}^p \phi_j \log (y_{t-j}^*) - \mathbf{x}'_{t-j} \beta) + \sum_{j=1}^q \theta_j \log \frac{y_{t-j}^*}{\mu_{t-j}}. \quad (3.6)$$

Para maximizar a função de log-verossimilhança são utilizados métodos numéricos.

3.1.2 Propriedades dos modelos GARMA

Benjamin et al (2003) investigaram a estacionariedade do modelo fazendo a análise para algumas funções de ligação específicas.

Os autores provam que, para o caso da função de ligação identidade, a média marginal de y_t do modelo GARMA definido em 3.1 e 3.2 é dada por

$$E(y_t) = x'_t \beta, \quad (3.7)$$

e no caso da distribuição Poisson a variância marginal é dada por

$$Var(y_t) = \psi^{(2)}(1)\beta_0, \quad (3.8)$$

onde $\psi^{(2)} = 1 + \psi_1^2 B + \psi_2^2 B^2 + \dots$ e $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots = \phi(B)^{-1} \theta(B)$.

De acordo com Benjamim et al.(2003) modelos com função de ligação diferente da identidade não permitem uma forma direta e tratável para o cálculo da média e variância para verificação da estacionariedade. Desta forma, estes autores fizeram simulações com 16 combinações de ϕ e $\theta = (-0.4, 0, 0.4, 0.8)$ para verificar estacionariedade em modelos com função de ligação diferente da identidade e não encontraram nenhuma evidência de não estacionariedade para as combinações de parâmetros testadas.

3.1.3 Previsão no modelo GARMA

Após a verificação do modelo é possível fazer previsões para o modelo GARMA ajustado, assim como no modelo ARMA.

A previsão no período de origem t para um horizonte de l períodos de tempo futuros, $\hat{y}_t(l)$ é dada por

$$\hat{y}_t(l) = E_t[y_{t+l}] = \mu_{t+l} = \exp(x'_{t+l} \beta + \sum_{j=1}^p \phi_j [g(y_{t+l-j}) - x'_{t+l-j}] + \sum_{j=1}^q \theta_j [g(y_{t+l-j}) - \eta_{t+l-j}]). \quad (3.9)$$

3.2 Modelo Autoregressivo Média Móvel Linear Generalizado (GLARMA)

A classe de modelos GLARMA foi proposta inicialmente por Shephard (1995) e generalizada por Davis et al (2003). Estes procedimentos propõem uma alternativa para modelar séries que possuem uma estrutura de dependência. Da mesma forma que o GARMA, o modelo é uma extensão da estrutura usual de MLG, que permite a correlação serial entre as observações.

Davis et al (2003) apresentam o modelo para os dados condicionais ao passado tendo uma distribuição de probabilidades Poisson.

Sejam y_t as observações e $F_{t-1} = (y^{(t-1)}, x^{(t)})$, onde $y^{(t-1)}$ é o passado do processo de contagem e $x^{(t)}$ é o passado e o presente das variáveis regressoras.

A distribuição condicional de y_t é dada por

$$y_t | F_{t-1} \sim P(\mu_t), \quad (3.10)$$

onde $P(\mu_t)$ é uma Poisson com média μ_t . Como os autores trabalham com a distribuição de Poisson, a função de ligação utilizada é a logarítmica.

No caso do modelo GLARMA, o preditor é dado por

$$\eta_t = \log(\mu_t) = \mathbf{x}'_t \beta + \sum_{i=1}^{\infty} \tau_i \varepsilon_{t-i} \quad (3.11)$$

onde τ é um vetor de parâmetros e

$$\varepsilon_t = \frac{(y_t - \mu_t)}{\mu_t^\lambda}, \quad (3.12)$$

com $\lambda \in (0,1]$.

A estrutura média móvel infinita pode ser especificada em termos de um número finito

de parâmetros. Uma forma de parametrizar os pesos τ é expressá-los como coeficientes de um filtro autoregressivo médias móveis (Box & Jenkins, 1972),

$$\tau(B) = \sum_{i=1}^{\infty} \tau_i B^i = \frac{\theta(B)}{\phi(B)} - 1, \quad (3.13)$$

sendo $\phi(B)$ e $\theta(B)$ os componentes autoregressivos e médias móveis do modelo de Box e Jenkins, dados em (2.3).

Dessa forma o modelo é

$$\eta_t = \log(\mu_t) = \mathbf{x}'_t \beta + Z_t, \quad (3.14)$$

onde

$$Z_t = \phi_1(Z_{t-1} + \varepsilon_{t-1}) + \dots + \phi_p(Z_{t-p} + \varepsilon_{t-p}) + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}. \quad (3.15)$$

3.2.1 Estimação nos modelos GLARMA

Para estimar os parâmetros $\gamma' = (\beta', \theta', \phi')$ é utilizado o método de máxima verossimilhança. A log verossimilhança para o modelo GLARMA Poisson é a mesma do modelo GARMA, apresentada em 3.5, porém com μ_t dado por

$$\mu_t = \exp(\eta_t) = \exp(\mathbf{x}'_t \beta + \phi_1(Z_{t-1} + \varepsilon_{t-1}) + \dots + \phi_p(Z_{t-p} + \varepsilon_{t-p}) + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}). \quad (3.16)$$

O estimador de máxima verossimilhança não possui uma forma analítica, então são utilizados métodos numéricos para maximização.

3.2.2 Propriedades dos modelos GLARMA

Para provar propriedades do modelo, Davis et al(2003) reduziram o modelo GLARMA a um modelo sem covariável e com $q=1$. Substituindo 3.12 em 3.11 tem-se

$$\eta_t = \beta + \tau_1(y_{t-1} - \exp(\eta_{t-1}))(\exp(-\lambda\eta_{t-1})). \quad (3.17)$$

De 3.17 foi possível encontrar o valor da esperança para η_t

$$E(\eta_t) = E[E[(\eta_t|\eta_{t-1})]] = \beta.$$

Para a variância os autores obtiveram o resultado

$$Var(\eta_t) = Var[E(\eta_t|\eta_{t-1})] + E[Var(\eta_t|\eta_{t-1})] = \theta^2 E[\exp(1 - 2\lambda)\eta_{t-1}].$$

É possível perceber que a variância depende do valor de λ , então Davis et al(2003) concluíram que

- Se $\frac{1}{2} \leq \lambda < 1$ existe uma distribuição estacionária para y_t ,
- $\lambda < \frac{1}{2}$ a existência de estacionariedade não está estabelecida.

3.2.3 Previsão no modelo GLARMA

Assim como nos modelos ARMA e GARMA, após a identificação, estimação dos parâmetros e verificação do modelo é possível fazer previsões.

Adaptando a metodologia de previsões de modelos ARMA para o modelo GLARMA, têm-se que o valor previsto de origem t e horizonte l , $\hat{y}_t(l)$ é dado por

$$\hat{y}_t(l) = E_t[y_{t+l}] = \mu_{t+l} = \exp(x'_{t+l}\beta + E_t[Z_{t+l}]), \quad (3.18)$$

onde

$$E_t[Z_{t+l}] = \phi_1(E_t[Z_{t+l-1}] + E_t[\varepsilon_{t+l-1}]) + \dots + \phi_p(E_t[Z_{t+l-p}] + E_t[\varepsilon_{t+l-p}]) + \theta_1 E_t[\varepsilon_{t+l-1}] + \dots + \theta_q E_t[\varepsilon_{t+l-q}],$$

(3.19)

sendo $E_t[\varepsilon_{t-j}]$ as esperanças condicionais dos ε_t .

Assim como nos modelos ARMA deve-se considerar

- $E_t[\varepsilon_{t-j}] = \varepsilon_{t-j}$
- $E_t[\varepsilon_{t+j}] = 0$.
- $E_t[Z_{t-j}] = Z_{t-j}$
- $E_t[Z_{t+j}] = \hat{Z}_t(j)$,

para $j > 0$.

Capítulo 4

Estudo de simulação

Nesta seção será apresentado um estudo Monte Carlo, no qual foram implementados os modelos GARMA Poisson(p,q) e GLARMA Poisson(p,q). Foram gerados modelos com p e q iguais a (1,0), (0,1), (2,0) e (0,2) afim de comparar as estimativas geradas pelo método de máxima verossimilhança nos modelos GARMA, GLARMA e MLG.

Todos os programas foram desenvolvidos na linguagem de programação R versão (3.1.0). Para a obtenção da estimativa de máxima verossimilhança foi utilizada a função `nlm`, que minimiza funções arbitrárias escritas em R.

Os valores dos parâmetros adotados para simular séries dos modelos GARMA e GLARMA foram fixados em $\beta_0 = 1$, $\beta_1 = 0,3$, três valores para θ e ϕ (-0,5, 0,25 e 0,5) nos modelos MA(1) e AR(1) e as combinações (-0,5 e 0,25) e (0,5 e 0,25) nos casos MA(2) e AR(2). O parâmetro λ não foi estimado e foi fixado para a geração da série GLARMA Poisson em 0,5 e para estimação foram utilizados 0,5 e 1,0.

As séries foram simuladas com e sem covariável de tendência, dada por $\frac{i}{n}$, $i=1,\dots,n$.

O número de repetições de Monte Carlo utilizado foi 1.000 e o tamanho da série foi fixado em 100. Para cada série simulada foram calculadas a média e o erro quadrático médio (EQM) das 1.000 estimativas dos parâmetros.

De maneira geral, com o estudo de simulação foi possível chegar a algumas conclusões à cerca dos modelos. Na maior parte das simulações o próprio modelo, dos quais os

dados foram gerados, apresenta melhores estimativas, como esperado.

Uma característica importante destes modelos é que o parâmetro β_1 , relativo à covariável, é sempre bem estimado, independente do modelo gerador dos dados. O coeficiente β_0 , entretanto, só é bem estimado pelo modelo que gera os dados.

Outra conclusão interessante refere-se à estimação dos parâmetros ϕ e θ em modelos AR(1) e MA(1). Quando o modelo gerador é o GARMA com ϕ ou θ positivos, o GLARMA com $\lambda = 1,0$ fornece boas estimativas para estes parâmetros. Quando o modelo gerador é o GLARMA com $\lambda = 0,5$, e ϕ ou θ positivos, o modelo GARMA estima satisfatoriamente estes parâmetros. Porém, em outras situações, não foi possível identificar um padrão único na estimação dos modelos.

Finalmente, parece que a variabilidade das estimativas aumenta à medida que o valor dos parâmetros θ ou ϕ aumentam.

Estas são conclusões gerais que foram observadas no estudo de simulação. A seguir são apresentados, detalhadamente, os resultados para cada modelo. Na Seção 4.1 estão os resultados para séries geradas do modelo GARMA e na Seção 4.2 do modelo GLARMA.

4.1 Séries geradas pelo modelo GARMA

4.1.1 GARMA(1,0)

Esta seção apresenta os resultados para séries geradas de um modelo GARMA com um componente autoregressivo de ordem 1 na média do processo. Estas séries foram então ajustadas utilizando-se os modelos MLG, GARMA, e GLARMA com $\lambda = 0,5$ e $1,0$. Os resultados para o modelo sem covariável são apresentados na Figura 4.1, enquanto a Figura 4.2 mostra os resultados para o modelo com covariável tendência. Tabelas com estas mesmas comparações são apresentadas no Anexo 7.2 para se ter um melhor detalhamento das estimativas médias e seus EQM's.

A primeira coluna de gráficos em cada figura corresponde às estimativas dos parâmetros obtidas quando ϕ foi fixado em $-0,5$, a segunda coluna com $\phi = 0,25$ e a terceira

$\phi = 0,5$. O valor de β_0 foi mantido sempre igual a 1,0 e quando existe covariável β_1 foi fixado em 0,3. Deve-se observar que com o método MLG não é possível estimar o parâmetro ϕ .

A primeira conclusão que pode ser tirada das Figuras 4.1 e 4.2 é que, como esperado, as melhores estimativas são obtidas pelo modelo GARMA.

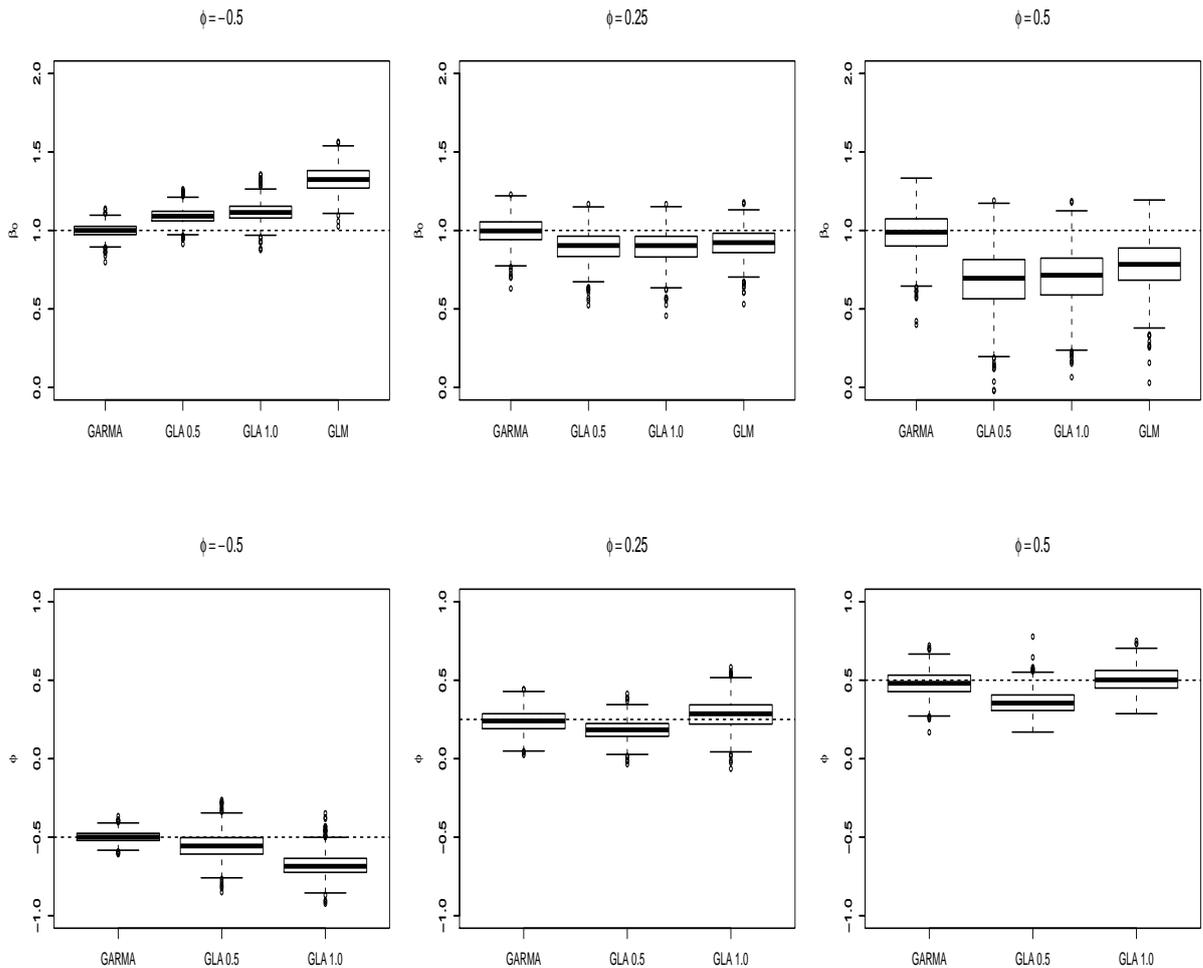


Figura 4.1: Boxplot das estimativas para séries geradas pelo GARMA AR(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro ϕ .

Uma análise detalhada da Figura 4.1 mostra que, à respeito das estimativas para β_0 , o MLG não fornece boas estimativas e os modelos GLARMA com $\lambda = 0,5$ e $1,0$ resultam em estimativas semelhantes. Além disso, é possível observar que quando o parâmetro

ϕ é negativo as estimativas obtidas para β_0 apresentam menor dispersão em todos os modelos. Com relação ao parâmetro ϕ , as estimativas obtidas apresentaram pequena variabilidade. A estimativa obtida pelo GLARMA com $\lambda = 1,0$ para ϕ quando seu real valor é $-0,5$ apresenta um vício maior do que as demais estimativas geradas pelos outros modelos. É possível observar o mesmo comportamento para a estimativa fornecida por GLARMA com $\lambda = 0,5$ quando o real valor de $\phi = 0,5$.

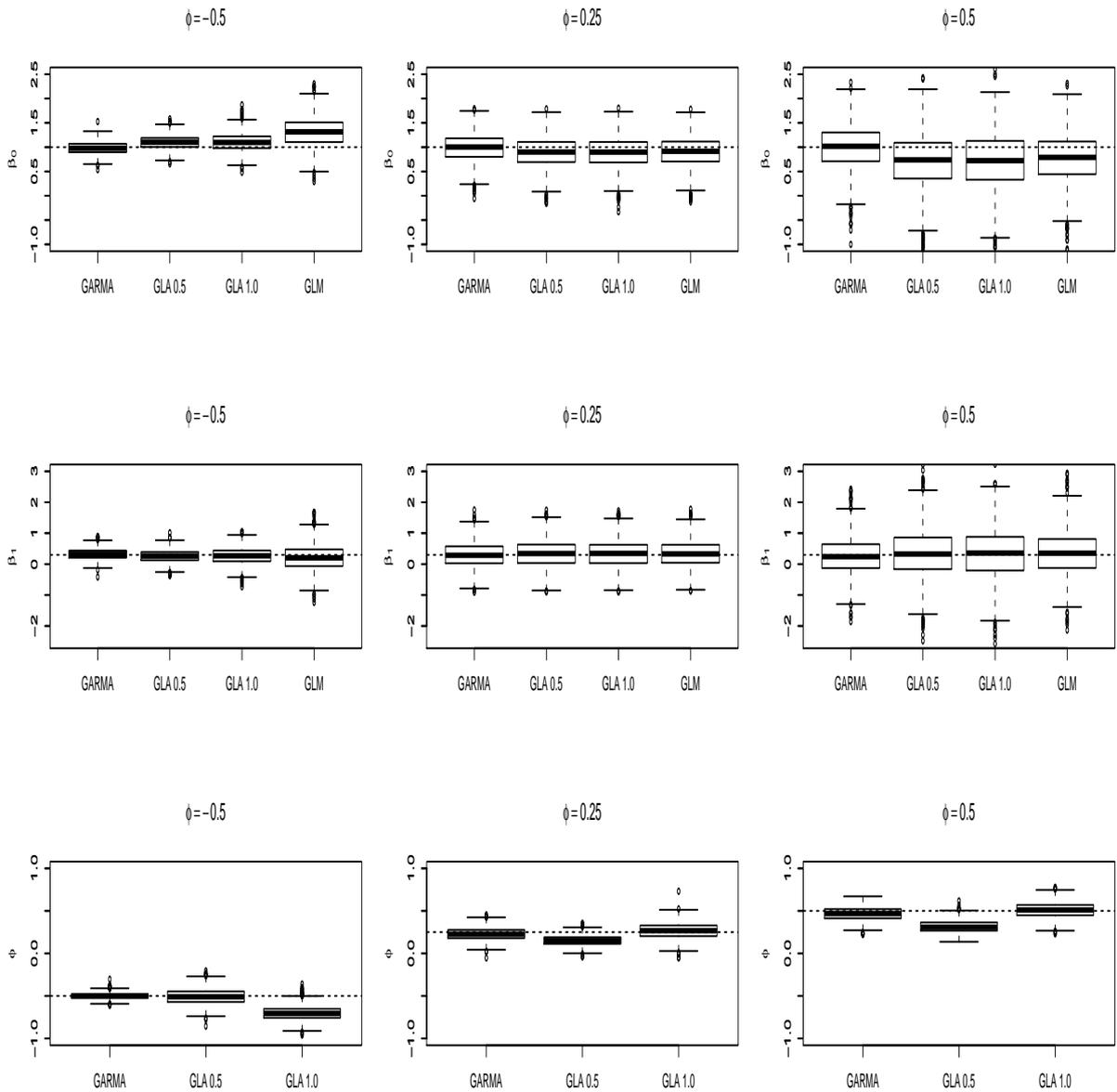


Figura 4.2: Boxplot das estimativas para séries geradas pelo GARMA AR(1) com covariável tendência. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 e os da terceira linha para ϕ .

A Figura 4.2 mostra os resultados para o modelo GARMA AR(1) com covariável.

Observando a Tabela 7.2 no Anexo 7.2 é possível perceber que em quase todas as estimativas do modelo MLG ele apresenta um erro quadrática médio maior do que as estimativas dos demais modelos.

As estimativas para β_0 e β_1 apresentam maior variabilidade à medida que o valor real de ϕ aumenta. As estimativas obtidas para β_1 são boas independente do modelo ajustado, mesmo pelo MLG.

As estimativas obtidas por GLARMA com $\lambda = 1,0$, para ϕ positivo, são melhores do que quando $\lambda = 0,5$. Todavia quando ϕ é negativo GLARMA com $\lambda = 0,5$ fornece melhores estimativas. Não é possível observar o mesmo padrão para as estimativas de β_0 e β_1 , onde as estimativas geradas pelos dois modelos comportam-se de maneira semelhante.

4.1.2 GARMA(0,1)

Os resultados das Tabelas 7.4 e 7.3, no Anexo 7.2, e nas Figuras 4.3 e 4.5 referem-se às estimativas obtidas para séries geradas de uma GARMA com MA(1) sem e com covariável, respectivamente.

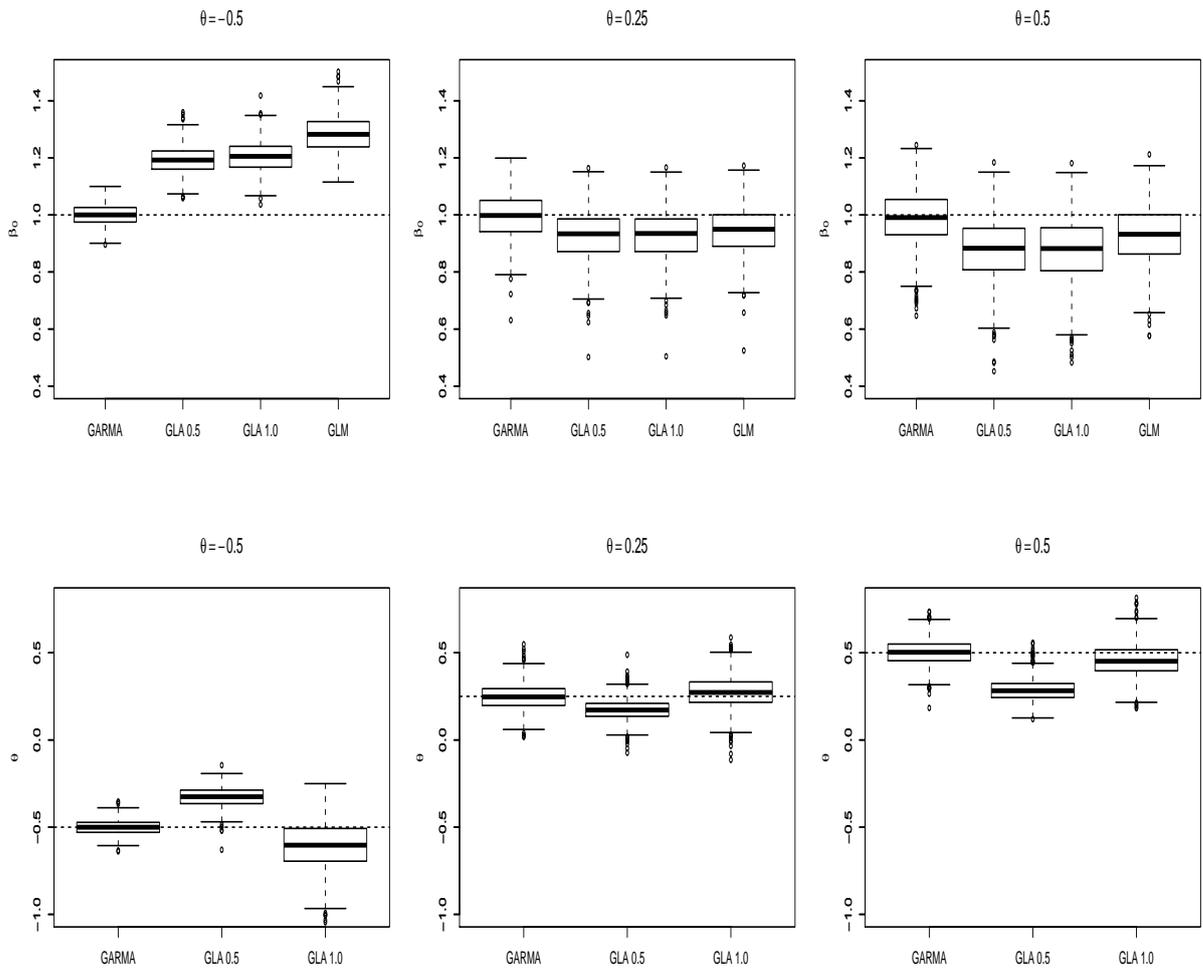


Figura 4.3: Boxplot para séries geradas pelo GARMA MA(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro θ .

O melhor modelo novamente para as estimativas dos parâmetros foi o GARMA, dos quais os dados foram gerados. Todas as estimativas geradas para β_0 , quando θ tem valor real positivo, são subestimadas pelos modelos GLARMA e MLG, e essas estimativas são superestimadas, quando θ é negativo. Já as estimativas para β_0 do modelo GLARMA com $\lambda = 0,5$ e $\lambda = 1,0$ apresentam valores semelhantes. Com relação ao parâmetro θ as estimativas obtidas pelos modelos GLARMA são melhores quando $\lambda = 1,0$. O MLG não resultou em estimativas satisfatórias para os parâmetros.

Figura 4.4: Boxplot para séries geradas pelo GARMA MA(1) com covariável

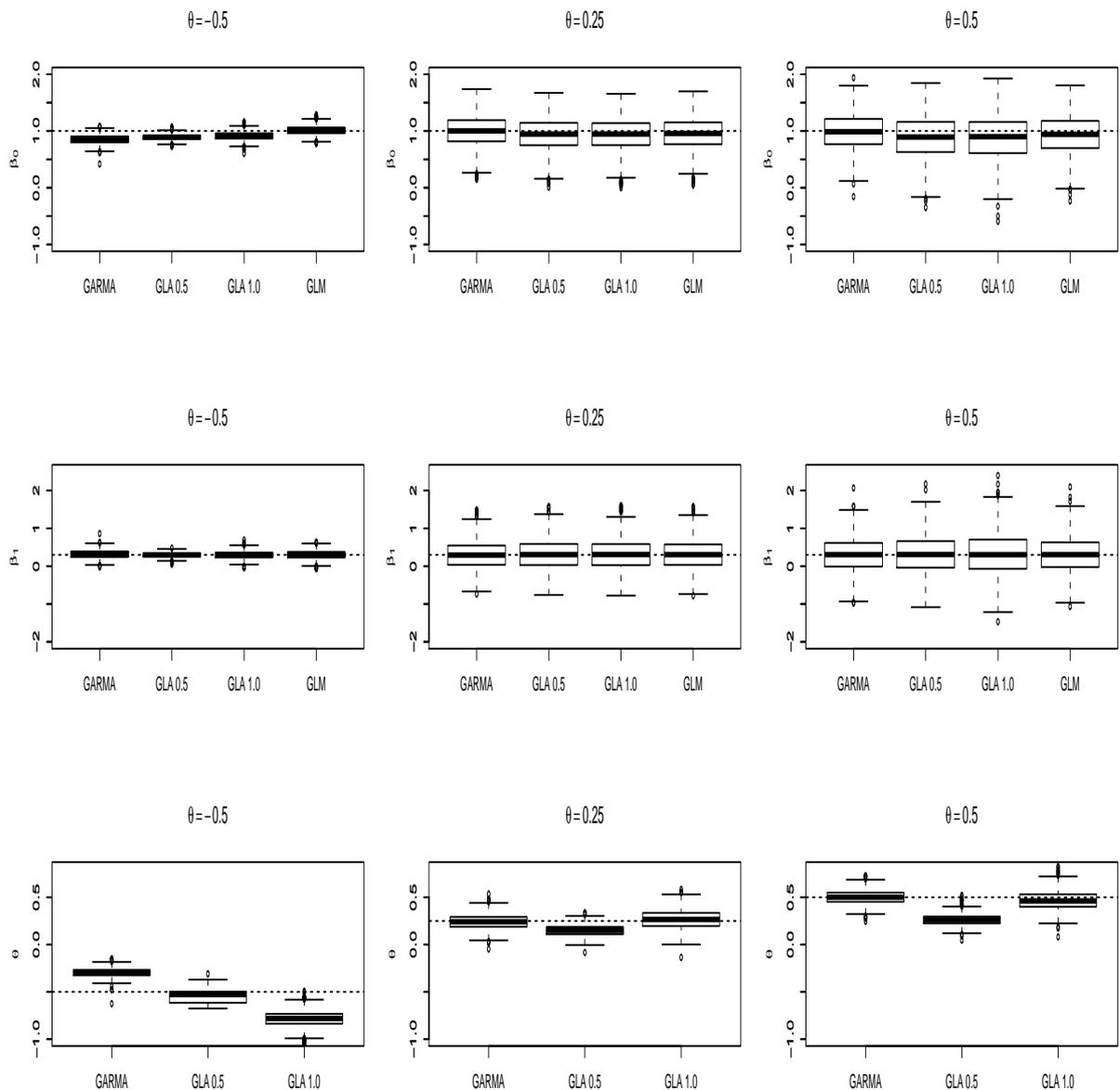


Figura 4.5: Boxplot para séries geradas pelo GARMA MA(1) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro β_1 e os da terceira linha para o parâmetro θ .

Uma análise da Figura 4.5 mostra que o modelo GARMA, dos quais os dados foram gerados, apresentou as melhores estimativas somente nos casos em que os valores reais de θ eram positivos.

É possível observar que a estimativa para β_1 é satisfatória para qualquer modelo. Já com relação a β_0 as estimativas são subestimadas quando θ é positivo. As estimativas para β_0 e β_1 foram semelhantes para o modelo GLARMA com os dois valores de λ . No entanto, quando a estimativa é para θ , o modelo GLARMA com $\lambda = 1,0$ apresenta melhores resultados quando o real valor do parâmetro é positivo. E GLARMA com $\lambda = 0,5$ apresenta melhores resultados quando θ é negativo.

A variabilidade das estimativas de β aumenta à medida que θ aumenta.

4.1.3 GARMA(2,0)

Nas Figuras 4.6 e 4.7 são apresentados os resultados encontrados para as estimativas de um modelo GARMA AR(2,0) sem e com covariável, respectivamente.

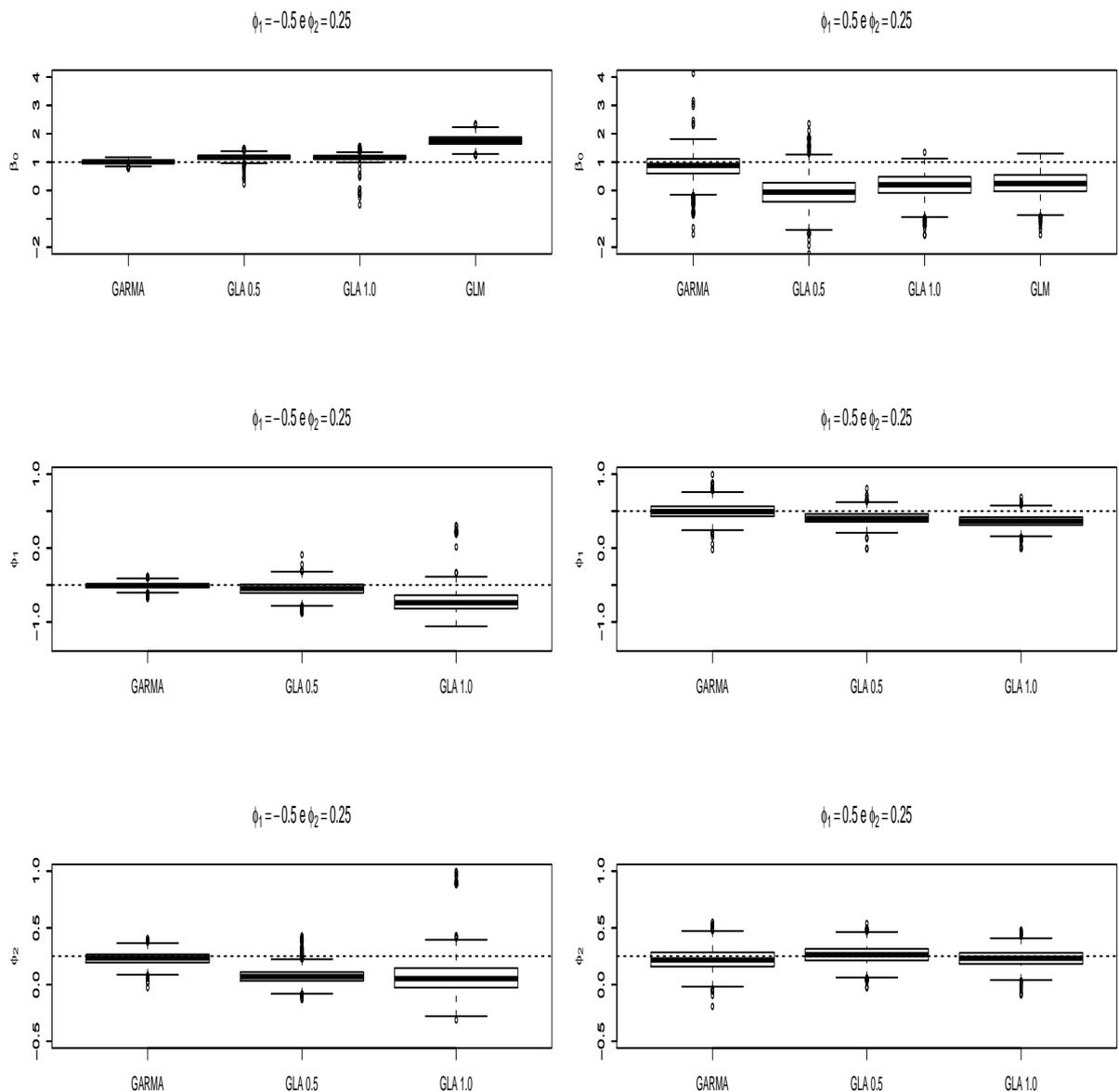


Figura 4.6: Boxplot para séries geradas pelo GARMA AR(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro ϕ_1 e os da terceira para ϕ_2 .

Nos modelos em que os valores reais de $\phi_1 = 0,5$ e $\phi_2 = 0,25$ nenhum dos modelos apresentaram boas estimativas para β_0 , o que pode ser melhor verificado na Tabela 7.5. Porém, as estimativas para ϕ_1 e ϕ_2 foram satisfatórias para todos os modelos.

As melhores estimativas obtidas quando $\phi_1 = -0,5$ e $\phi_2 = 0,25$ foram obtidas pelo

modelo GARMA. As estimativas geradas pelos modelos GLARMA com parâmetros $\lambda = 0,5$ e $1,0$ apresentam estimativas semelhantes para todos os parâmetros. Já o modelo MLG não apresenta estimativas satisfatórias.

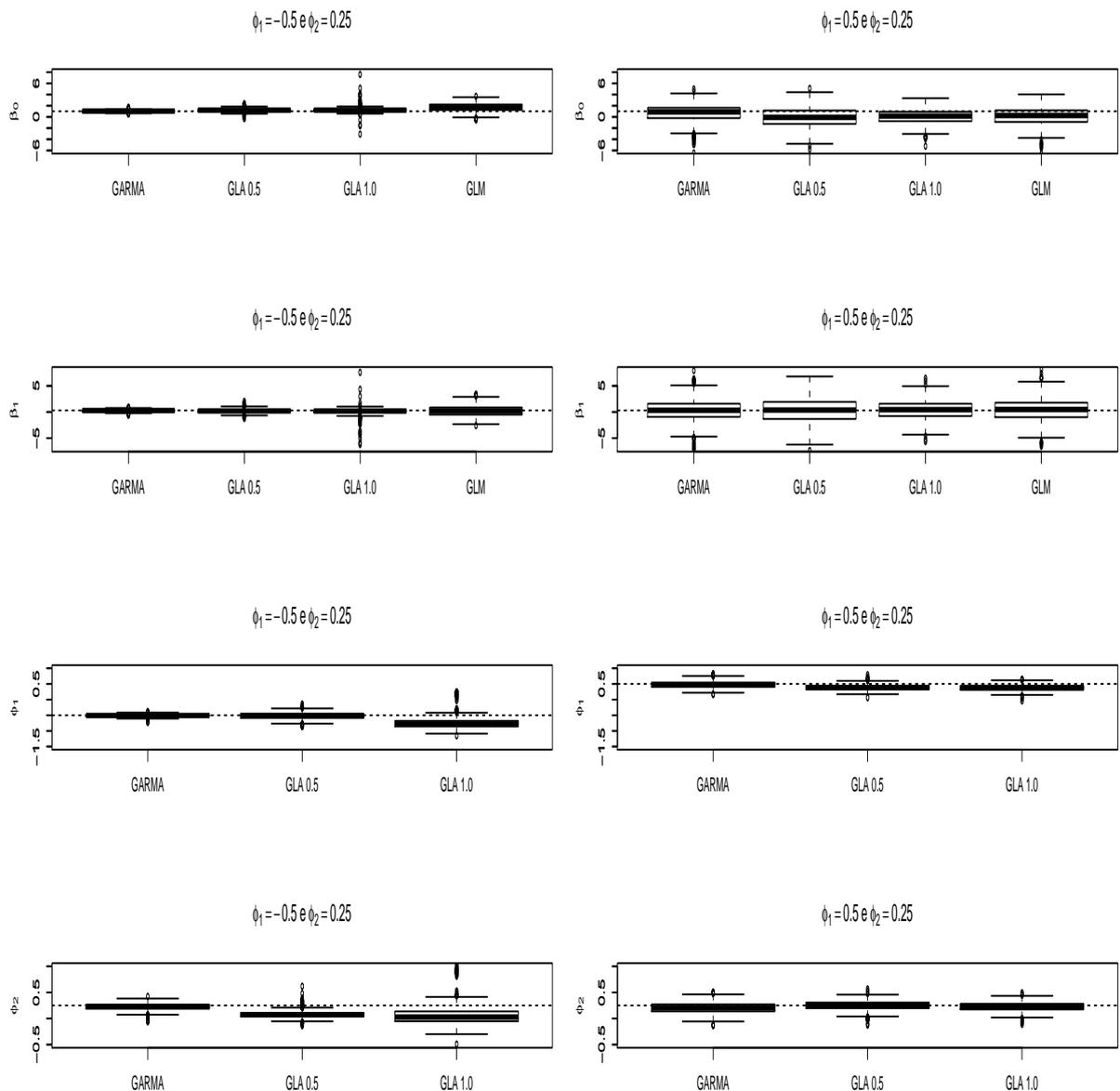


Figura 4.7: Boxplot para séries geradas pelo GARMA AR(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira para parâmetro ϕ_1 e os da quarta para o parâmetro ϕ_2 .

A Figura 4.7 é referente aos resultados do modelo com covariável. As estimativas obtidas por todos os modelos para β_0 apresentam um vício grande quando o modelo é

gerado com $\phi_1 = 0,5$ e $\phi_2 = 0,25$, como pode ser melhor verificado na Tabela 7.6.

Para o modelo gerado com $\phi_1 = -0,5$ e $\phi_2 = 0,25$ as melhores estimativas encontradas foram pelo modelo GARMA. De maneira geral, as estimativas para β_1 apresentaram um EQM bem superior ao EQM das outras estimativas.

As estimativas obtidas pelos modelos GLARMA para ϕ_2 e β_1 , no caso em que $\phi_1 = -0,5$ e $\phi_2 = 0,25$ não foram satisfatórias. Assim como o MLG não apresentou boas estimativas para β_0 .

4.1.4 GARMA(0,2)

As Tabelas 7.7 e 7.8, em Anexo 7.2 e as Figuras 4.8 e 4.9 apresentam as estimativas geradas pelos modelos quando as séries são geradas de um GARMA MA(2) sem e com covariável, respectivamente.

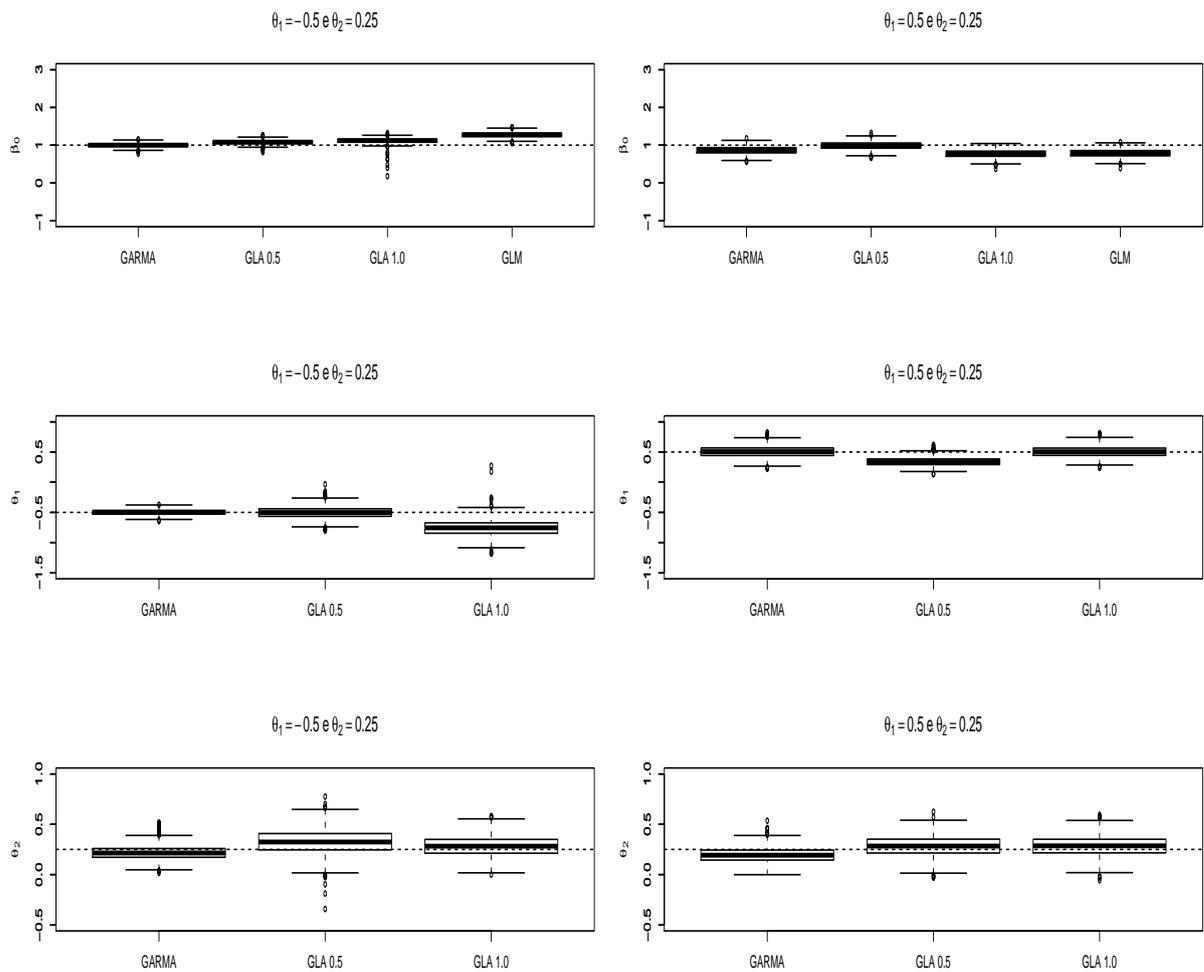


Figura 4.8: Boxplot para séries geradas pelo GARMA MA(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro θ_1 e os da terceira para o parâmetro θ_2

As melhores estimativas foram obtidas pelo modelo GARMA, cujos os dados foram gerados. As estimativas obtidas para β_0 no caso que $\theta_1 = 0,5$ e $\theta_2 = 0,25$ são todas subestimadas, também é possível observar que o MLG não apresenta boas estimativas para β_0 .

Novamente, observa-se que estimativas dos modelo GLARMA com $\lambda = 0,5$ são muito semelhantes as estimativas do mesmo modelo com $\lambda = 1,0$ para os parâmetros β_0 e θ_2 .

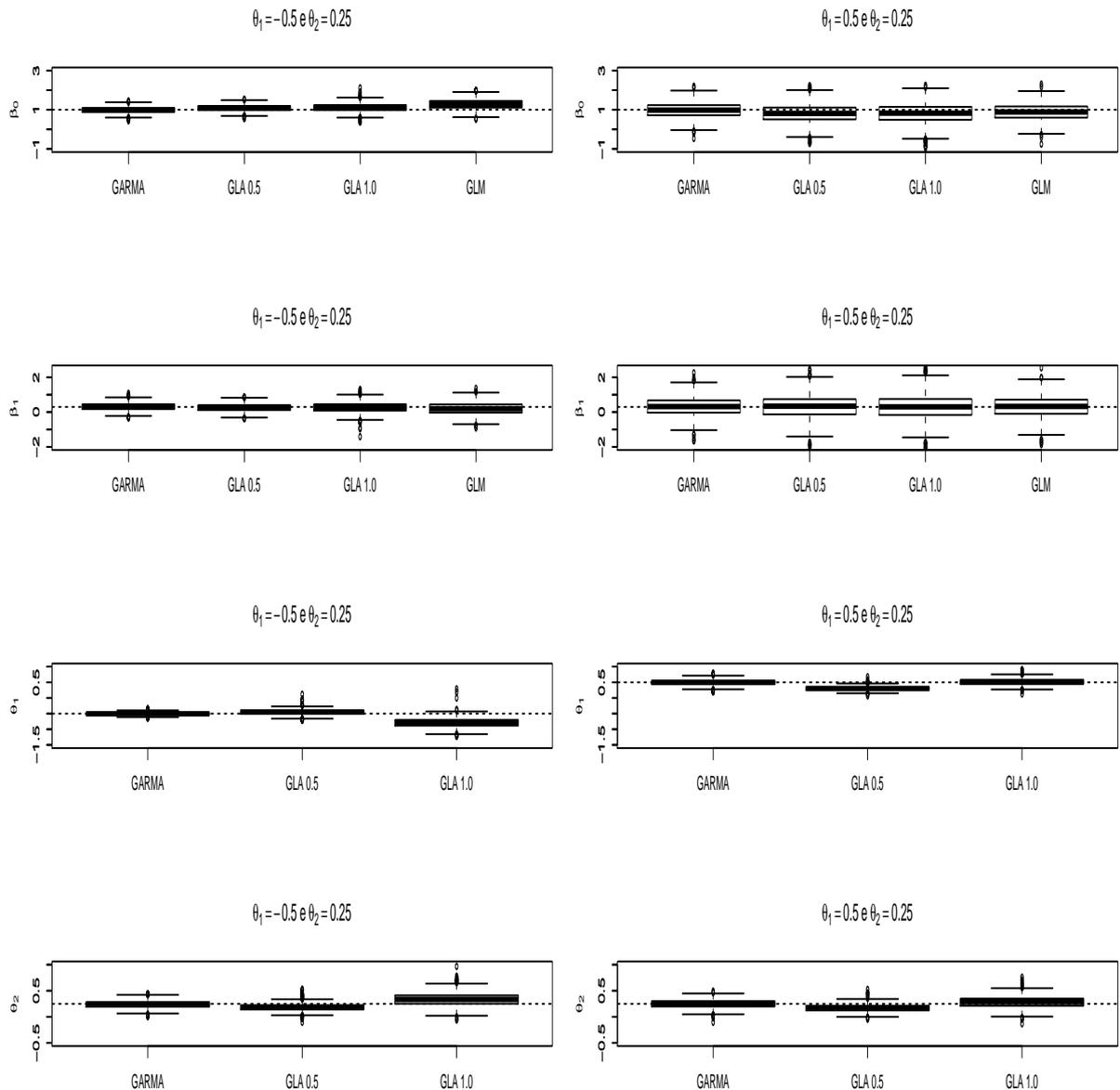


Figura 4.9: Boxplot para séries geradas pelo GARMA MA(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira para o parâmetro θ_1 e os da quarta linha para o parâmetro θ_2

A Figura 4.9 refere-se aos resultados do modelo com covariável. As melhores estimativas foram encontradas pelo modelo GARMA, com exceção para β_1 na simulação em que θ_1 e θ_2 são positivos. Com relação às estimativas obtidas para β_1 , todos os modelos resultaram em valores próximos aos reais, embora seja possível observar que quando θ_1 e

θ_2 são positivos, os valores estimados para o parâmetro são superestimados em todos os modelos ajustados. Em contrapartida, para esta mesma simulação encontra-se valores subestimados para β_0 .

Novamente, nota-se que as estimativas dos modelos GLARMA com parâmetro $\lambda = 0,5$ e $\lambda = 1,0$ apresentam estimativas semelhantes para β_0 e para β_1 . Também é possível observar que quando $\theta_1 = -0,5$ e $\theta_2 = 0,25$, encontra-se melhores estimativas para esses parâmetros pelo modelo GLARMA com $\lambda = 0,5$ e quando $\theta_1 = 0,5$ e $\theta_2 = 0,25$ as melhores estimativas são obtidas pelo modelo GLARMA com $\lambda = 1,0$.

4.2 Séries geradas pelo modelo GLARMA

4.2.1 GLARMA(1,0)

Nas Tabelas 7.10 e 7.9, Anexo 7.2, e nas Figuras 4.11 e 4.10 são apresentados os resultados das estimativas de séries geradas por um modelo GLARMA AR(1,0) sem e com covariável, respectivamente.

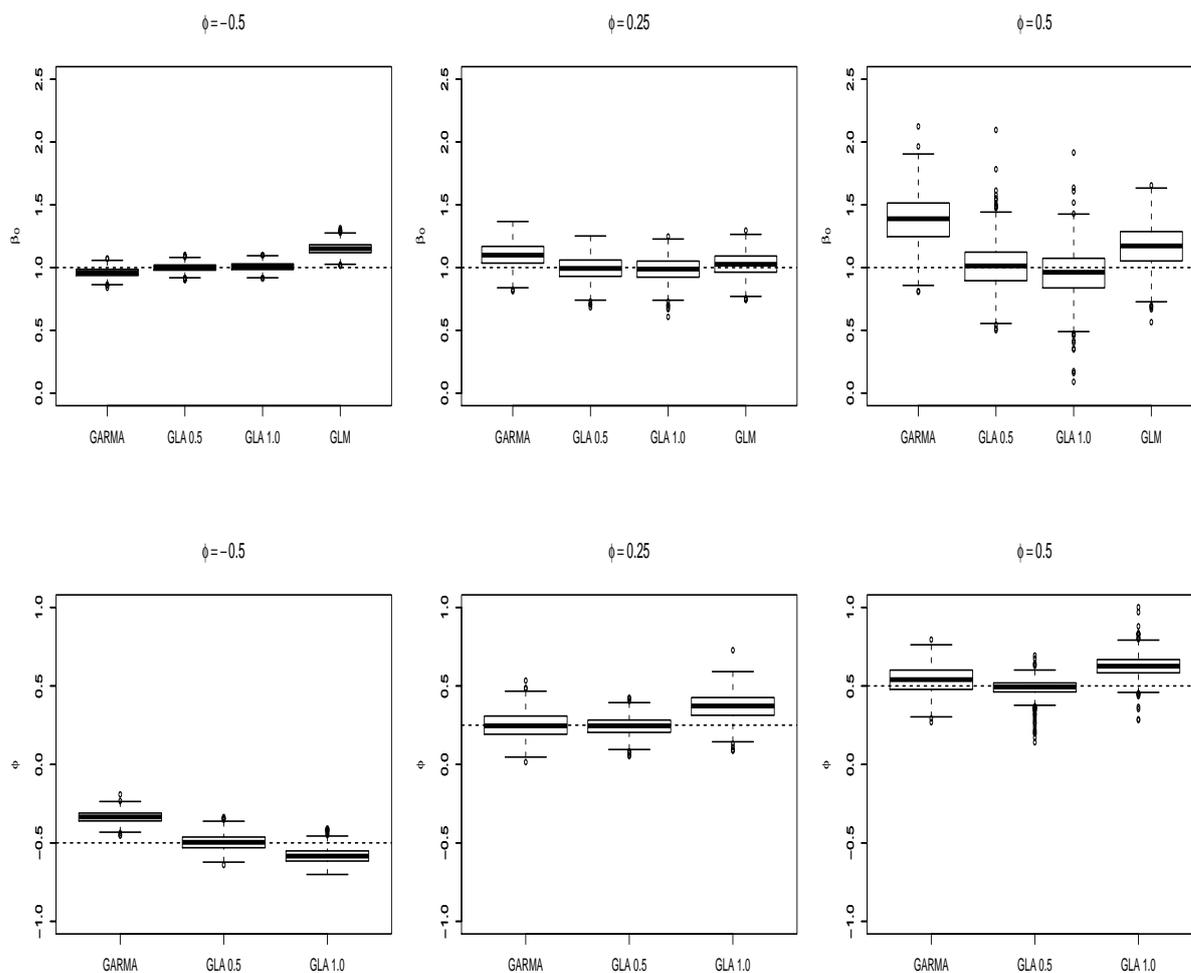


Figura 4.10: Boxplot para séries geradas pelo GLARMA AR(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e os da segunda linha para o parâmetro ϕ

As melhores estimativas são obtidas pelo modelo que os dados foram gerados, GLARMA com $\lambda = 0,5$, com exceção do parâmetro ϕ , quando seu real valor é $0,25$. É possível observar que à medida que ϕ aumenta as estimativas apresentam maior variabilidade.

O modelo MLG não fornece estimativas satisfatórias para β_0 quando $\phi = -0,5$ e $\phi = 0,5$. Assim como as estimativas encontradas para β_0 , pelo modelo GARMA, também apresentam um vício maior para $\phi = 0,25$ e $\phi = 0,5$. Com relação ao parâmetro ϕ , o modelo GARMA só não fornece estimativas ruins quando o real valor do parâmetro é $\phi = 0,25$. Já as estimativas obtidas para o modelo GLARMA com $\lambda = 1,0$ apresentam

vício grande.

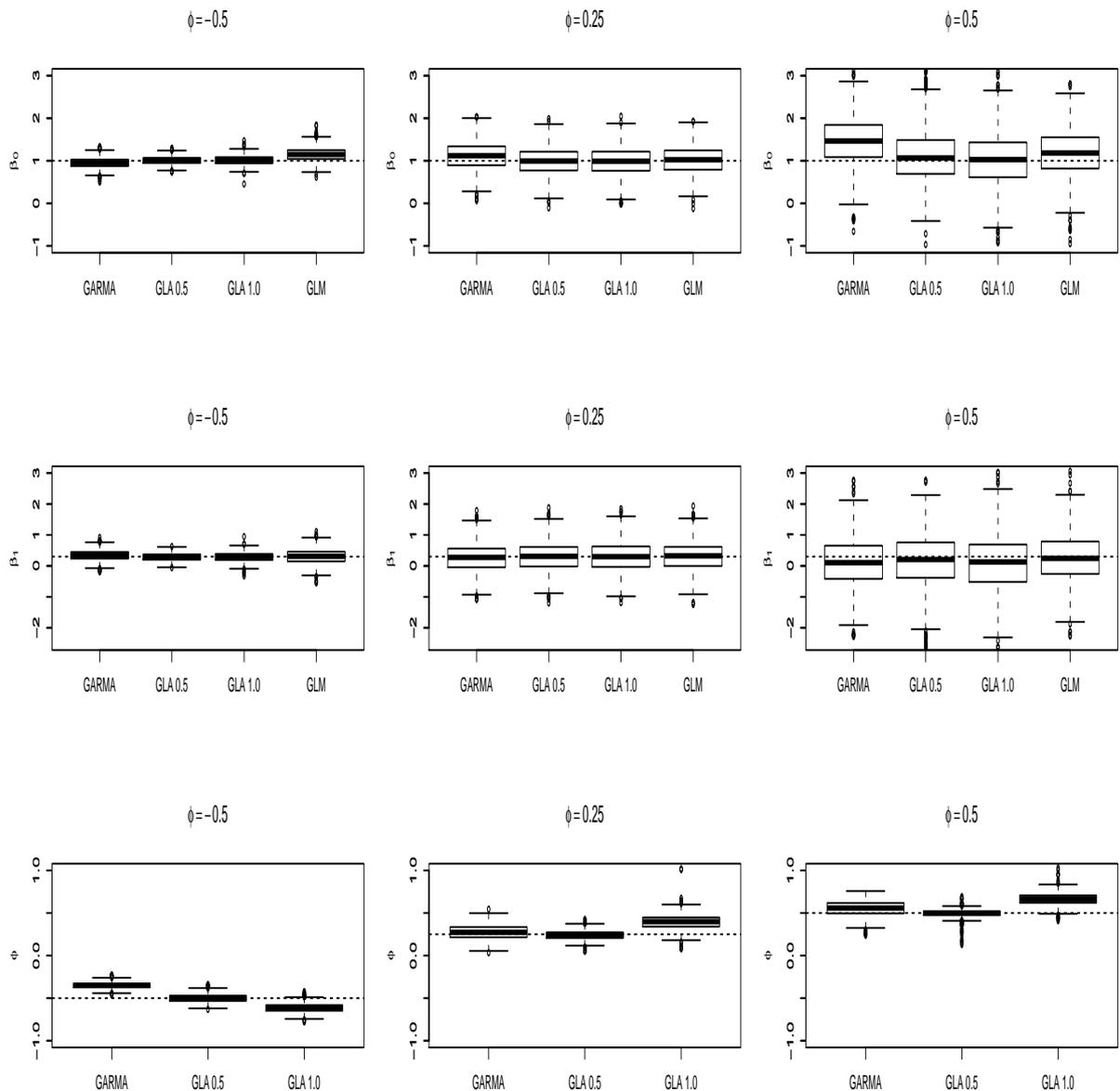


Figura 4.11: Boxplot para séries geradas pelo GLARMA AR(1) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 e os da terceira para o parâmetro ϕ .

A Figura 4.11 apresenta os resultados para o modelo com covariável. As melhores estimativas são obtidas pelo modelo GLARMA com $\lambda = 0,5$. É possível perceber que a variabilidade das estimativas de β_0 e β_1 aumenta com o valor do parâmetro ϕ . Também observa-se que o parâmetro β_1 é estimado adequadamente independente do

modelo ajustado, já para β_0 todos os modelos fornecem estimativas superestimadas para quando o parâmetro $\phi = 0,5$. E com relação ao parâmetro ϕ , as estimativas geradas pelo modelo GLARMA com $\lambda = 1,0$ para apresentam maior vício.

4.2.2 GLARMA(0,1)

As Tabelas 7.12 e 7.11 e as Figuras 4.13 e 4.12 apresentam os resultados das estimativas para séries geradas por um modelo GLARMA Poisson(0,1) sem e com covariável, respectivamente.

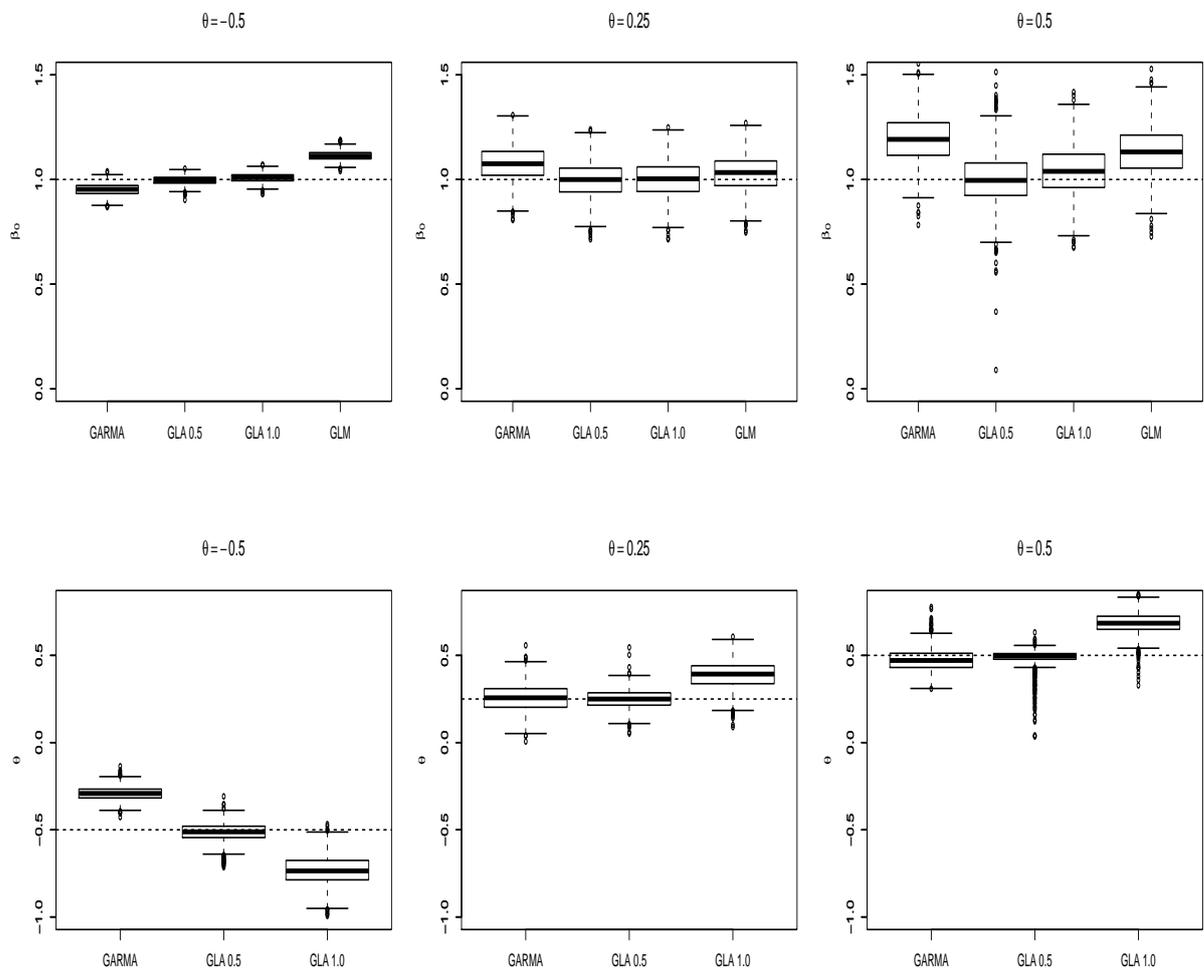


Figura 4.12: Boxplot para séries geradas pelo GLARMA MA(1) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 e da segunda linha para o parâmetro θ .

É possível perceber que o modelo que retorna as melhores estimativas é o modelo GLARMA com $\lambda = 0,5$, dos quais os dados foram gerados. Novamente, é notável que à medida que o valor do parâmetro θ aumenta, também aumenta a variabilidade das estimativas.

O modelo MLG fornece estimativas superestimadas para β_0 . Assim com o GARMA também resulta em estimativas com maior vício para β_0 , e para θ quando seu valor real é $-0,5$. E o modelo GLARMA com $\lambda = 1,0$ também encontra estimativas com maior vício para o parâmetro θ , para qualquer que seja seu valor real.

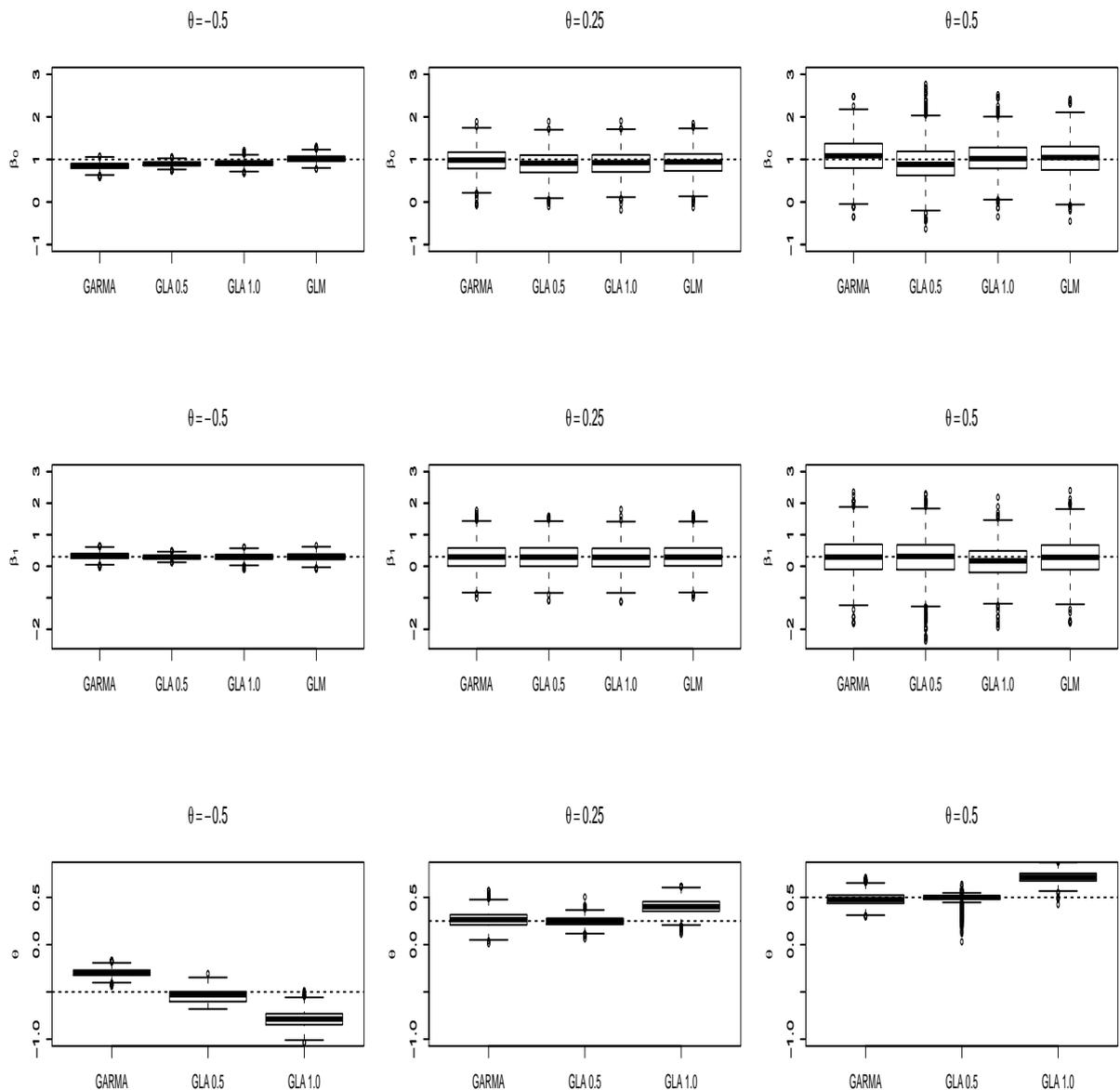


Figura 4.13: Boxplot para séries geradas pelo GLARMA MA(1) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 e os da terceira linha para θ .

Uma observação detalhada da Figura 4.13 para dados com covariável mostra uma maior variabilidade dos dados quando θ é maior. Nota-se boas estimativas para β_1 , com exceção do modelo GLARMA com $\lambda = 1,0$ quando o parâmetro $\theta = 0,5$. Assim como também é possível perceber que esse modelo fornece estimativas com maior vício para todos os valores de θ .

De maneira geral, o GARMA apresenta estimativas satisfatórias para os parâmetros, com exceção da estimativa de β_0 e θ , quando o modelo foi gerado com $\theta = -0,5$. Nessa simulação o modelo MLG ajustado para os dados também resulta em estimativas satisfatórias.

4.2.3 GLARMA(2,0)

Os resultados encontrados para as estimativas do modelo acima encontram-se nas Figuras 4.14 e 4.15, sem e com covariável, respectivamente .

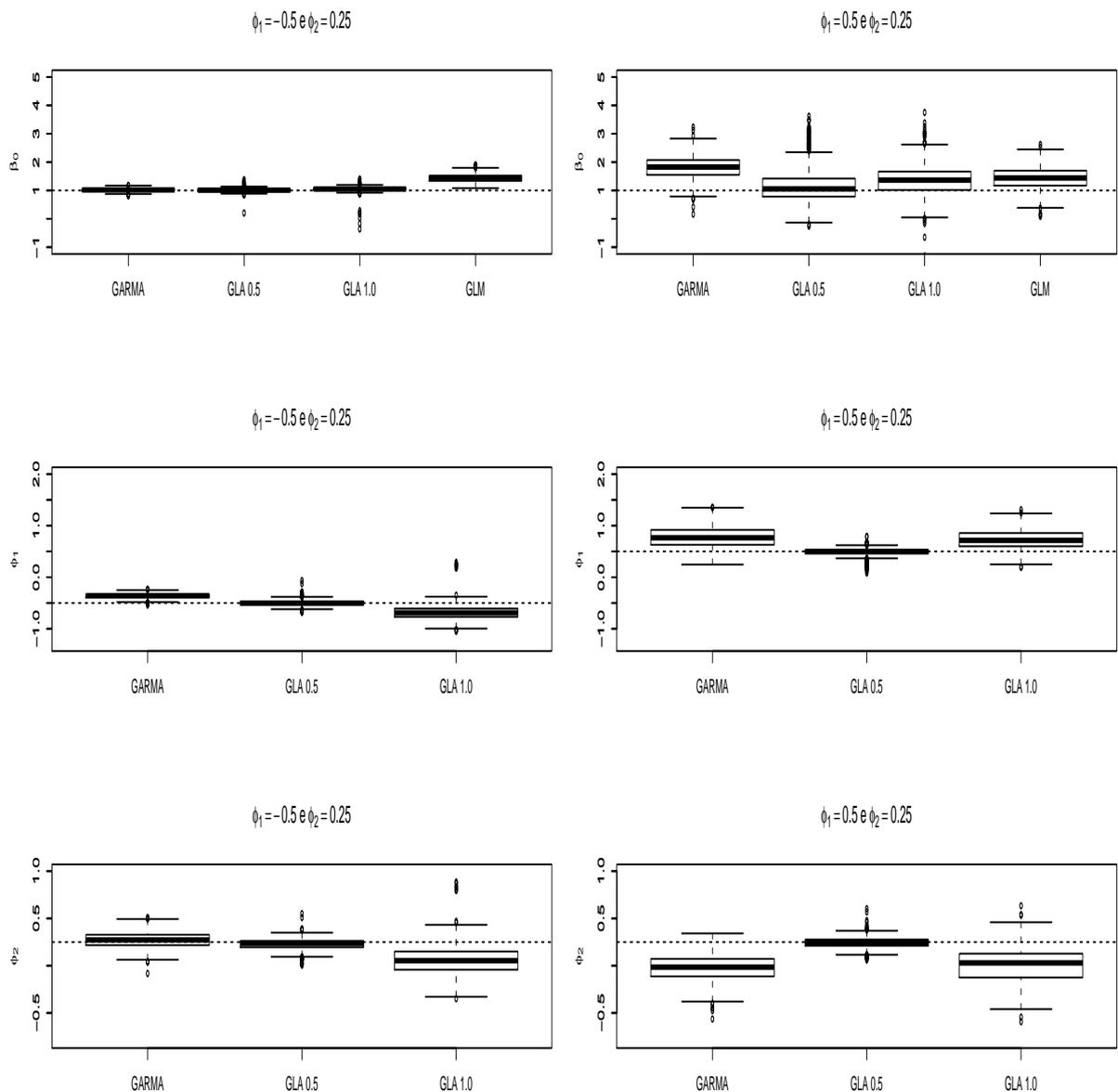


Figura 4.14: Boxplot para séries geradas pelo GLARMA AR(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro ϕ_1 e os da terceira para o parâmetro ϕ_2 .

As melhores estimativas encontradas foram as do modelo GLARMA, com $\lambda = 0,5$, dos quais os dados foram gerados.

As estimativas para todos os parâmetros obtidas pelo modelo GARMA, quando $\phi_1 = 0,5$ e $\phi_2 = 0,25$, apresentaram maior vício, assim como as estimativas obtidas pelo

modelo MLG para β_0 e as estimativas obtidas pelo modelo GLARMA com $\lambda = 1,0$ para ϕ_1 e ϕ_2 .

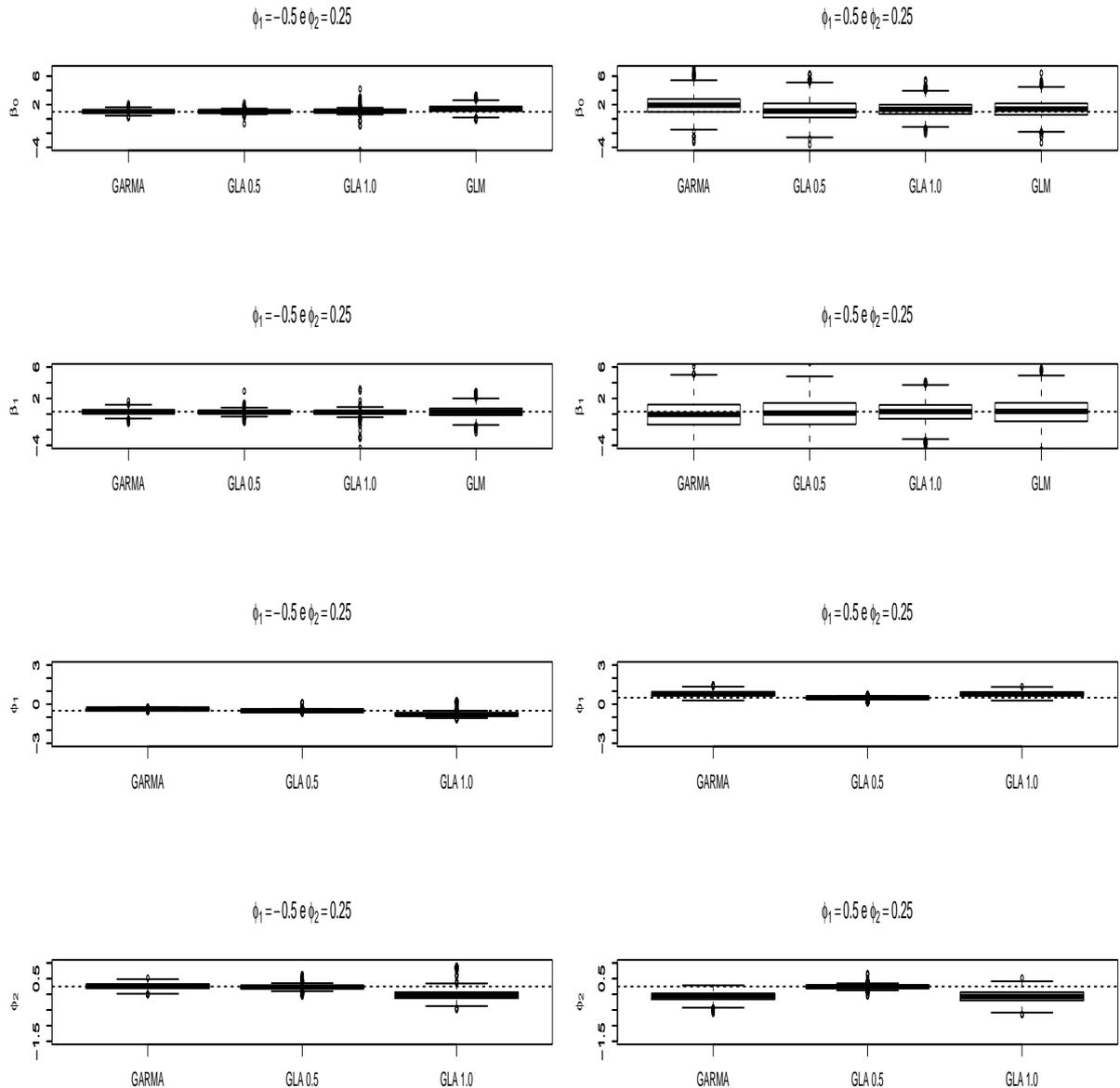


Figura 4.15: Boxplot para séries geradas pelo GLARMA AR(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira para o parâmetro ϕ_1 e os da quarta linha para o parâmetro ϕ_2 .

A Figura 4.15 apresenta os resultados para o modelo com covariável. De maneira geral, é possível perceber estimativas melhores quando os dados são gerados com $\phi_1 = -0,5$

e $\phi_2 = 0,25$. O modelo GARMA não apresenta estimativas satisfatórias para nenhum dos parâmetros, no caso em que $\phi_1 = 0,5$ e $\phi_2 = 0,25$. Assim, como quando $\phi_1 = -0,5$ e $\phi_2 = 0,25$, esse modelo não fornece estimativas satisfatórias para ϕ_1

O modelo MLG fornece estimativas com maior vício para β_0 , assim como o modelo GLARMA com $\lambda = 1,0$ gerou estimativas muito viciadas para todos os parâmetros.

4.2.4 GLARMA(0,2)

Nos Gráficos 4.17 e 4.16 são mostrados os resultados para as estimativas geradas para um modelo GLARMA(0,2), sem e com covariável, respectivamente.

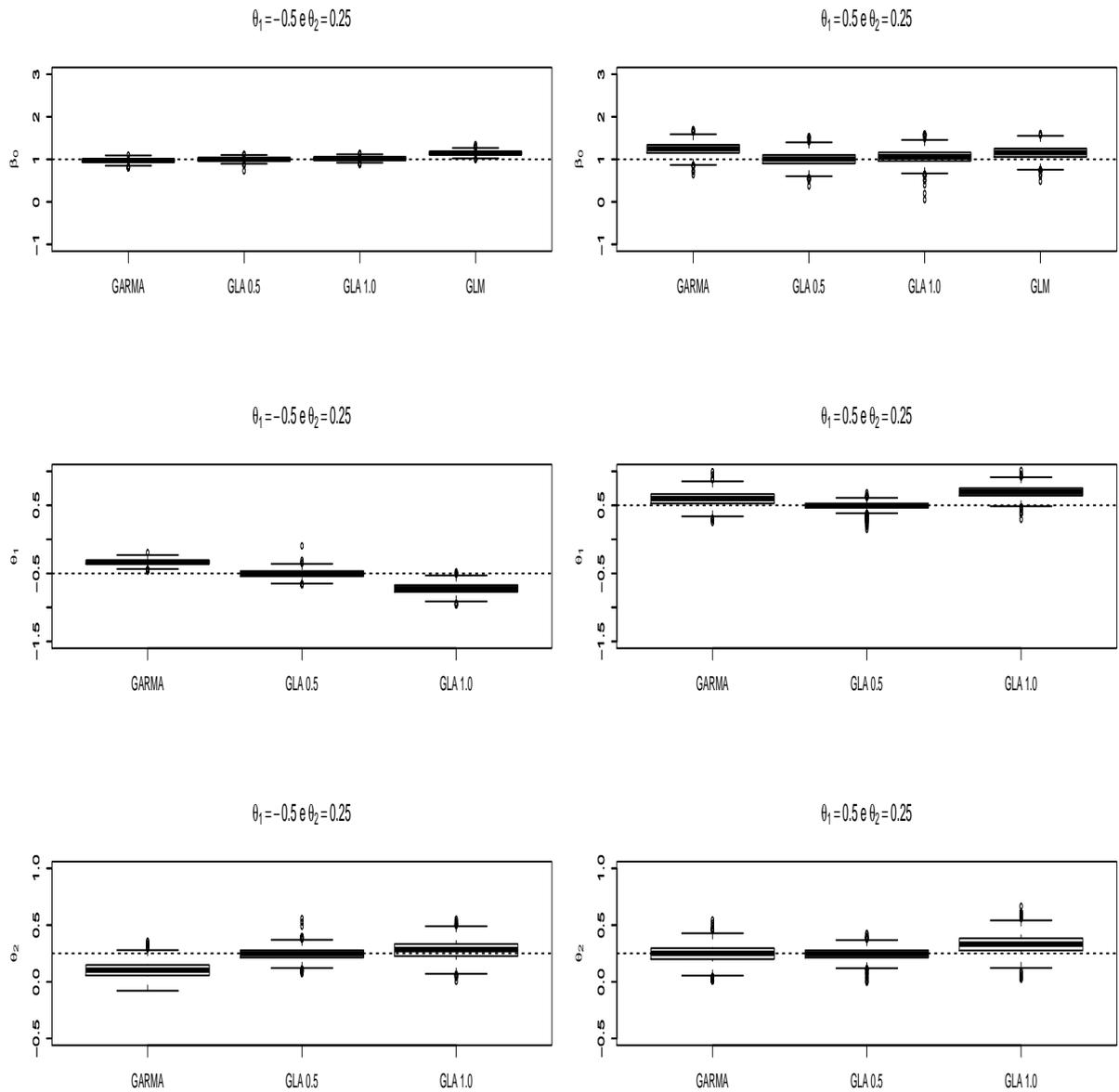


Figura 4.16: Boxplot para séries geradas pelo GLARMA MA(2) sem covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$, respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro θ_1 e os da terceira para o parâmetro θ_2

As melhores estimativas obtidas foram fornecidas pelo modelo GLARMA com $\lambda = 0,5$. Já o modelo MLG não fornece estimativas satisfatórias para β_0 .

O modelo GLARMA com $\lambda = 1,0$ retorna estimativas com maior vício para θ_1 e θ_2 independente de seu valor real. Assim como o modelo GARMA não fornece boas

estimativas para os parâmetros, com exceção do parâmetro β_0 , quando os dados são gerados por $\theta_1 = -0,5$ e $\theta_2 = 0,25$, e do θ_2 quando os dados são gerados por $\theta_1 = 0,5$ e $\theta_2 = 0,25$.

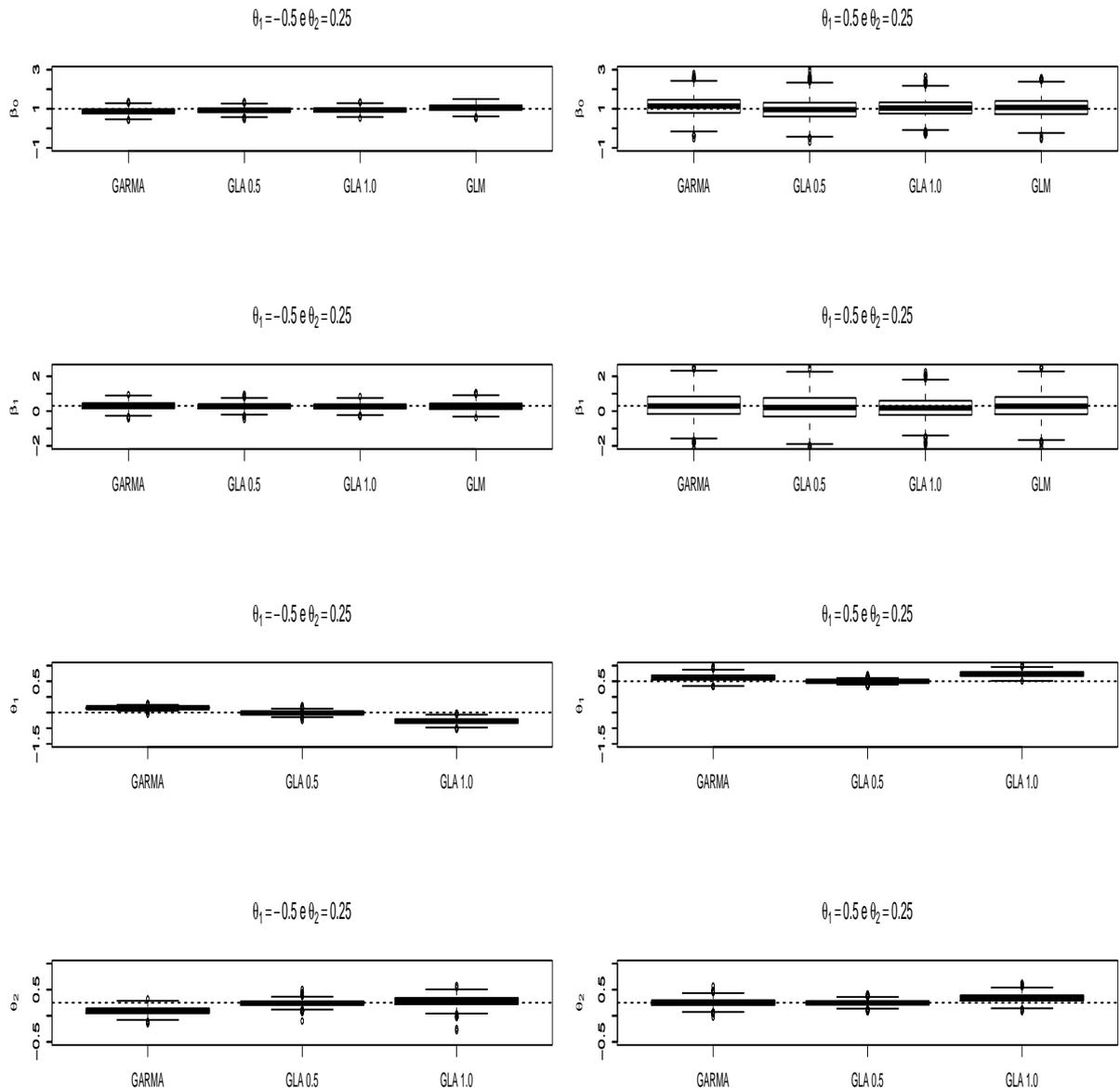


Figura 4.17: Boxplot para séries geradas pelo GLARMA MA(2) com covariável. Nos Gráficos Boxplot as siglas GLA0,5 e GLA1,0 representam as médias das estimativas geradas pelo modelo GLARMA com parâmetro $\lambda = 0,5$ e $1,0$ respectivamente. Os gráficos da primeira linha apresentam as estimativas para o parâmetro β_0 , os da segunda linha para o parâmetro β_1 , os da terceira linha para o parâmetro θ_1 e os da quarta linha para o parâmetro θ_2

A Figura 4.17 apresenta os resultados para o modelo com covariável. É possível

perceber uma variabilidade maior para β_0 e β_1 quando os parâmetros θ são ambos positivos. As estimativas obtidas para β_0 são todas subestimadas quando $\theta_1 = -0,5$ e $\theta_2 = 0,25$ e superestimadas para $\theta_1 = 0,5$ e $\theta_2 = 0,25$.

O modelo GARMA não fornece boas estimativas para o parâmetro θ_1 quando o valor real é $\theta_1 = -0,5$ e $\theta_2 = 0,25$, já o modelo MLG estimou satisfatoriamente os parâmetros.

Em relação ao parâmetro β_1 , as estimativas encontradas são boas quando $\theta_1 = -0,5$ e $\theta_2 = 0,25$, já no segundo caso, os dois modelos GLARMA retornam estimativas com vício negativo.

Capítulo 5

Aplicação a dados reais

Neste capítulo seguem duas aplicações a dados reais para ilustrar o uso dos modelos descritos neste trabalho.

5.1 Poliometite

Os dados referem-se ao número mensal de casos de poliometite nos EUA nos anos de 1970 a 1983. Na introdução deste trabalho os mesmos foram analisados pelo MLG e a suposição de resíduos não correlacionados não foi satisfeita ver Figura 1.2. Portanto agora serão ajustados outros modelos, GARMA e GLARMA com $\lambda = 0,5$ e $\lambda = 1,0$, na tentativa de eliminar a correlação encontrada nos resíduos pelo MLG.

Com o objetivo de comparação dos modelos, serão feitas previsões para valores futuros da série. Desta forma, as últimas 6 observações (jul/83 a dez/83) serão retiradas da análise.

Tratando-se de uma série de contagem, a distribuição suposta para os dados é Poisson. As covariáveis utilizadas foram seno e cosseno, anual e semestral, para modelar a sazonalidade e um componente de tendência linear.

Na Figura 5.1 estão a FAC e a FACP da série do número de casos de poliometite. É possível perceber no lag 1 um pico no gráfico da FACP e um decrescimento na FAC, o que

de acordo com Box & Jenkins (1976) sugere um modelo com componente autoregressivo de ordem 1.

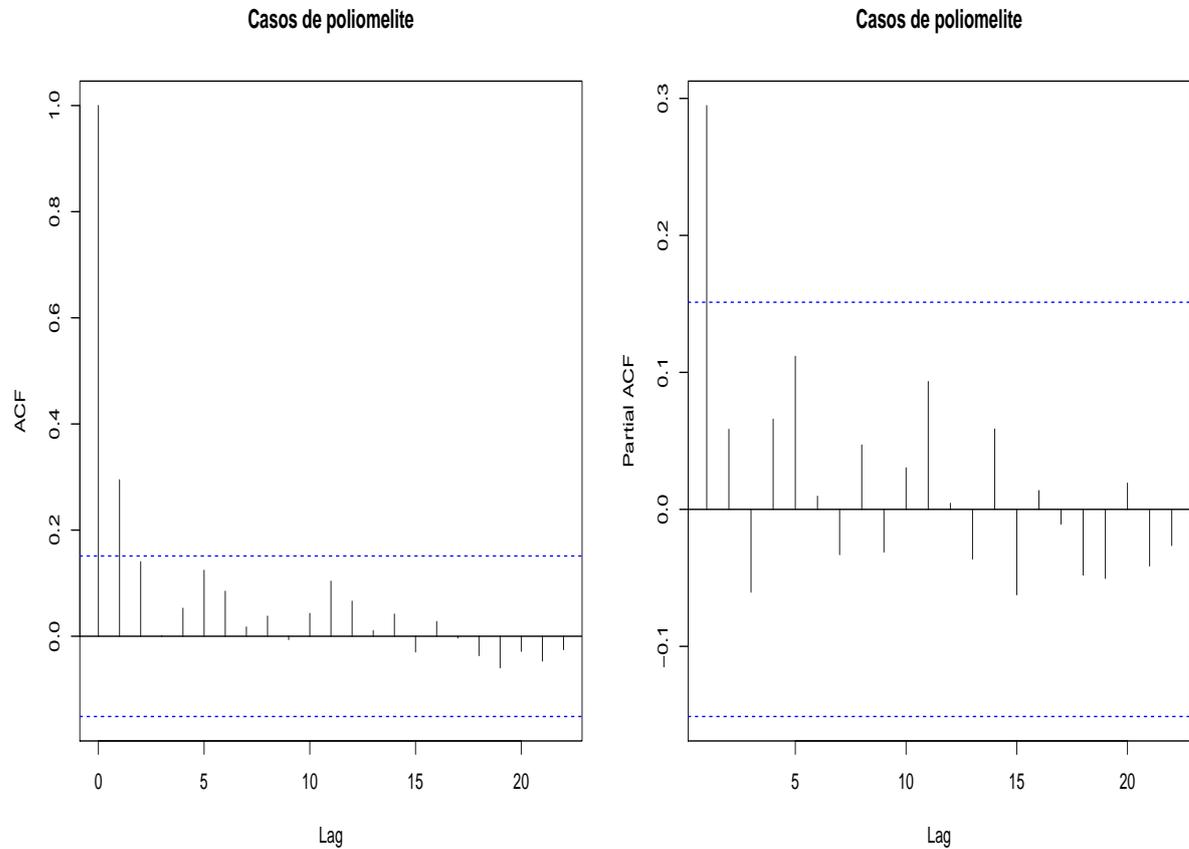


Figura 5.1: Autocorrelação e autocorrelação parcial da série dos casos de poliometite

Desta forma, foram ajustados aos dados modelos GARMA e GLARMA com um componente AR(1) e as covariáveis citadas anteriormente. Na tabela 5.1 encontram-se os AIC de cada modelo. O modelo que apresentou menor AIC foi GLARMA Poisson AR(1) com $\lambda = 1,0$. Porém uma análise de resíduos mostrou que esse modelo não conseguiu eliminar a correlação presente nos mesmos, assim como os modelos GARMA ajustados também apresentaram resíduos correlacionados.

Tabela 5.1: AIC dos modelos ajustados para os dados de casos de poliomielite

GLM	557,90
GARMA AR(1)	541,79
GLARMA AR(1) $\lambda = 0,5$	538,35
GLARMA AR(1) $\lambda = 1,0$	533,97

Como o modelo GLARMA Poisson AR(1) com $\lambda = 0,5$ apresentou o segundo menor AIC e resíduos não correlacionados, serão apresentadas as estimativas para os parâmetros deste modelo (ver Tabela 5.2), assim como a análise de resíduos detalhada.

Tabela 5.2: Ajuste GLARMA AR(1) para os dados de Poliomielite

	Estimativa	Erro padrão	p-valor
Intercepto	0,137	0,103	0,184
Tendência	-4,227	1,950	0,030
Cosseno anual	-0,121	0,122	0,323
Seno anual	-0,543	0,145	0,000
Cosseno semestral	0,278	0,113	0,014
Seno semestral	-0,413	0,114	0,000
AR(1)	0,237	0,046	0,000

Com exceção do cosseno anual, todas as outras covariáveis são significativas para o modelo. Entretanto a covariável cosseno anual não será retirada, já que o seno anual apresentou um coeficiente significativo.

Conforme foi observado nas simulações na Seção 4 era esperado que as estimativas dos parâmetros β fossem semelhantes às estimativas obtidas pelo MLG, que se encontram na Tabela 1.1, com exceção do β_0 .

A análise de resíduos deste modelo é apresentada na Figura 5.2. O modelo ajustado apresentou alguns outliers nos resíduos, mas em geral estes apresentam um comportamento aleatório em torno de zero.

É possível perceber através da análise da FAC e FACP dos resíduos, ver Figura 5.3, que nesse ajuste não existe correlação residual, o que indica que o modelo conseguiu capturar a dependência temporal dos dados, já que trata-se de uma série temporal. Esta é uma das principais vantagens do ajuste destes modelos em relação ao MLG em que a suposição de não correlação não foi satisfeita.

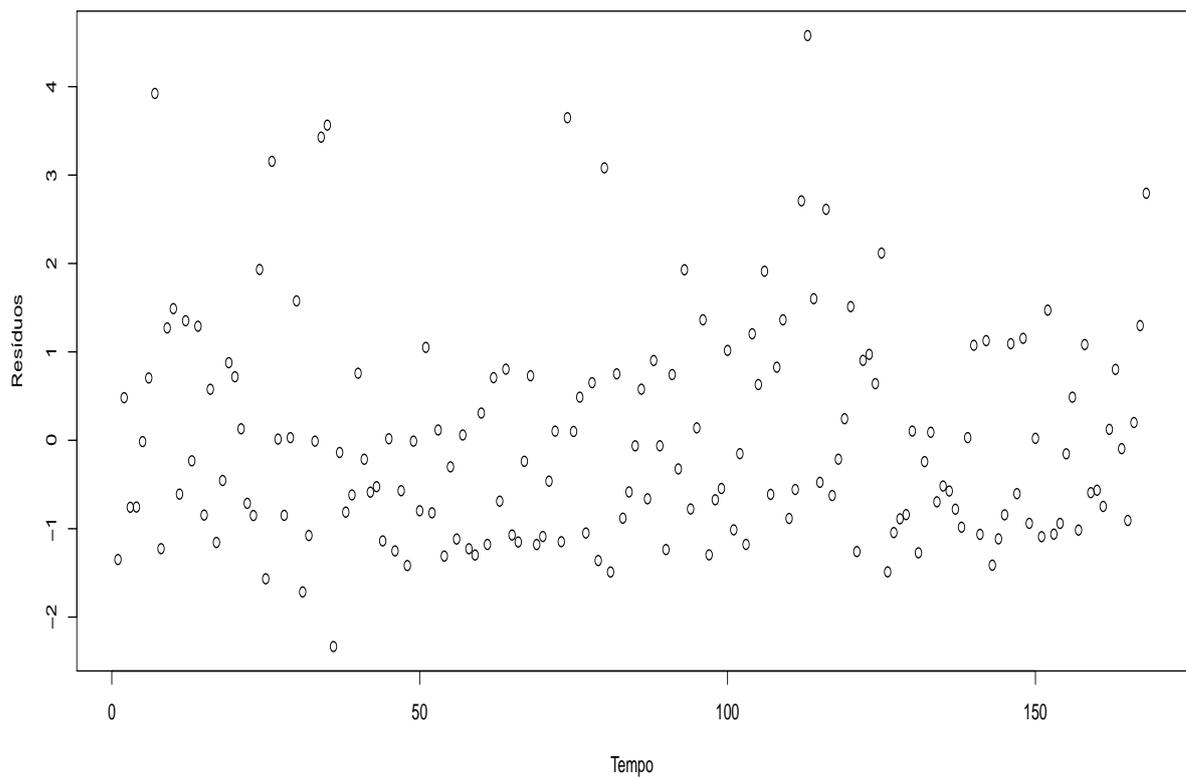


Figura 5.2: Análise dos resíduos para o modelo GLARMA AR(1) ajustado à série de número mensal de casos de poliometite

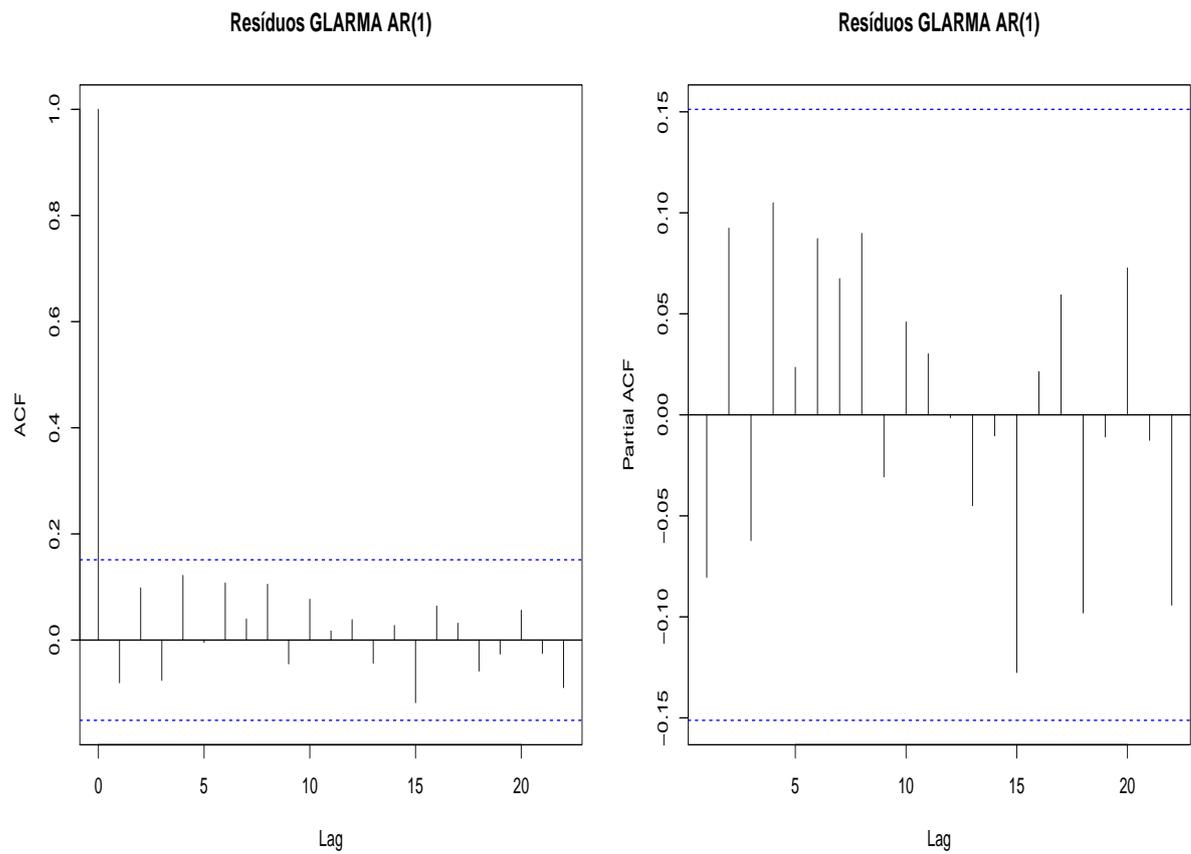


Figura 5.3: Autocorrelograma e autocorrelograma parcial dos resíduos do ajuste à série de número mensal de casos de poliomelite

A Figura 5.4 apresenta a série temporal dos dados de poliomelite e o modelo ajustado. É possível perceber que o modelo descreve bem os dados da série.

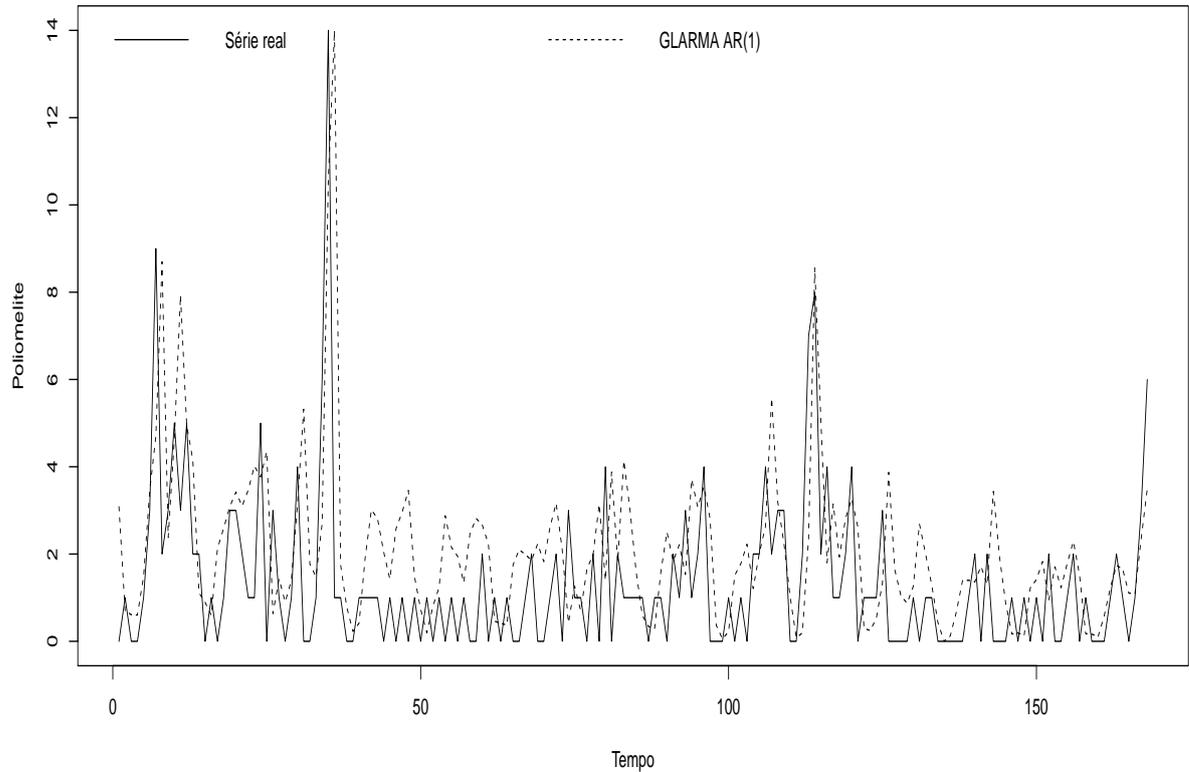


Figura 5.4: Série temporal do número mensal de casos de poliomelite. A linha contínua representa a série real e a linha tracejada o modelo GLARMA AR(1) ajustado à série.

Com o intuito de comparar os modelos com relação à habilidade de predição de valores futuros, serão feitas previsões 1 passo à frente para todos os modelos ajustados à série de casos de poliomelite.

No modelo GLARMA AR(1) ajustado, a previsão a 1 passo à frente é dada por, $\hat{y}_t(1) = \mu_{t+1}$, onde,

$$\mu_{t+1} = \exp(x'_{t+1}\beta + E_t[Z_{t+1}]), \quad (5.1)$$

e

$$E_t[Z_{t+1}] = \phi_1(Z_t + \varepsilon_t). \quad (5.2)$$

Para se obter as previsões do componente Z_{t+1} serão utilizados os procedimentos de

previsão descritos na Seção 3.2.3.

Para o modelo MLG é feita uma substituição dos parâmetros estimados e $\hat{y}_t(1)$ resulta em

$$\hat{y}_t(1) = \mu_{t+1} = \exp(x'_{t+1}\beta). \quad (5.3)$$

No modelo GARMA a previsão é feita utilizando-se dos procedimentos já citados na Seção 3.9. Para um GARMA AR(1) têm-se $\hat{y}_t(l)$

$$\hat{y}_t(l) = \mu_{t+1} = \exp(x'_{t+1}\beta + \phi_1[\log(y_t^*) - x'_t\beta]), \quad (5.4)$$

onde $y_t^* = \max(y_t, c)$, com $0 < c < 1$.

A Tabela 5.3 mostra os resultados obtidos para os modelos GLARMA AR(1) com $\lambda = 0,5$ (GLA 0,5), GLARMA AR(1) com $\lambda = 1,0$ (GLA 1,0), GARMA AR(1) (GA) e MLG ajustados. A primeira coluna refere-se ao valor real.

Tabela 5.3: Previsão para dados de poliomelite utilizando os modelos GLARMA AR(1), GARMA e MLG

	Valor Real	MLG	GLA 0,5	GLA 1,0	GA
$\hat{Y}_{162}(1)$	2,00	0,95	1,00	1,19	0,38
$\hat{Y}_{163}(1)$	1,00	0,82	1,03	1,18	0,38
$\hat{Y}_{164}(1)$	0,00	0,82	0,78	0,80	0,31
$\hat{Y}_{165}(1)$	1,00	0,80	0,74	0,62	0,23
$\hat{Y}_{166}(1)$	3,00	0,99	1,23	1,22	0,49
$\hat{Y}_{167}(1)$	6,00	1,27	1,73	1,96	0,61
EQM de previsão		4,71	3,84	3,49	6,5

É possível perceber que as previsões dos modelos GLARMA AR(1) resultaram em valores mais próximos dos reais e esse mesmo modelo com $\lambda = 1,0$ resultou em um EQM

de previsão menor. Neste caso, o modelo que forneceu o melhor ajuste foi também o que apresentou melhores previsões.

5.2 Falência de empresas Norte Americanas

Utilizando o número mensal de falências de empresas Norte Americanas no período compreendido entre janeiro de 1985 e dezembro de 2012 disponíveis em UCLA-LoPucki Bankruptcy Research Database, e dados macroeconômicos obtidos em St. Louis Fed online foram feitos ajustes GARMA e GLARMA afim de encontrar o modelo que melhor se adequa aos dados.

A variável resposta é a contagem do número mensal de casos de falências, e as covariáveis são os Índices de Produção Industrial (INDPRO), Novas licenças de habitação (PERMIT), Títulos de investimento(BAA), um Índice publicado pelo Conselho da Reserva Federal com base no rendimento médio de uma série de títulos do Tesouro, ajustado para um prazo de 10 anos (GS10), Taxa dos fundos federais (FEDFUNDS), Índice de Preços ao Produtor: Bens Finais (PPIFGS).

A Figura 5.5 apresenta a série temporal do número de empresas que decretaram falência entre 1985 e 2012.

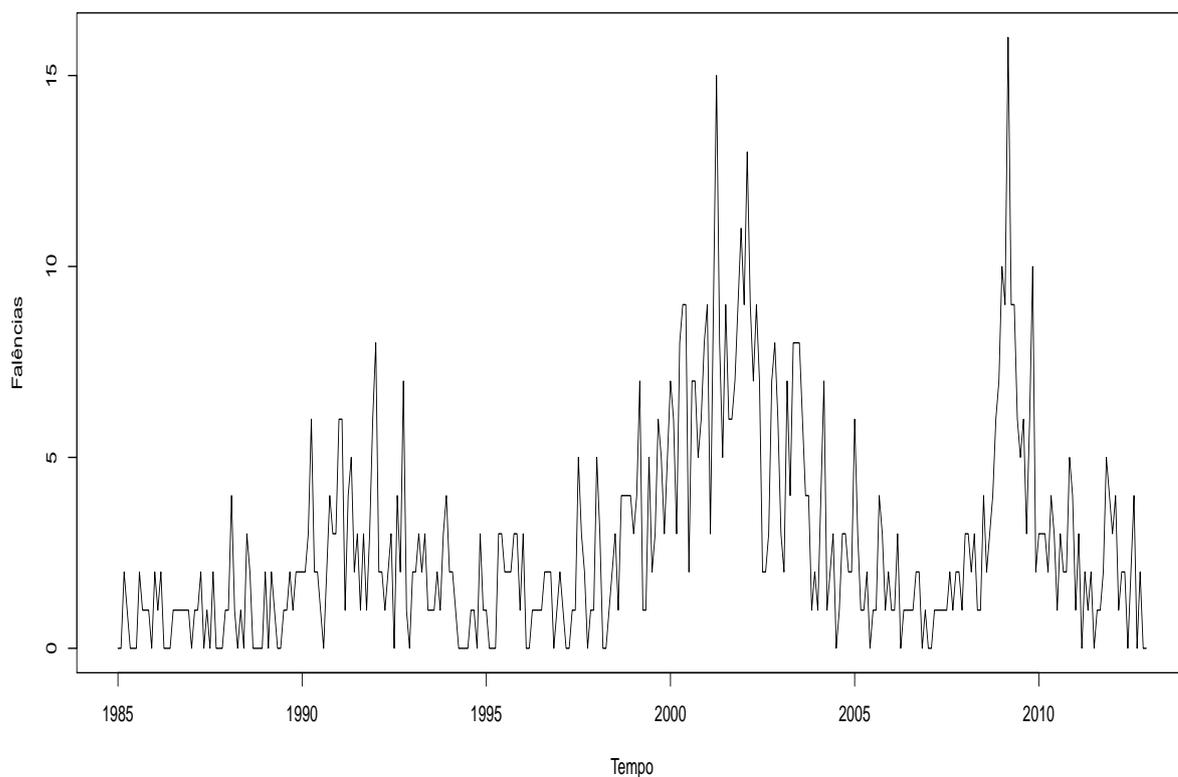


Figura 5.5: Série temporal do número de casos de falências de empresas Norte Americanas

É possível perceber dois picos na série, um referente à abril de 2001 e outro em março de 2009 com 15 e 16 empresas que foram à falência, respectivamente. Por isso, foram acrescentadas duas covariáveis indicadoras ao modelo, que nas tabelas encontram-se como Ind1 e Ind 2, referentes ao primeiro e segundo períodos citados, respectivamente.

Os dados referem-se a uma série temporal de dados de contagem, por isso a distribuição utilizada para Y_t é Poisson.

A Figura 5.6 apresenta as funções de autocorrelação e autocorrelação parcial do banco de dados de falência. É possível perceber um decaimento exponencial na FAC e picos na FACP, que de acordo com Box & Jenkins(1976) sugerem uma estrutura autoregressiva de ordem 3.

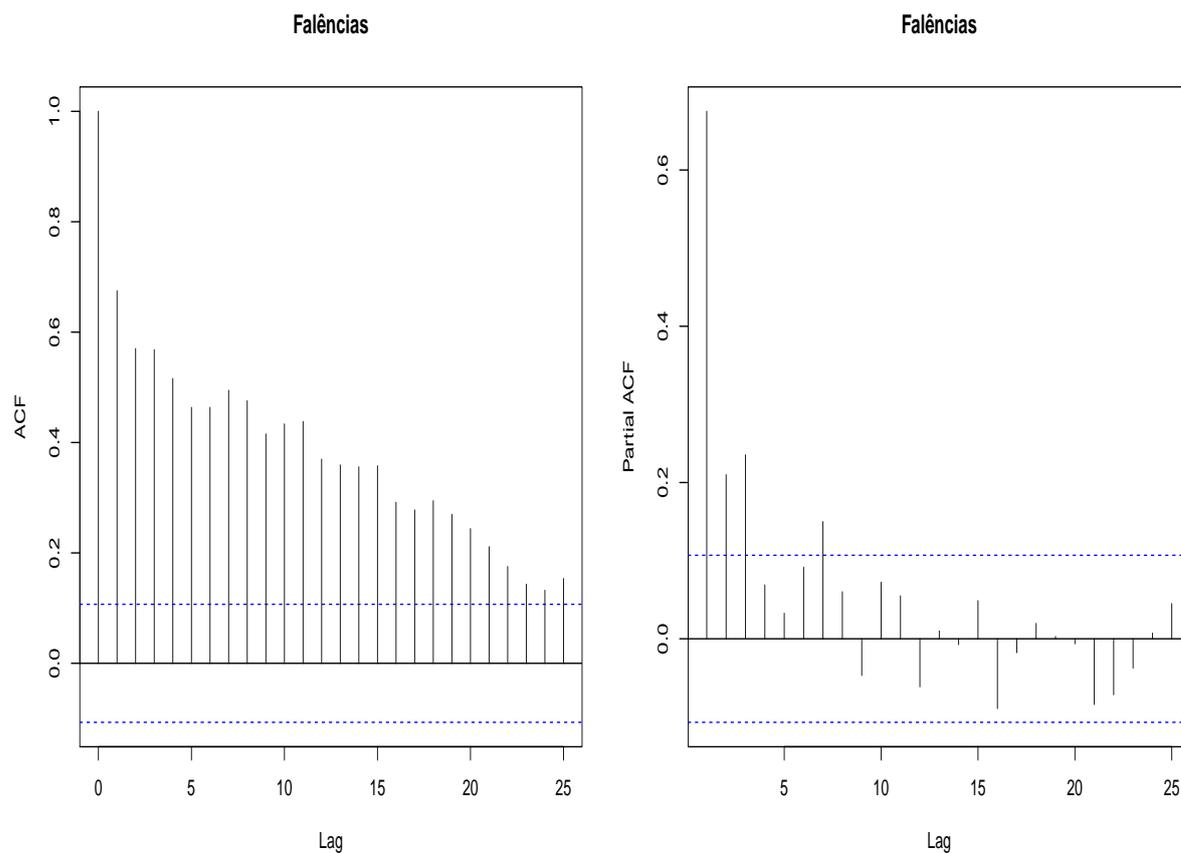


Figura 5.6: Função de autocorrelação e autocorrelação parcial do número de empresas que decretaram falência.

Assim, os modelos GARMA e GLARMA de ordem AR(3) foram ajustados e os AIC encontram-se na Tabela 5.4.

Tabela 5.4: AIC dos modelos ajustados à série do número de empresas Norte Americanas que decretaram falência

GARMA	1254,01
GLARMA $\lambda = 0,5$	1232,47
GLARMA $\lambda = 1,0$	1247,27

Utilizando-se do AIC, o melhor modelo encontrado foi um GLARMA AR(3) com $\lambda = 0,5$. Como este modelo foi o que apresentou os resíduos mais adequados, suas estimativas encontram-se na Tabela 5.5.

Tabela 5.5: Ajuste GLARMA AR(3) para o número de casos de falências de empresas Norte Americanas

	Estimativa	p-valor
Intercepto	3,330	0,033
INDPRO	0,063	0,000
PERMIT	-0,001	0,002
BAA	0,337	0,000
GS10	-0,385	0,001
FEDFUNDS	-0,100	0,053
PPIFGS	-0,046	0,000
Ind1	0,610	0,029
Ind2	0,466	0,042
AR(1)	0,199	0,000
AR(2)	0,074	0,019
AR(3)	0,062	0,049

Todas as variáveis foram consideradas significativas para a análise, apesar do coeficiente da variável FEDFUNDS estar no limite do nível de significância adotado de 5%.

Os gráficos de resíduos do modelo GLARMA AR(3) ajustado podem ser vistos nas Figuras 5.7 e 5.8.

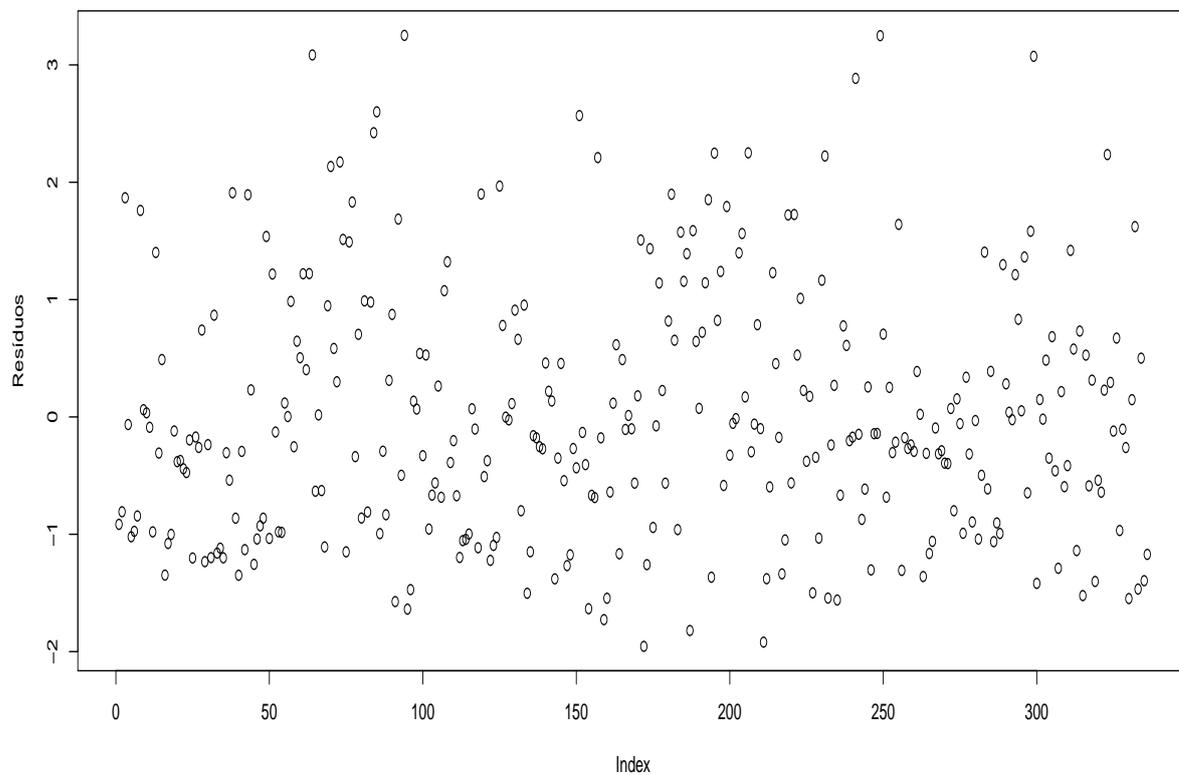


Figura 5.7: Resíduos GLARMA AR(3)

Os resíduos parecem ser homoscedásticos (Figura 5.7) já que estão aleatoriamente distribuídos em torno de zero.

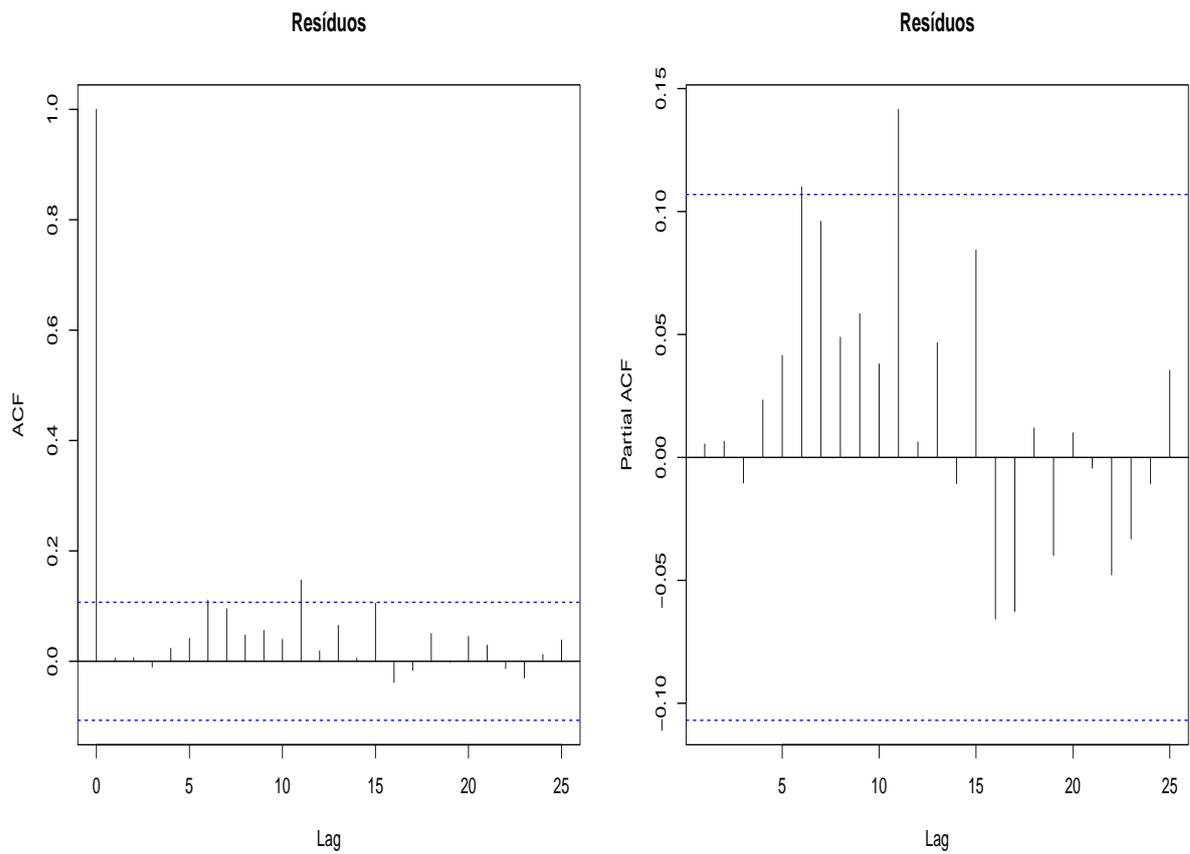


Figura 5.8: FAC e FACP GLARMA AR(3)

A FACP na Figura 5.8 mostra apenas um pico significativo no lag 11, que pode ser justificado como perturbações aleatórias ao modelo, mas é possível perceber a não correlação dos resíduos, o que indica uma adequação dos dados ao modelo proposto.

A Figura 5.9 apresenta a série temporal dos casos de falência e o modelo ajustado. É possível perceber que o modelo descreve bem os dados da série.

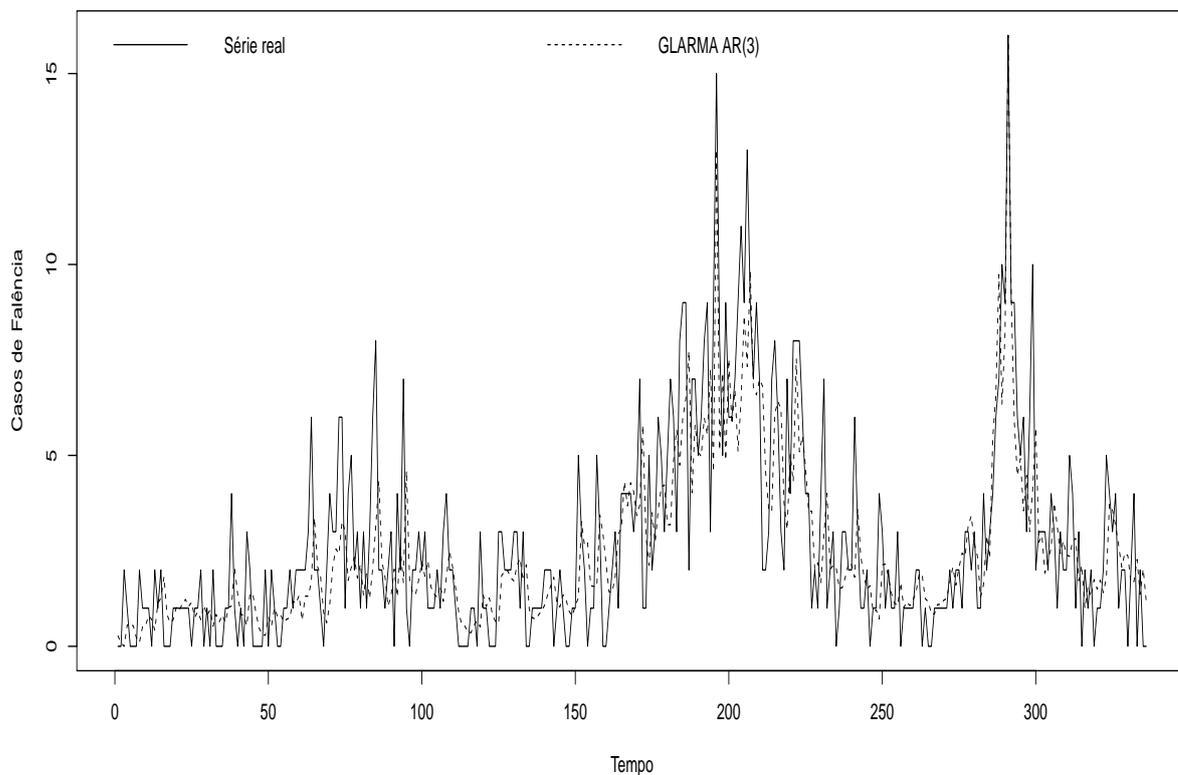


Figura 5.9: Série temporal do número de casos de falência. A linha contínua representa a série real e a linha tracejada o modelo GLARMA AR(3) ajustado à série.

Assim como no exemplo anterior, afim de comparar a capacidade de previsão dos modelos, serão feitas previsões a 1 passo a frente. Foram retiradas as últimas seis observações (julho/2012 a dezembro/2012) e usando cada um dos modelos serão calculadas previsões e seus EQM de previsão.

Para GLARMA AR(3) foram utilizados os mesmos procedimentos descritos em 3.2.3. A equação de previsão é dada por

$$\mu_{t+1} = \exp(x'_{t+1}\beta + E_t[Z_{t+1}]), \quad (5.5)$$

e

$$E_t[Z_{t+1}] = \phi_1(E_t[Z_t] + E_t[\varepsilon_t]) + \phi_2(E_t[Z_{t-1}] + E_t[\varepsilon_{t-1}]) + \phi_3(E_t[Z_{t-2}] + E_t[\varepsilon_{t-2}]). \quad (5.6)$$

A equação de previsão para um modelo GARMA AR(3), já citada na Seção 3.9, é dada por

$$\hat{y}_t(l) = \mu_{t+1} = \exp(x'_{t+1}\beta + \phi_1[\log(y_t^*) - x'_t\beta] + \phi_2[\log(y_{t-1}^*) - x'_{t-1}\beta] + \phi_3[\log(y_{t-2}^*) - x'_{t-2}\beta]). \quad (5.7)$$

onde $y_t^* = \max(y_t, c)$, com $0 < c < 1$.

A Tabela 5.6 apresenta a previsão para os modelos GARMA (GA), GLARMA com $\lambda = 0,5$ (GLA 0,5), GLARMA com $\lambda = 1,0$ (GLA 1,0) e MLG.

Tabela 5.6: Previsão para dados de falência utilizando os modelos GLARMA AR(3) e GARMA AR(3)

	Valor real	GLA 0,5	GLA 1,0	GA	MLG
$\hat{Y}_{331}(1)$	2,00	2,32	2,44	0,73	2,33
$\hat{Y}_{332}(1)$	4,00	1,93	2,24	0,92	1,93
$\hat{Y}_{333}(1)$	0,00	1,62	2,58	1,31	1,62
$\hat{Y}_{334}(1)$	2,00	1,45	1,90	0,68	1,45
$\hat{Y}_{335}(1)$	0,00	1,62	2,41	1,04	1,63
$\hat{Y}_{336}(1)$	0,00	1,65	1,80	0,56	1,66
EQM de previsão		3,15	3,16	2,65	2,11

Neste caso, embora o melhor ajuste tenha sido o do modelo GLARMA AR(3) com $\lambda = 0,5$, o modelo GLM foi o que apresentou as melhores previsões.

Capítulo 6

Conclusão e considerações finais

Neste trabalho foi proposto o ajuste de modelos GARMA e GLARMA para séries temporais para dados de contagem com distribuição Poisson.

As estimativas dos parâmetros foram verificadas através de séries simuladas. Os resultados demonstram que as estimativas possuem maior variabilidade à medida que o parâmetro θ ou ϕ cresce no caso onde só existe AR(1) ou MA(1). De maneira geral, as estimativas para os parâmetros são semelhantes para o modelo GLARMA com $\lambda = 0,5$ e $\lambda = 1,0$, quando as séries são geradas de um modelo GARMA, já no caso em que as séries são geradas por um GLARMA com $\lambda = 0,5$ algumas estimativas geradas pelo modelo GLARMA com $\lambda = 1,0$ são mais viciadas. As estimativas encontradas para β_1 , em grande maioria, são bem próximas do valor real fixado em $\beta_1 = 0,3$, independente do modelo. Em contrapartida, para β_0 , em muitas simulações foram encontradas estimativas com maior vício, principalmente geradas do modelo MLG.

As abordagens propostas neste trabalho foram aplicadas na análise de duas séries reais. A primeira, dados mensais de poliomelite de um hospital dos Estados Unidos que teve como melhor modelo GLARMA(1,0) que apresentou o menor AIC entre os modelos MLG e GARMA e a análise de resíduos do modelo não apresentou autocorrelação. A segunda série temporal real tratou-se de dados econômicos à cerca do número mensal de empresas que decretaram falência nos Estados Unidos no período de 1985 a 2012, novamente usando o critério AIC o melhor ajuste foi de um modelo GLARMA AR(3) e

também a análise de resíduos indicou que a dependência dos dados havia sido capturada pelo modelo, não apresentando autocorrelação significativa.

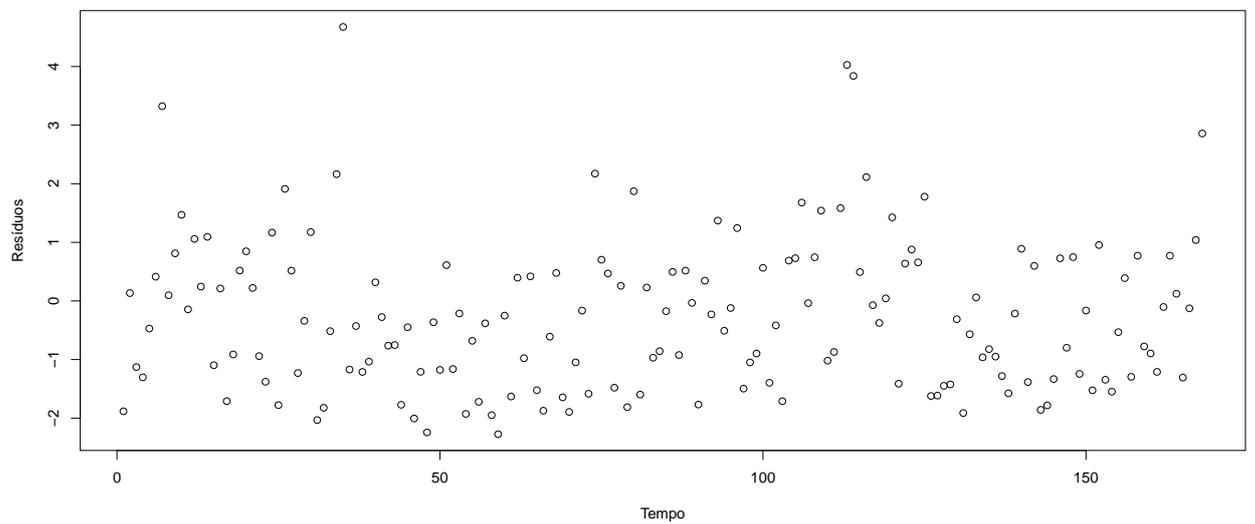
Trabalhos futuros incluem estimar o vício das estimativas dos parâmetros geradas pelos modelos GARMA e GLARMA usando Bootstrap e efetuar as correções necessárias, construir intervalos de confiança para os parâmetros usando o bootstrap, além de avaliar as aplicações destes modelos em conjuntos de dados reais usando distribuição Binomial Negativa.

Capítulo 7

Anexo

7.1 Anexo 1

Figura 7.1: Resíduos do ajuste MLG dos casos de poliomelite



7.2 Anexo 2

Tabela 7.1: Estimativas para séries geradas pelo GARMA AR(1) sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,325 (0,112)	0,998 (0,001)	1,092 (0,011)	1,118 (0,017)
$\phi = -0,5$		-0,497 (0,001)	-0,554 (0,010)	-0,676 (0,036)
$\beta_0 = 1,0$	0,917 (0,015)	0,994 (0,007)	0,896 (0,020)	0,895 (0,021)
$\phi = 0,25$		0,239 (0,005)	0,183 (0,008)	0,283 (0,010)
$\beta_0 = 1,0$	0,777 (0,075)	0,982 (0,018)	0,683 (0,136)	0,702 (0,120)
$\phi = 0,5$		0,480 (0,007)	0,358 (0,026)	0,506 (0,007)

Tabela 7.2: Estimativas para séries geradas pelo GARMA AR(1) com covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,311 (0,185)	0,982 (0,016)	1,099 (0,030)	1,106 (0,046)
$\beta_1 = 0,3$	0,219 (0,183)	0,322 (0,030)	0,257 (0,041)	0,266 (0,069)
$\phi = -0,5$		-0,499 (0,001)	-0,505 (0,009)	-0,699 (0,047)
$\beta_0 = 1,0$	0,906 (0,103)	0,990 (0,083)	0,890 (0,111)	0,885 (0,117)
$\beta_1 = 0,3$	0,336 (0,192)	0,298 (0,167)	0,339 (0,199)	0,343 (0,208)
$\phi = 0,25$		0,229 (0,006)	0,154 (0,012)	0,266 (0,009)
$\beta_0 = 1,0$	0,759 (0,350)	0,990 (0,214)	0,693 (0,450)	0,698 (0,455)
$\beta_1 = 0,3$	0,362 (0,565)	0,267 (0,406)	0,352 (0,686)	0,349 (0,693)
$\phi = 0,5$		0,465 (0,007)	0,316 (0,039)	0,510 (0,008)

Tabela 7.3: Estimativas para séries geradas pelo modelo GARMA MA(1) sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,283 (0,084)	1,000 (0,001)	1,192 (0,039)	1,204 (0,044)
$\theta = -0,5$		-0,500 (0,002)	-0,327 (0,033)	-0,609 (0,030)
$\beta_0 = 1,0$	0,946 (0,010)	0,995 (0,006)	0,929 (0,012)	0,929 (0,012)
$\theta = 0,25$		0,247 (0,006)	0,172 (0,010)	0,273 (0,009)
$\beta_0 = 1,0$	0,927 (0,015)	0,988 (0,009)	0,877 (0,027)	0,875 (0,028)
$\theta = 0,5$		0,502 (0,005)	0,289 (0,049)	0,458 (0,011)

Tabela 7.4: Estimativas para séries geradas pelo modelo GARMA MA(1) com covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,014 (0,007)	0,849 (0,029)	0,891 (0,014)	0,912 (0,013)
$\beta_1 = 0,3$	0,302 (0,013)	0,287 (0,020)	0,298 (0,004)	0,295 (0,010)
$\theta = -0,5$		-0,296 (0,043)	-0,545 (0,006)	-0,786 (0,089)
$\beta_0 = 1,0$	0,951 (0,078)	0,995 (0,068)	0,937 (0,082)	0,934 (0,087)
$\beta_1 = 0,3$	0,306 (0,158)	0,295 (0,142)	0,308 (0,163)	0,311 (0,172)
$\theta = 0,25$		0,241 (0,006)	0,149 (0,013)	0,264 (0,010)
$\beta_0 = 1,0$	0,926 (0,128)	0,978 (0,108)	0,884 (0,149)	0,881 (0,171)
$\beta_1 = 0,3$	0,317 (0,240)	0,313 (0,210)	0,320 (0,265)	0,320 (0,313)
$\theta = 0,5$		0,499 (0,005)	0,263 (0,059)	0,465 (0,011)

Tabela 7.5: Estimativas para séries geradas pelo modelo GARMA AR(2) sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,757 (0,605)	1,005 (0,004)	1,170 (0,040)	1,164 (0,050)
$\phi_1 = -0,5$		-0,509 (0,002)	-0,548 (0,011)	-0,715 (0,073)
$\phi_2 = 0,25$		0,228 (0,004)	0,078 (0,035)	0,068 (0,057)
$\beta_0 = 1,0$	0,230 (0,806)	-1,353 (2338,701)	-0,038 (1,432)	0,165 (0,921)
$\phi_1 = 0,5$		0,492 (0,012)	0,412 (0,015)	0,364 (0,026)
$\phi_2 = 0,25$		0,219 (0,011)	0,260 (0,008)	0,227 (0,007)

Tabela 7.6: Monte Carlo para GARMA AR(2) com covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,736 (1,042)	1,034 (0,023)	1,222 (0,132)	1,283 (0,386)
$\beta_1 = 0,3$	0,213 (1,030)	0,256 (0,040)	0,193 (0,154)	-0,036 (2,629)
$\phi_1 = -0,5$		-0,510 (0,002)	-0,520 (0,010)	-0,741 (0,099)
$\phi_2 = 0,25$		0,225 (0,005)	0,082 (0,033)	0,059 (0,075)
$\beta_0 = 1,0$	0,118 (3,061)	-1,553 (4892,844)	-0,086 (3,996)	0,079 (2,381)
$\beta_1 = 0,3$	0,402 (4,550)	0,266 (7,713)	0,295 (5,576)	0,400 (3,039)
$\phi_1 = 0,5$		0,478 (0,012)	0,383 (0,021)	0,375 (0,024)
$\phi_2 = 0,25$		0,205 (0,012)	0,251 (0,007)	0,225 (0,007)

Tabela 7.7: Estimativas para séries geradas pelo modelo GARMA MA(2) sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,275	0,999	1,078	1,117
	0,080	0,003	0,009	0,019
$\theta_1 = -0,5$		-0,498	-0,503	-0,754
		0,002	0,009	0,084
$\theta_2 = 0,25$		0,221	0,329	0,387
		0,005	0,009	0,023
$\beta_0 = 1,0$	0,868	0,981	0,777	0,783
	0,035	0,014	0,073	0,069
$\theta_1 = 0,5$		0,506	0,341	0,507
		0,008	0,030	0,008
$\theta_2 = 0,25$		0,196	0,281	0,286
		0,009	0,011	0,001

Tabela 7.8: Estimativas para séries geradas pelo modelo GARMA MA(2) com covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,281 (0,137)	0,986 (0,021)	1,088 (0,032)	1,111 (0,054)
$\beta_1 = 0,3$	0,207 (0,127)	0,313 (0,041)	0,258 (0,047)	0,266 (0,086)
$\theta_1 = -0,5$		-0,505 (0,002)	-0,454 (0,009)	-0,790 (0,105)
$\theta_2 = 0,25$		0,238 (0,007)	0,187 (0,021)	0,337 (0,019)
$\beta_0 = 1,0$	0,875 (0,200)	0,967 (0,152)	0,807 (0,263)	0,805 (0,307)
$\beta_1 = 0,3$	0,314 (0,370)	0,315 (0,298)	0,312 (0,445)	0,308 (0,539)
$\theta_1 = 0,5$		0,494 (0,008)	0,303 (0,043)	0,514 (0,010)
$\theta_2 = 0,25$		0,249 (0,006)	0,170 (0,011)	0,283 (0,014)

Tabela 7.9: Estimativas para séries geradas pelo modelo GLARMA AR(1) sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,151 (0,025)	0,961 (0,003)	0,999 (0,001)	1,007 (0,001)
$\phi = -0,5$		-0,335 (0,029)	-0,495 (0,002)	-0,581 (0,009)
$\beta_0 = 1,0$	1,027 (0,009)	1,102 (0,019)	0,993 (0,009)	0,987 (0,009)
$\phi = 0,25$		0,250 (0,006)	0,243 (0,003)	0,368 (0,021)
$\beta_0 = 1,0$	1,167 (0,057)	1,383 (0,185)	1,012 (0,034)	0,956 (0,035)
$\phi = 0,5$		0,537 (0,009)	0,487 (0,003)	0,626 (0,021)

Tabela 7.10: Estimativas para séries geradas pelo modelo GLARMA AR(1) com covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,153 (0,049)	0,947 (0,016)	1,007 (0,008)	1,014 (0,012)
$\beta_1 = 0,3$	0,303 (0,053)	0,342 (0,028)	0,288 (0,015)	0,288 (0,021)
$\phi = -0,5$		-0,351 (0,024)	-0,501 (0,002)	-0,615 (0,015)
$\beta_0 = 1,0$	1,021 (0,112)	1,119 (0,122)	0,995 (0,110)	0,961 (0,544)
$\beta_1 = 0,3$	0,310 (0,218)	0,264 (0,207)	0,302 (0,214)	0,382 (4,280)
$\phi = 0,25$		0,275 (0,008)	0,241 (0,003)	0,394 (0,028)
$\beta_0 = 1,0$	1,180 (0,338)	1,464 (0,558)	1,121 (0,514)	1,060 (0,607)
$\beta_1 = 0,3$	0,249 (0,620)	0,123 (0,685)	0,120 (0,966)	0,017 (2,326)
$\phi = -0,5$		0,555 (0,011)	0,493 (0,002)	0,664 (0,032)

Tabela 7.11: Estimativas para séries geradas pelo modelo GLARMA MA(1) sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,114 (0,013)	0,952 (0,003)	0,996 (0,000)	1,008 (0,001)
$\theta = -0,5$		-0,291 (0,045)	-0,515 (0,003)	-0,734 (0,062)
$\beta_0 = 1,0$	1,031 (0,009)	1,077 (0,013)	0,998 (0,008)	1,002 (0,008)
$\theta = 0,25$		0,257 (0,006)	0,249 (0,003)	0,388 (0,025)
$\beta_0 = 1,0$	1,131 (0,030)	1,193 (0,050)	0,999 (0,017)	1,040 (0,016)
$\theta = 0,5$		0,474 (0,005)	0,482 (0,004)	0,684 (0,038)

Tabela 7.12: Estimativas para séries geradas pelo modelo GLARMA MA(1) covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,018 (0,007)	0,846 (0,030)	0,894 (0,014)	0,912 (0,013)
$\beta_1 = 0,3$	0,296 (0,014)	0,323 (0,012)	0,294 (0,004)	0,294 (0,011)
$\theta = -0,5$		-0,296 (0,043)	-0,544 (0,006)	-0,785 (0,089)
$\beta_0 = 1,0$	0,930 (0,096)	0,975 (0,088)	0,902 (0,100)	0,911 (0,097)
$\beta_1 = 0,3$	0,298 (0,189)	0,293 (0,181)	0,294 (0,186)	0,285 (0,184)
$\theta = 0,25$		0,267 (0,007)	0,246 (0,003)	0,402 (0,029)
$\beta_0 = 1,0$	1,029 (0,165)	1,080 (0,178)	0,936 (0,264)	1,033 (0,153)
$\beta_1 = 0,3$	0,291 (0,342)	0,298 (0,356)	0,242 (0,508)	0,153 (0,317)
$\theta = 0,5$		0,482 (0,005)	0,483 (0,004)	0,716 (0,050)

Tabela 7.13: Estimativas para séries geradas pelo modelo GLARMA AR(2) sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,438 (0,192)	1,016 (0,000)	1,014 (0,000)	1,051 (0,003)
$\phi_1 = -0,5$		-0,362 (0,019)	-0,496 (0,000)	-0,681 (0,033)
$\phi_2 = 0,25$		0,278 (0,001)	0,227 (0,001)	0,229 (0,001)
$\beta_0 = 1,0$	1,405 (0,164)	-0,787 (3,195)	1,138 (0,019)	1,324 (0,105)
$\phi_1 = 0,5$		0,767 (0,071)	0,484 (0,000)	0,721 (0,049)
$\phi_2 = 0,25$		-0,021 (0,074)	0,246 (0,000)	0,246 (0,000)

Tabela 7.14: Estimativas para séries geradas pelo modelo para GLARMA AR(2) com covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,407 (0,166)	1,020 (0,000)	1,050 (0,002)	1,105 (0,011)
$\beta_1 = 0,3$	0,335 (0,001)	0,303 (0,000)	0,240 (0,004)	0,182 (0,014)
$\phi_1 = -0,5$		-0,374 (0,016)	-0,500 (0,000)	-0,762 (0,069)
$\phi_2 = 0,25$		0,258 (0,000)	0,226 (0,001)	-0,007 (0,066)
$\beta_0 = 1,0$	1,358 (0,128)	-1,638 (6,961)	1,206 (0,042)	1,382 (0,146)
$\beta_1 = 0,3$	0,311 (0,000)	-1,060 (1,851)	0,033 (0,072)	0,144 (0,024)
$\phi_1 = 0,5$		0,828 (0,107)	0,470 (0,001)	0,800 (0,090)
$\phi_2 = 0,25$		-0,084 (0,112)	0,239 (0,000)	-0,071 (0,103)

Tabela 7.15: Estimativas para séries geradas pelo GLARMA MA2 sem covariável

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	1,144 (0,021)	0,969 (0,001)	0,999 (0,000)	1,015 (0,000)
$\theta_1 = -0,5$		-0,333 (0,028)	-0,504 (0,000)	-0,719 (0,048)
$\theta_2 = 0,25$		0,101 (0,022)	0,244 (0,000)	0,272 (0,001)
$\beta_0 = 1,0$	1,162 (0,026)	1,255 (0,065)	1,010 (0,000)	1,064 (0,004)
$\theta_1 = 0,5$		0,598 (0,010)	0,495 (0,000)	0,699 (0,039)
$\theta_2 = 0,25$		0,248 (0,000)	0,247 (0,000)	0,332 (0,007)

Tabela 7.16: Estimativas para séries geradas pelo GLARMA MA2 com covariável tendência

Parâmetro	GLM (EQM)	GARMA (EQM)	GLARMA 0.5 (EQM)	GLARMA 1.0 (EQM)
$\beta_0 = 1,0$	0,985 (0,000)	0,807 (0,037)	0,888 (0,013)	0,921 (0,006)
$\beta_1 = 0,3$	0,386 (0,007)	0,397 (0,009)	0,318 (0,000)	0,287 (0,000)
$\theta_1 = -0,5$		-0,337 (0,026)	-0,510 (0,000)	-0,764 (0,070)
$\theta_2 = 0,25$		0,082 (0,028)	0,223 (0,001)	0,270 (0,000)
$\beta_0 = 1,0$	1,080 (0,006)	1,144 (0,021)	1,037 (0,001)	1,071 (0,005)
$\beta_1 = 0,3$	0,272 (0,001)	0,295 (0,000)	0,122 (0,032)	0,145 (0,024)
$\theta_1 = 0,5$		0,624 (0,015)	0,501 (0,000)	0,742 (0,058)
$\theta_2 = 0,25$		0,258 (0,000)	0,250 (0,000)	0,349 (0,010)

Referências Bibliográficas

- [1] Al-Osh, M. A., Alzaid, A. A. (1987). First order integer valued autoregressive (INAR (1)) process. *Journal of Time Series Analysis*, 8(3), 261-275.
- [2] Box, G. E., Jenkins, G. M. (1976). *Time series analysis*, revised edition. San Francisco.
- [3] Benjamin, M. A., Rigby, R. A., Stasinopoulos, D. M. (2003). Generalized autoregressive moving average models. *Journal of the American Statistical association*, 98(461), 214-223.
- [4] Davis, R. A., Dunsmuir, W. T., Streett, S. B. (2003). Observation driven models for Poisson counts. *Biometrika*, 90(4), 777-790.
- [5] Davis, R. A., Dunsmuir, W. T., Wang, Y. (2000). On autocorrelation in a Poisson regression model. *Biometrika*, 87(3), 491-505.
- [6] Gamerman, D., Santos, T. R. Franco, G. C. (2013). A non-Gaussian family of state-space models with exact marginal likelihood. *Journal of Times Series Analysis*, 34, 625-645.
- [7] Gauss, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium auctore Carolo Friderico Gauss.* sumtibus Frid. Perthes et IH Besser.
- [8] Harvey, A.C., Fernandes, C., 1989. Time series models for count or qualitative observations. *Journal of Business Economic Statistics*, 7(4), 407-417.

- [9] Harvey, A. C., Harvey, A. C. (1993). *Time series models* (Vol. 2). New York. Harvester Wheatsheaf.
- [10] Legendre, A. M. (1805). Nouvelles méthodes pour la détermination des orbites des comètes (No. 1). Paris. F. Didot.
- [11] McCullagh, P., Nelder, J. A. (1989). *Generalized linear models*. London/New York. Chapman and Hall.
- [12] Morettin, P. A., Toloi, C. (2006). *Análise de séries temporais*. São Paulo. Blucher.
- [13] Morettin, P. A., Toloi, C. D. C. (1987). *Previsão de séries temporais*. São Paulo: Atual.
- [14] Nelder, J. A., Wedderburn, R. W. M. (1972). General linearized models. *Journal Royal Statistical Society Series A*, 135, 370-384.
- [15] Silva, A. N., Reisen, V. A. (2011). Associação entre os delitos registrados na RGV uma aplicação do modelo GLARMA Poisson (pp. 37-44). *Revista da Estatística da Universidade Federal de Ouro Preto*, 1(1).
- [16] Tao, Y., Zhong, L., Huang, X., Lu, S. E., Li, Y., Dai, L., ... Huang, W. (2011). Acute mortality effects of carbon monoxide in the Pearl River Delta of China. *Science of the Total Environment*, 410, 34-40.
- [17] Veselý, V., Tonner, J., Hrdlivková, Z., Michálek, J., Kolá, M. (2009). Analysis of PM10 air pollution in Brno based on generalized linear model with strongly rank deficient design matrix. *Environmetrics*, 20(6), 676-698.
- [18] West, M., Harrison, P. J., Migon, H. S. Dynamic Generalized Linear Models and Bayesian Forecasting. *Journal of the American Statistical Association*. Volume 80, Issue 389, 1985.
- [19] Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika*, 75(4), 621-629.