



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA

Ewellyn Karollyne Alves de Alencar Ferreira

O USO DA DISTRIBUIÇÃO LINDLEY GENERALIZADA
NA CONSTRUÇÃO DE UM NOVO PROCESSO
AUTORREGRESSIVO DE PRIMEIRA ORDEM

Recife

2024

Ewellyn Karollyne Alves de Alencar Ferreira

O USO DA DISTRIBUIÇÃO LINDLEY GENERALIZADA NA
CONSTRUÇÃO DE UM NOVO PROCESSO
AUTORREGRESSIVO DE PRIMEIRA ORDEM

Dissertação submetida ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Mestre em Estatística. Área de concentração: Estatística Aplicada.

Orientadora: Prof^a. Dr^a. Maria do Carmo Soares de Lima

Recife

2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Ferreira, Ewellyn Karollyne Alves de Alencar.

O uso da distribuição Lindley generalizada na construção de um novo processo autorregressivo de primeira ordem / Ewellyn Karollyne Alves de Alencar Ferreira. - Recife, 2024.

57f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Ciências Exatas e da Natureza, Programa de Pós-graduação em Estatística, 2024.

Orientação: Maria do Carmo Soares de Lima.

1. Processo autorregressivo de primeira ordem; 2. Série temporal; 3. Simulação de Monte Carlo; 4. Teoria de novas distribuições. I. Lima, Maria do Carmo Soares de. II. Título.

UFPE-Biblioteca Central

Ewellyn Karollyne Alves de Alencar Ferreira

O USO DA DISTRIBUIÇÃO LINDLEY GENERALIZADA NA
CONSTRUÇÃO DE UM NOVO PROCESSO
AUTORREGRESSIVO DE PRIMEIRA ORDEM

Dissertação submetida ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Mestre em Estatística. Área de concentração: Estatística Aplicada.

Aprovada em: 28 de junho de 2024.

BANCA EXAMINADORA

Prof^a. Dr^a. Maria do Carmo Soares de Lima
Universidade Federal de Pernambuco

Prof. Dr. Abraão David Costa do
Nascimento
Universidade Federal de Pernambuco

Prof^a. Dr^a. Agatha Sacramento Rodrigues
Universidade Federal do Espírito Santo

Agradecimentos

Primeiramente, a Deus, por permitir com que eu chegasse até aqui.

À minha família e namorado, por todo apoio, incentivo e compreensão.

À minha orientadora, Profa Maria do Carmo, por toda ajuda e todo ensinamento.
Sem a senhora nada disso seria possível.

Aos professores do departamento, por todo o conhecimento compartilhado.

À CAPES, pelo apoio financeiro.

"Todos somos leigos, o que nos diferencia é o objeto de nossa ignorância."

(Silva, 2018, p. 186) ^[1]

Resumo

A Teoria de Novas distribuições foi amplamente explorada nas últimas décadas com o intuito de se propor modelos mais flexíveis e que cujos ajustes ofertados fossem superiores aos obtidos através de distribuições estabelecidas na literatura. Dessa forma, com o intuito de trazer alguma contribuição para a área de novas distribuições aplicada a séries temporais, o presente trabalho propõe um novo processo autorregressivo de ordem um com marginal Lindley Generalizada. Para o modelo aqui apresentado, é provado que a distribuição dos erros aleatórios é dada por uma mistura de duas distribuições Gama. Algumas medidas estatísticas do processo proposto são estudadas: esperança e variância condicional, distribuição conjunta, densidade espectral e funções de autocorrelação e autocorrelação parcial. Os parâmetros do processo são estimados através dos métodos de estimação Gaussiano e de mínimos quadrados condicionais. Simulações de Monte Carlo são realizadas em cinco diferentes cenários com o objetivo de avaliar o comportamento dos estimadores encontrados. Enfim, uma aplicação a dados reais é realizada, na qual o ajuste do processo autorregressivo aqui proposto é comparado com os ajustes obtidos para outros seis modelos autorregressivos não-Gaussianos de ordem um.

Palavras-chave: Processo autorregressivo de primeira ordem; série temporal; simulação de Monte Carlo; teoria de novas distribuições.

Abstract

The theory of new distributions have been widely used in the last decades with the intention of proposing more flexible models whose fitted values to real data are superior to other ones previously established in the literature. Therefore, in order to bring some contribution to the area of new distributions applied to time series, the present work proposes a new first-order autoregressive process with marginal distribution Generalized Lindley. For the model presented here, it is proven that the distribution of the random errors is given by a mixture of two Gamma distributions. Some statistical measurements of the proposed process are studied: conditional mean, conditional variance, joint distribution, spectral density and the functions of autocorrelation and partial autocorrelation. The parameters of the process are estimated through the Gaussian method and the conditional least squares method. Monte Carlo simulations are implemented in five different scenarios with the aim of evaluating the behavior of the estimators found. Finally, a real data application is carried out, in which the fit of the autoregressive process here proposed is compared with the fit obtained for six other non-Gaussians autoregressive models of order one.

Key-words: First-order autoregressive process; Monte Carlo simulation; theory of new distributions; time series.

Lista de ilustrações

Figura 1 – Função de autocorrelação do processo proposto em três cenários diferentes.	30
Figura 2 – Função de autocorrelação parcial do processo proposto em três cenários diferentes.....	31
Figura 3 – Nível do Lago Huron entre os anos de 1875 e 1972.....	44
Figura 4 – Funções de autocorrelação (a) e autocorrelação parcial (b) dos dados analisados.	44

Lista de tabelas

Tabela 1 – Estimativa e Erro Quadrático Médio dos parâmetros do modelo LGAR sob o método de Mínimos Quadrados Condicionais	39
Tabela 2 – Estimativa e Erro Quadrático Médio dos parâmetros do modelo GLAR sob o método Gaussiano.....	40
Tabela 3 – Critérios de comparação para a aplicação 1	45
Tabela 4 – Estimativas dos parâmetros para a aplicação.....	46

Lista de siglas e abreviaturas

acf	função de autocorrelação.
AIC	critério de informação de Akaike.
AR	autorregressivos.
BIC	critério de informação Bayesiano.
EAR(1)	processo autorregressivo Exponencial de ordem um.
EQM	erro quadrático médio.
fda	função distribuição acumulada.
fdp	função densidade de probabilidade.
GaLAR(1)	processo Gama-Lindley autorregressivo de primeira ordem.
GAR(1)	processo Gama autorregressivo de primeira ordem.
GLAR(1)	processo autorregressivo de ordemum Laplace generalizada.
i.i.d.	independentes e identicamente distribuídas.
IGAR(1)	processo autorregressivo de primeira ordem com distribuição marginal Gaussiana inversa.
LAR(1)	processo autorregressivo de primeira ordem Lindley.
LGAR(1)	processo autorregressivo de primeira ordem com distribuição marginal Lindley Generalizada.
mqc	mínimos quadrados condicionais.
pacf	função de autocorrelação parcial.
PLINAR(1)	processo autorregressivo Poisson-Lindley de ordem um.

Sumário

1	Introdução	12
1.1	Objetivos	18
1.2	Plataforma Computacional	19
1.3	Organização do trabalho	19
2	A distribuição Lindley Generalizada	20
2.1	Introdução	20
2.2	A distribuição Lindley Generalizada	21
3	Novo processo autorregressivo de primeira ordem	24
3.1	Introdução	24
3.2	Construção do processo AR(1) proposto	25
3.3	Propriedades estatísticas do processo	27
3.3.1	Medidas estatísticas condicionais	27
3.3.2	Distribuição conjunta	29
3.4	Funções de autocorrelação e autocorrelação parcial	29
3.5	Densidade espectral	31
4	Estimação do Processo LGAR(1)	33
4.1	Introdução	33
4.2	Método dos Mínimos Quadrados Condicionais	34
4.3	Método Gaussiano	35
4.4	Resultados das Estimações	36
5	Aplicação a dados reais	41
5.1	Introdução	41
5.2	Modelos candidatos	41
5.3	Critérios de comparação	43
5.4	Aplicação a dados reais	43
6	Conclusão	47
6.1	Trabalhos Futuros	47
	Referências	49
	APÊNDICE A Proposição 3	52
	APÊNDICE B Método de estimação dos mínimos quadrados condicionais	55
	APÊNDICE C Transformada de Laplace condicional de X_{t+k} dado X_{t-1}	57

1 Introdução

A distribuição Lindley, proposta por Lindley ^[1] em 1958, é amplamente utilizada na área da Teoria de Novas Distribuições como *baseline*. Isso se dá em grande parte graças a sua simplicidade na representação, uma vez que suas funções densidade e acumulada não são expressas em função de quantidades de difícil computação. Além disso, esta distribuição é indexada por apenas um parâmetro, o que simplifica bastante o processo de estimação e simulação. Seja X uma variável aleatória contínua definida nos reais positivos. Dizemos que X tem distribuição uniparamétrica Lindley se sua função densidade de probabilidade (fdp) e função distribuição acumulada (fda) são dadas, respectivamente, por:

$$f_L(x; \theta) = \frac{\theta^2(1+x)e^{-\theta x}}{1+\theta}$$

e

$$F_L(x; \theta) = 1 - \left(1 + \frac{\theta x}{1+\theta}\right) e^{-\theta x},$$

em que $x > 0$ e $\theta > 0$ é um parâmetro de escala.

Apesar de ter sido proposta em 1958, suas principais características só vieram a ser estudadas em 2007 por Ghitany, Atieh e Nadarajah^[2]. Dentre as propriedades da distribuição Lindley examinadas pelos autores, podem-se destacar função característica, função taxa de falha, momentos, curva de Lorenz, distribuições de somas, produtos e razões e uma fórmula fechada para o estimador de máxima verossimilhança do único parâmetro indexador na distribuição.

Além disso, visto que não se pode gerar dados diretamente da distribuição Lindley através do método da inversa, os autores em questão utilizaram-se do fato que a distribuição Lindley pode ser vista como uma mistura das distribuições Exponencial(θ) e Gama($2, \theta$) para propor uma forma de gerar dados da Lindley(θ). Por fim, os autores mostraram que a distribuição Lindley pode oferecer melhores ajustes que a distribuição Exponencial para dados de tempo de espera através de uma aplicação a dados reais na qual a Lindley apresentou performance superior.

Como mencionado anteriormente, a distribuição Lindley é bastante utilizada como parte constituinte de novas distribuições. Um bom exemplo é a distribuição Binomial Negativa-Lindley proposta por Zamani e Ismail^[3] em 2010 e obtida através da mistura das distribuições Binomial Negativa e Lindley. Os autores apresentaram a forma fechada na nova distribuição, seu momento fatorial e os estimadores dos parâmetros através dos métodos dos momentos e por máxima verossimilhança. Duas aplicações a dados reais foram realizadas em dados de contagem em que a probabilidade em zero é alta, em ambos

os casos a distribuição proposta ofereceu melhor ajuste que as distribuições Poisson e Binomial Negativa.

Com o intuito de se melhor trabalhar com dados de sobrevivência na presença de riscos competitivos, Mazucheli e Achcar^[4] (2011) propuseram um modelo com base na distribuição Lindley. Em tais dados, estuda-se a distribuição do tempo até a falha, levando-se em consideração que os indivíduos podem chegar à falha através de mais de uma causa e o mesmo indivíduo não pode chegar à falha por mais de um fator. No modelo proposto pelos autores, os riscos têm distribuição Lindley e são não correlacionados. Os autores apresentaram a função densidade de probabilidade e a função de log-verossimilhança com suas derivadas parciais de primeira ordem. Uma aplicação a dados reais foi realizada com um conjunto de dados do tempo de vida de pacientes com carcinoma de células escamosas. Os autores compararam o ajuste oferecido pelo modelo de riscos competitivos Lindley com o de outras duas distribuições muito competitivas, a saber, Exponencial e Weibull e observaram que o modelo proposto gerou melhores resultados.

Uma extensão da distribuição Lindley foi proposta por Bakouch^[5] e colaboradores em 2012, a qual foi chamada de Lindley estendida. Os autores desenvolveram a função densidade de probabilidade, função taxa de falha, momentos, estatísticas de ordem e curvas de Bonferroni e Lorenz e estudaram os estimadores de máxima verossimilhança. Ademais, foi executada uma aplicação a dados reais na qual a distribuição Lindley estendida apresentou melhor ajuste que as distribuições Weibull, Gama, Exponencial Exponencializada, Geométrica Exponencial complementar e Weibull modificada de acordo com quatro diferentes critérios de comparação.

Em 2013, Ghitany *et al.*^[6] propuseram a distribuição Power Lindley. Os autores mostraram que a nova distribuição é mais flexível que a Lindley e realizaram uma aplicação a dados reais de resistência à tração de fibras de carbono, na qual a Power Lindley apresentou um melhor ajuste que outras distribuições biparamétricas como Gama e Weibull.

No mesmo ano, Shankar, Sharma e Shanker^[7] propuseram a Lindley biparamétrica. A distribuição foi ajustada a três bancos de dados de tempo de espera e de sobrevivência, os quais já haviam sido previamente utilizados na literatura com a distribuição Lindley. Nos três cenários abordados, a Lindley biparamétrica ofertou melhores ajustes que a Lindley uniparamétrica.

Ainda em 2013, Shanker e Mishra^[8] desenvolveram a distribuição Quasi Lindley. A nova distribuição é indexada por um vetor bidimensional de parâmetros e possui a distribuição Lindley como caso particular. Os momentos, função taxa de falha, função de vida residual média e estimadores dos parâmetros pelos métodos dos momentos e máxima verossimilhança foram estudados para a distribuição proposta. A distribuição Quasi Lindley foi ajustada a dois bancos de dados previamente utilizados na literatura para

demonstrar bondade de ajuste da distribuição Lindley e, em ambos os casos, a distribuição Quasi Lindley apresentou melhores resultados.

Çakmakyapan e Kadilar^[9] propuseram a Kumaraswamy-Lindley. As autoras não realizaram aplicação a dados reais, mas apresentaram as divergências encontradas ao usar duas distintas funções de otimização, *multroot* e *nlm*, disponíveis no software R para estimar os parâmetros da distribuição. As autoras utilizaram a função *multroot* para resolver o sistema de equações não lineares formado pelas derivadas parciais da função de log-verossimilhança da distribuição proposta, mas alguns dos parâmetros não foram bem estimados. Como alternativa, elas aplicaram um algoritmo tipo Newton-Raphson através da função *nlm*, o que levou à estimativas muito melhores que as obtidas pelo primeiro método. Apesar disso, a melhoria obtida com o segundo método de estimação pode ser explicada pelo fato de que a função *nlm* estima apenas um parâmetro por vez e supõe-se que os demais são conhecidos.

Uma outra extensão da Lindley é a Beta-Lindley proposta por Merovci e Sharma^[10] (2014). Os autores em questão fizeram duas aplicações a dados reais para avaliar o ajuste fornecido pelo modelo proposto. O primeiro conjunto de dados utilizado diz respeito ao tempo de remissão de câncer de bexiga com dados não censurados, enquanto o segundo banco representa o tempo de sobrevivência de porquinhos da Índia infectados com vírus. Em ambos os casos, a distribuição proposta ofereceu melhor ajuste que as distribuições Lindley e Beta-Exponencial.

Ainda em 2014, Gui, Zhang e Lu^[11] apresentaram a Lindley Poisson. Os autores realizaram duas aplicações a dados reais. A primeira analisa o intervalo de tempo entre tremores de terra sucessivos na Universidade de Bosphoros. Já a segunda, trata de porquinhos da Índia infectados com vírus. Em ambos os casos, a distribuição proposta forneceu melhor performance que as distribuições Lindley, Power Lindley e Lindley Generalizada.

A distribuição Lindley Exponencial foi proposta por Bhati, Malik e Vaman^[12] em 2015. Tal distribuição é biparamétrica e foi obtida quando os autores aplicaram a função de distribuição acumulada de uma distribuição Exponencial na fórmula da família de distribuições Gama-G, proposta por Zografos e Balakrishnan^[13] (2009). Os autores realizaram duas aplicações a dados reais, uma com relação ao tempo de remissão e a outra com relação ao tempo de espera. Em ambos os casos a distribuição proposta gerou melhores ajustes que as distribuições Power Lindley, Nova Lindley Generalizada e Lindley.

No mesmo ano, Lima^[14] desenvolveu a distribuição Gama Lindley, a qual tem a distribuição Lindley como *baseline*. Para encontrar a nova distribuição, a autora aplicou as funções densidade de probabilidade e acumulada da distribuição Lindley nas respectivas fdp e fda da família de distribuições Gama-G. Além disso, a autora fez duas aplicações a dados reais para testar a bondade de ajuste do modelo proposto. Em ambas as aplicações, uma com tempo entre falhas e outra com imagens de radar, a distribuição Gama Lindley

apresentou melhores ajustes que as distribuições Weibull, Lindley e Exponencial Geométrica complementar.

Em 2016, Popovic e Bakouch^[15] propuseram um processo estacionário autorregressivo de primeira ordem com distribuição marginal Lindley. Os autores estudaram algumas propriedades do novo processo, tais como, densidade espectral, medidas condicionais, *run probabilities* e ergodicidade e estimaram os parâmetros do processo através de mínimos quadrados condicionais, estimação não paramétrica e pelo método Gaussiano. Ademais, duas aplicações a dados reais foram realizadas, nas quais o modelo proposto ofereceu melhores resultados que outros modelos autorregressivos de primeira ordem não Gaussianos.

Barco, Mazucheli e Janeiro^[16] (2016) apresentaram a distribuição *power* Lindley inversa, a qual foi obtida a partir de duas generalizações da distribuição Lindley, mais especificamente, das distribuições *power* Lindley (Ghitany *et al.*, 2013)^[6] e Lindley inversa (Sharma *et al.*, 2015)^[17]. Os autores estudaram o comportamento das funções densidade de probabilidade, taxa de falha e quantil da distribuição proposta. Além disso, um estudo de simulação foi realizado para avaliar a performance dos estimadores de máxima verossimilhança dos parâmetros, os quais se mostraram assintoticamente não viesados. Foram realizadas duas aplicações a conjuntos de dados conhecidos na literatura. Para o primeiro conjunto de dados observado, a distribuição *power* Lindley inversa apresentou melhor ajuste que as distribuições *power* Lindley, Lindley inversa, Lindley generalizada e Lindley. Enquanto que no segundo caso, a distribuição *power* Lindley inversa foi comparada com as distribuições *power* Lindley, Lindley inversa, Lindley e Fréchet transmutada. As distribuições *power* Lindley inversa e Fréchet transmutada apresentaram os melhores resultados, variando qual distribuição oferece melhor ajuste a depender do critério de comparação utilizado.

A distribuição Poisson Lindley foi proposta por Özel e Çakmakyapan^[18] em 2017. Tal distribuição pode ser obtida ao se aplicar a distribuição Lindley como *baseline* na família Poisson-X. Devido a sua grande flexibilidade, pode-se utilizar a distribuição Poisson Lindley para modelar dados em diversas áreas do conhecimento como na economia, engenharia e análise sobrevivência. Os autores ajustaram a distribuição proposta a um conjunto de dados e os resultados mostraram que a Poisson Lindley ofereceu melhores ajustes que os outros 12 modelos com os quais a distribuição em questão foi comparada.

A distribuição Lindley discreta, proposta por Abebe e Shanker^[19] (2018), é análoga a distribuição Lindley contínua e é obtida utilizando o método de discretização de séries infinitas. Além de mostrar que o método dos momentos e de máxima verossimilhança geram os mesmos estimadores, os autores estudaram diversas propriedades da distribuição proposta, dentre as quais pode-se destacar a função geradora de momentos e os coeficientes de variação, assimetria e curtose. Por fim, duas aplicações a dados reais das ciências biológicas foram feitas e em ambos os casos a distribuição Lindley discreta apresentou

melhores ajustes que as distribuições Poisson e Poisson-Lindley.

Em 2019, Pena-Ramirez, Guerra e Cordeiro^[20] propuseram a distribuição Nadarajah-Haghighi Lindley, a qual é obtida através da composição das distribuições Nadarajah-Haghighi e Lindley, é indexada por um vetor de três parâmetros e possui as distribuições Lindley, Nadarajah-Haghighi e Exponencial como casos especiais. Os autores realizaram duas aplicações e em ambos os casos a distribuição proposta ofereceu melhores ajustes que as distribuições Exponencial Nadarajah-Haghighi, Exponencial Weibull, Weibull, Nadarajah-Haghighi, Lindley e Exponencial.

Além disso, unir a área de novas distribuições com a de séries temporais possibilitou aos pesquisadores o desenvolvimento de novos processos autorregressivos muito importantes, dentre os quais podemos destacar as contribuições a seguir.

Em 1980, Gaver e Lewis^[21] propuseram um processo autorregressivo Exponencial de ordem um (EAR(1)), o qual possui distribuição marginal Exponencial(λ). Seja a sequência de inovações $\{\varepsilon_t\}$ composta por variáveis independentes e identicamente distribuídas, então o processo EAR(1) é dado por

$$X_t = \rho X_{t-1} + \varepsilon_t,$$

em que $\rho \in [0, 1)$ e $t = 0, \pm 1, \pm 2, \dots$.

O processo em questão é indexado por dois parâmetros, um advindo da distribuição marginal (λ) e o outro da própria estrutura do modelo (ρ). Além disso, ele possui algumas propriedades muito atraentes, tais como a facilidade na realização de simulações e no tratamento analítico e o fato de levar a uma extensão cuja distribuição marginal segue uma exponencial e cuja correlação tem a estrutura de um processo autorregressivo com médias móveis.

Sua equação diferença é dada por

$$X_t = \rho X_{t-1} + I_t E_t,$$

em que I_t é uma sequência i.i.d. de variáveis aleatórias com distribuição Bernoulli($1 - \rho$) e $\{E_t\}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com distribuição Exponencial(λ).

No mesmo trabalho, os autores apresentaram o processo Gama autorregressivo de primeira ordem (GAR(1))^[21]. O modelo em questão possui distribuição marginal Gama(λ, k) e é indexado por uma tríade de parâmetros, um proveniente da estrutura autorregressiva (ρ) e os demais da distribuição marginal (λ, k). Assim como o modelo anterior, o processo GAR(1) é obtido através da Equação (3.1) para o caso em que $\rho \in [0, 1)$ e $x \geq 0$.

Abraham e Balakrishna^[22] (1999) propuseram o processo autorregressivo de primeira ordem com distribuição marginal Gaussiana inversa (IGAR(1)). Seja x_0 uma variável

aleatória com distribuição Gaussiana inversa independente de ε_i , em que $\{\varepsilon_t\}$ é uma sequência de variáveis aleatórias i.i.d. não negativas. O processo IGAR(1) é obtido através da equação

$$X_t = \rho X_{t-1} + \varepsilon_t,$$

em que $\rho \in [0, 1)$ e $t = 1, 2, \dots$. S

O processo autorregressivo de ordemum Laplace generalizada (GLAR(1)) foi proposto por Jose e Thomas^[23] em 2011. O modelo em questão segue a estrutura aditiva usual dos processos autorregressivos de primeira ordem dada na Equação (3.1) condicionada ao caso em que $\{X_t, t = 0, \pm 1, \pm 2, \dots\}$ e tem distribuição marginal estacionária Laplace generalizada com vetor paramétrico $(\theta, k, \sigma, \tau)$.

Anos depois, em 2016, Popovic e Bakouch^[15] propuseram um novo processo autorregressivo de primeira ordem Lindley (LAR(1)). O processo em questão tem distribuição marginal Lindley e é indexado por um vetor biparamétrico, a saber, ρ advindo da estrutura autorregressiva e λ da distribuição marginal. Os autores abordaram diversas propriedades do novo modelo, tais como, densidade espectral, medidas condicionais e *run probabilities*. Os parâmetros do modelo foram estimados através dos métodos de mínimos quadrados condicionais, Gaussiano, e estatística não paramétrica. Além disso, duas aplicações a dados reais foram realizadas e o processo proposto apresentou resultados superiores a outros 6 modelos autorregressivos de primeira ordem Não Gaussianos conhecidos na literatura.

O processo autorregressivo Poisson-Lindley de ordem um (PLINAR(1)) foi proposto por Mohammadpour, Bakouch e Shirozhan^[24] em 2018. O processo então apresentado consiste de um modelo estacionário autorregressivo de primeira ordem de valores inteiros com distribuição marginal discreta Poisson-Lindley. Considere o processo autorregressivo de primeira ordem de valores inteiros INAR(1) definido pela equação

$$X_t = \rho \circ X_{t-1} + \varepsilon_t, \tag{1.1}$$

em que $t \geq 1$, $\rho \circ$ é o operador *binomial thinning operator* e $\{\varepsilon_t\}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas. Dessa forma, o modelo PLINAR(1) é definido pela Equação (1.1) em que se assume que $\{X_t\}$ tem distribuição Poisson-Lindley de parâmetro θ .

Além disso, também podemos destacar o processo Gama-Lindley autorregressivo de primeira ordem (GaLAR(1)) proposto por Mello, Lima e Nascimento^[25] (2021). O processo em questão é obtido através da Equação (3.1), possui distribuição marginal Gamma-Lindley e é indexado por três parâmetros, um proveniente da estrutura autorregressiva (ρ) e os demais da distribuição marginal (λ, β). Os autores estimaram os parâmetros desconhecidos do processo através do método Gaussiano e de mínimos quadrados condicionais, além de realizarem uma aplicação a dados hidrológicos, na qual o modelo proposto apresentou resultados superiores a outros 6 processos autorregressivos de primeira ordem.

Dessa forma, é notável a importância e a diversidade das possíveis aplicações não só da distribuição Lindley como principalmente das distribuições que a têm como *baseline* e/ou caso particular. Assim, este trabalho de dissertação dará ênfase ao uso da distribuição Lindley Generalizada, proposta por Zakerzadeh e Dolati^[26] (2009) e que será discutida em mais detalhes no capítulo a seguir, como marginal de um processo autorregressivo de ordem 1.

1.1 Objetivos

Ao se propor uma nova distribuição, tem-se como foco principal encontrar um modelo mais flexível e que apresente ajustes superiores a determinados tipos de dados reais quando comparado a outras distribuições conhecidas na literatura.

Nesse sentido, o presente trabalho busca unir a Teoria de Novas Distribuições à área de Séries Temporais ao propor um novo processo autorregressivo de primeira ordem cuja marginal segue uma distribuição Lindley Generalizada^[26] ao tomar como base trabalhos como os de Popovic e Bakouch^[15] (2016) e Mello, Lima e Nascimento^[25] (2021) os quais propuseram processos autorregressivos de ordem um com marginais Lindley e Gama-Lindley, respectivamente. A motivação para a escolha de tal *baseline* encontra-se no Capítulo 2, o qual é dedicado à distribuição Lindley Generalizada^[26].

Os objetivos específicos são destacados a seguir:

- Propor um novo processo autorregressivo de primeira ordem com marginal Lindley Generalizada;
- Desenvolver a distribuição dos erros do modelo proposto levando em consideração a marginal mencionada previamente;
- Estudar algumas medidas estatísticas condicionais do processo, a saber, esperança, variância e transformada de Laplace;
- Avaliar as funções de autocorrelação e autocorrelação parcial com o intuito de confirmar a ordem do processo;
- Estimar os parâmetros do modelo através dos métodos Gaussiano e de Mínimos Quadrados Condicionais;
- Desenvolver simulações de Monte Carlo a fim de avaliar o desempenho dos estimadores Gaussianos e de Mínimos Quadrados Condicionais e
- Através de aplicações a dados reais, comparar os ajustes ofertados pelo modelo aqui proposto com os ajustes obtidos para outros modelos autorregressivos de ordem um encontrados na literatura.

1.2 Plataforma Computacional

O presente trabalho foi redigido utilizando o *software* livre \LaTeX através do editor de texto *online* Overleaf. O *software* em questão foi desenvolvido pelo cientista da computação Leslie B. Lamport em 1983 e desde então tem sido amplamente empregado na produção de literatura científica.

O *software* R (Ross Ihaka e Robert Gentleman), através da IDE RStudio, foi utilizado na construção de gráficos, no desenvolvimento de simulações e nas aplicações a dados reais computadas. As principais bibliotecas empregadas foram: *base*, *graphics* e *stats*.

Dentre todas as funções computadas, vale ressaltar as funções `ts()`, usada para transformar um *dataframe* em um objeto do tipo série temporal; `acf()` e `pacf()`, usadas para desenvolver os gráficos da função de autocorrelação (`acf`) e função de autocorrelação parcial (`pacf`), respectivamente e a função `optim()`, empregada nas simulações apresentadas no Capítulo 4.

1.3 Organização do trabalho

A presente dissertação está organizada da seguinte maneira. No capítulo 2, a distribuição Lindley Generalizada utilizada como marginal do processo proposto é estudada. No capítulo 3, o processo autorregressivo de primeira ordem com distribuição marginal Lindley Generalizada (LGAR(1)) é apresentado e algumas de suas propriedades estatísticas são investigadas, a saber: medidas estatísticas condicionais, distribuição conjunta, funções de autocorrelação e autocorrelação parcial e densidade espectral. No capítulo 4, os parâmetros do processo proposto são estimados através dos métodos Gaussiano e de mínimos quadrados condicionais e simulações de Monte Carlo são desenvolvidas para avaliar a performance dos estimadores obtidos. No capítulo 5, três aplicações a dados reais são realizadas e a performance do modelo proposto é comparada à de outros 7 processos autorregressivos de primeira ordem. Por fim, no capítulo 6, as conclusões finais são apresentadas.

Note que, ao longo do trabalho algumas definições e resultados são destacados. As contribuições da autora são apresentadas majoritariamente através de proposições, enquanto que as previamente existentes na literatura são anunciadas em ambiente de definição.

2 A distribuição Lindley Generalizada

Resumo

Neste capítulo é apresentada a distribuição Lindley Generalizada, proposta por Zakerzadeh e Dolati^[26] (2009). Os autores estudaram diversas propriedades da nova distribuição, além de avaliar a flexibilidade do modelo através de duas aplicações a dados reais. O primeiro exemplo numérico abordou o tempo, em minutos, até a falha de uma amostra de componentes eletrônicos, enquanto o segundo conjunto de dados retratou o número de ciclos até a falha de diferentes tipos de fios de igual tamanho, testados sob determinado nível de tensão. Algumas das propriedades mencionadas são abordadas ao longo deste e dos demais capítulos.

Palavras-chave: Distribuição Lindley Generalizada. Propriedades estatísticas. Transformada de Laplace.

2.1 Introdução

Extensões da distribuição Lindley foram amplamente desenvolvidas e estudadas nas duas últimas décadas. Tais distribuições se mostraram muito eficientes para modelar dados das mais diversas áreas, como biologia, sociologia e engenharia.

Por outro lado, pouco progresso foi feito na construção de novos modelos autorregressivos cuja distribuição marginal seja uma extensão da Lindley. Apesar disso, os escassos trabalhos com o intuito mencionado mostraram que tais processos apresentam ajustes superiores a outros modelos autorregressivos (AR) estabelecidos na literatura.

Uma outra distribuição momentaneamente preterida pela academia é a Lindley Generalizada proposta por Zakerzadeh e Dolati^[26] em 2009. Dentre tantas outras de igual nomenclatura, pouco se encontra sobre a distribuição em questão o que vai de encontro com a sua alta flexibilidade e com o fato de que ela se mostrou competitiva com modelos muito populares, a saber, Gama, Weibull e Lognormal.

Desse modo, a distribuição Lindley Generalizada é estudada na seção a seguir, com o intuito de reparar as duas lacunas mencionadas acima e de se construir a base necessária para se utilizar tal distribuição como marginal de um novo processo autorregressivo.

2.2 A distribuição Lindley Generalizada

Diversos geradores de novas distribuições foram propostos nas últimas décadas, tais como a família Exponencializada-G, que foi um dos primeiros geradores a ser proposto (Lehmann, 1953 apud Cordeiro, Ortega e da Cunha^[27], 2013); a família Kumaraswamy-G proposta por Kumaraswamy^[28] em 1980, a qual traz ajustes superiores quando comparada a outras famílias na literatura; o proposto por Marshall e Olkin^[29] em 1997; a família Beta-G introduzida em 2002 por Eugene, Lee e Famoye^[30], que apesar de fornecer melhores ajustes que muitos dos demais geradores conhecidos na literatura, tem-se certa dificuldade na obtenção de resultados inerentes à simulação e estimação dos parâmetros tendo em vista a complexidade existente em sua estrutura; a família Gama-G (Zografos e Balakrishnan^[13], 2009) o gerador de famílias T-X (Alzaatreh, Lee e Famoye^[31], 2013) e o clássico método de misturas. Este último gera uma nova função de distribuição acumulada, $F(x)$, a partir de outras n fdas, digam-se $G_1(x), G_2(x), \dots, G_n(x)$, previamente definidas. Para tanto, a cada distribuição individual, denominada componente da mistura, é associado um peso ω_i e a distribuição de mistura é dada por:

$$F(x) = \sum_{i=1}^n \omega_i G_i(x),$$

em que $\sum_{i=1}^n \omega_i = 1$.

Sejam $g_1(x), g_2(x), \dots, g_n(x)$ as respectivas funções densidade de probabilidade associadas a cada componente da mistura; a densidade da mistura é obtida através de:

$$f(x) = \sum_{i=1}^n \omega_i g_i(x).$$

Conseqüentemente, se cada componente da mistura representar uma variável aleatória contínua, então a variável resultante também será contínua. Note que, apesar de não se abordar neste trabalho, há a possibilidade do conjunto dos componentes da mistura ser infinito enumerável, ou seja, $n = \infty$, ou infinito não enumerável.

O método em questão gera modelos com grande capacidade de se ajustar a conjuntos de dados heterogêneos e cujas populações possuem distribuições mais complexas. O que faz com que o método de misturas tenha uma adesão cada vez maior e com que as distribuições geradas através dele possam ser aplicadas nas mais diversas áreas do conhecimento.

Nesse sentido, Zakerzadeh e Dolati^[26] (2009) propuseram a variável aleatória X com densidade Lindley Generalizada como sendo uma mistura de duas distribuições Gama. Sejam $X_1 \sim \text{Gama}(\alpha, \theta)$ e $X_2 \sim \text{Gama}(\alpha + 1, \theta)$; assumamos que $X = X_1$ com probabilidade $\theta/(\theta + \gamma)$ e $X = X_2$ com probabilidade $\gamma/(\theta + \gamma)$, para $\gamma \geq 0$. Dessa forma, a função

densidade de probabilidade de X é dada por:

$$\begin{aligned} f_{LG}(x; \alpha, \theta, \gamma) &= \left[\frac{\theta^2(\theta x)^{\alpha-1}e^{-\theta x}}{(\theta + \gamma)\Gamma(\alpha)} \right] + \left[\frac{\gamma\theta(\theta x)^\alpha e^{-\theta x}}{(\theta + \gamma)\Gamma(\alpha + 1)} \right] \\ &= \frac{\theta^2(\theta x)^{\alpha-1}(\alpha + \gamma x)e^{-\theta x}}{(\theta + \gamma)\Gamma(\alpha + 1)}, \end{aligned} \quad (2.1)$$

em que $\alpha, \theta, x > 0$ e $\gamma \geq 0$. Esta distribuição possui como casos particulares as distribuições Lindley, quando $\alpha = \gamma = 1$; Gama com parâmetros α e θ , quando $\gamma = 0$ e Exponencial com parâmetro θ , quando $(\alpha, \gamma) = (1, 0)$. Ao longo do texto, a distribuição Lindley Generalizada será representada por LG ou $LG(\alpha, \theta, \gamma)$.

Os autores supracitados apresentaram uma expressão para a função geradora de momentos, $M(t)$, da distribuição LG, a qual é dada por:

$$M(t) = E(e^{tX}) = \left(\frac{\theta - t + \gamma}{\theta + \gamma} \right) \left(\frac{\theta}{\theta - t} \right)^{\alpha+1}. \quad (2.2)$$

Dessa forma, partindo da Equação (2.2), torna-se trivial a obtenção das expressões da esperança e variância da LG, as quais são dadas, respectivamente, por:

$$E(X) = M'(0) = \frac{\theta\alpha + \alpha\gamma + \gamma}{\theta(\theta + \gamma)}$$

e

$$\begin{aligned} V(X) &= E(X^2) - [E(X)]^2 = M''(0) - [M'(0)]^2 \\ &= \frac{\theta(2\alpha\gamma + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)}{\theta^2(\theta + \gamma)^2}. \end{aligned} \quad (2.3)$$

É importante salientar que há uma outra distribuição de mesma nomenclatura proposta por Nadarajah, Bakouch e Tahmasbi^[32] em 2011. Os autores em questão propuseram uma nova distribuição biparamétrica como alternativa às distribuições Gama, Weibull e Lognormal na modelagem de dados de tempo de vida e chamaram tal distribuição de Lindley Generalizada. O presente trabalho não irá abordar esta distribuição, mas sim a proposta por Zakerzadeh e Dolati.

Ademais, com o intuito de viabilizar cálculos futuros, obteve-se a transformada de Laplace de uma variável aleatória com distribuição Lindley Generalizada. A transformada de Laplace é uma transformada integral cujo nome foi dado em homenagem ao matemático Pierre-Simon Laplace, o qual empregou uma transformada similar em seus estudos. Ela foi desenvolvida de forma mais profunda entre os séculos XIX e XX por Matyáš Lerch, Oliver Heaviside e Thomas John I'Anson Bromwich.

A transformada de Laplace apresenta uma vasta aplicabilidade em diversas áreas como Matemática, Estatística, Física e Engenharias, pois através dela é possível diminuir

a complexidade do sistema que se está analisando. Dessa forma, seja X uma variável aleatória e ϕ_X a transformada de Laplace referente a X . Tem-se que a representação da forma pura da transformada de Laplace é dada por:

$$\begin{aligned}\phi_X(s) &= \int_{-\infty}^{\infty} e^{-sx} f(x) dx \\ &= \mathbb{E}[e^{-sx}].\end{aligned}\tag{2.4}$$

Proposição 1. *A transformada de Laplace de uma variável aleatória X com distribuição Lindley Generalizada é dada por:*

$$\phi_X(s) = \frac{\theta^{\alpha+1}(\theta + \gamma + s)}{(\theta + \gamma)(\theta + s)^{\alpha+1}}.\tag{2.5}$$

Demonstração. Aplicando a fdp da distribuição Lindley Generalizada presente em (2.1) na equação (2.4), tem-se que

$$\begin{aligned}\phi_X(s) &= \mathbb{E}[e^{-sx}] = \int_{-\infty}^{\infty} e^{-sx} \frac{\theta^2(\theta x)^{\alpha-1}(\alpha + \gamma x)e^{-\theta x}}{(\theta + \gamma)\Gamma(\alpha + 1)} dx \\ &= \frac{\theta^{\alpha+1}}{(\theta + \gamma)\Gamma(\alpha + 1)} \int_0^{\infty} \alpha x^{\alpha-1} e^{-(\theta+s)x} + \gamma x^{\alpha} e^{-(\theta+s)x} dx \\ &= \frac{\theta^{\alpha+1}}{(\theta + \gamma)\Gamma(\alpha + 1)} \left[\frac{\alpha\Gamma(\alpha)}{(\theta + s)^{\alpha}} + \frac{\gamma\Gamma(\alpha + 1)}{(\theta + s)^{\alpha+1}} \right] \\ &= \frac{\theta^{\alpha+1}(\theta + s + \gamma)}{(\theta + \gamma)(\theta + s)^{\alpha+1}}.\end{aligned}$$

□

Com base nas definições apresentadas, será desenvolvido um processo autorregressivo de primeira ordem com distribuição marginal Lindley Generalizada^[26] no capítulo a seguir.

3 Novo processo autorregressivo de primeira ordem

Resumo

Neste capítulo, um novo processo autorregressivo de primeira ordem é construído utilizando a Lindley Generalizada como distribuição das variáveis de saída. Para tal processo, são estudadas suas medidas estatísticas condicionais, distribuição conjunta e funções de autocorrelação, autocorrelação parcial e densidade espectral.

Palavras-chave: Distribuição Lindley Generalizada. Processo autorregressivo. Propriedades estatísticas.

3.1 Introdução

Uma série temporal pode ser descrita como um conjunto de observações referentes a uma determinada variável que foram coletadas e ordenadas ao longo do tempo. O tempo de coleta dos dados pode evoluir de forma discreta ou contínua. Nesse sentido, o tempo é visto como uma variável independente da variável de interesse e, tendo em vista que as observações são coletadas em momentos igualmente espaçados, há uma provável correlação entre as realizações da variável aleatória observada.

Tal correlação viola muitas das suposições estatísticas necessárias para a utilização dos métodos tradicionais empregados na análise de dados. Por essa razão, é preciso fazer uso de técnicas desenvolvidas especificamente para tratar amostras deste formato. Alguns dos mecanismos mais utilizados são representação gráfica das observações e das funções amostrais de autocorrelação e autocorrelação parcial, análise da presença de possíveis tendências e sazonalidades, além de averiguar o mecanismo gerador dos dados com o intuito de efetuar previsões a partir da amostra.

Dois conceitos muito importantes ao lidar com séries temporais são os de estacionaridade fraca, ou simplesmente estacionaridade, e estacionaridade forte, também chamada de estacionaridade estrita, os quais são apresentados a seguir.

Definição 1. *Uma série temporal $\{X(t); t \in \mathbb{T}\}$ é dita ser fracamente estacionária se:*

1. $\mathbb{E}[X(t)] = \mu(t) = \mu, \forall t \in \mathbb{T},$
2. $\mathbb{E}[X^2(t)] < \infty, \forall t \in \mathbb{T},$

$$3. \gamma_X(t_1, t_2) = \gamma_X(t_1 + \tau, t_2 + \tau), \forall t_1, t_2, \tau \in \mathbb{T},$$

em que $\gamma_X(t_1, t_2)$ é a função de autocovariância de X nos instantes t_1 e t_2 .

Definição 2. Uma série temporal $\{X(t); t \in \mathbb{T}\}$ é dita ser fortemente (estritamente) estacionária se as funções de distribuição conjuntas de $\{Y_{t_1}, \dots, Y_{t_n}\}$ e $\{Y_{t_1+h}, \dots, Y_{t_n+h}\}$ são idênticas $\forall n \in \mathbb{N}^+$ e $\forall t_1, \dots, t_n, t_1+h, \dots, t_n+h \in \mathbb{T}$. Dessa forma, todos os momentos finitos são constantes ao longo do tempo.

Ao fazer uma análise de séries temporais, é interessante que a série temporal estudada seja estacionária.

As seções a seguir buscam construir um processo autorregressivo de primeira ordem com a Lindley Generalizada como distribuição marginal.

3.2 Construção do processo AR(1) proposto

Sejam ε_t e X_t respectivamente as variáveis de entrada e saída do modelo no instante t . Um processo autorregressivo de ordem 1 é representado por $X_t \sim \text{AR}(1)$ e segue a estrutura dada a seguir

$$X_t = \rho X_{t-1} + \varepsilon_t, \quad (3.1)$$

em que $\rho, X_t, \varepsilon_t \in \mathbb{R}$, ρ é o parâmetro autorregressivo e $\{\varepsilon_t\}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) e independentes de $X_{t-h} \forall h \geq 1$. Note que, de (3.1) tem-se que $x_t = \rho^2 x_{t-2} + \rho \varepsilon_{t-2} + \varepsilon_t$.

Assim, ao substituir recursivamente a variável de saída presente no termo à direita da igualdade, obtém-se uma expressão para o t -ésimo instante de X que depende apenas do seu valor inicial (x_{-1}), do coeficiente autorregressivo e dos valores de ε_t em todos os instantes de tempo até t . Tal expressão é dada por:

$$x_t = \rho^{t+1} x_{-1} + \rho^t \varepsilon_0 + \rho^{t-1} \varepsilon_1 + \dots + \varepsilon_t.$$

A equação acima é preferível à (3.1), visto que depende da variável de saída apenas num valor supostamente conhecido e os termos desconhecidos (ρ e $\{\varepsilon_t\}$) possuem propriedades mais tratáveis que a variável de saída.

Ademais, para fins futuros, é importante ressaltar que o processo apresentado em (3.1) é estacionário se e só se $|\rho| < 1$ devido a natureza autorregressiva do processo. Uma outra informação que será necessária é a distribuição dos erros ($\{\varepsilon_t\}$). Para tanto, a

transformada de Laplace da distribuição de X_t será analisada. Dessa forma,

$$\begin{aligned}\phi_{X_t}(s) &= \mathbb{E}[\exp(-s\rho X_t)] \\ &\stackrel{(1)}{=} \mathbb{E}[\exp(-s\rho X_{t-1})] \mathbb{E}[\exp(-s\varepsilon_t)] \\ \therefore \phi_{X_t}(s) &\stackrel{(2)}{=} \phi_{X_{t-1}}(\rho s)\phi_{\varepsilon_t}(s).\end{aligned}\tag{3.2}$$

em que (1) e (2) representam, respectivamente, a utilização da independência das variáveis aleatórias e a definição da transformada de Laplace.

A transformada de Laplace da sequência de erros aleatórios é apresentada a seguir.

Proposição 2. *A transformada de Laplace do erro é dada por:*

$$\phi_{\varepsilon_t}(s) = \left(\frac{\theta + \gamma + s}{\theta + \gamma + \rho s}\right) \left(\frac{\theta + \rho s}{\theta + s}\right)^{\alpha+1}.\tag{3.3}$$

Demonstração. Como consequência direta de (3.2), tem-se que a transformada de Laplace dos erros ($\{\varepsilon_t\}$) pode ser obtida através de

$$\begin{aligned}\phi_{\varepsilon_t}(s) &= \frac{\phi_{X_t}(s)}{\phi_{X_{t-1}}(\rho s)} \\ &= \frac{\mathbb{E}[\exp(-s\rho X_t)]}{\mathbb{E}[\exp(-s\rho X_{t-1})]}.\end{aligned}$$

Pela estacionaridade do processo,

$$\begin{aligned}\phi_{\varepsilon_t}(s) &= \frac{\mathbb{E}[\exp(-s\rho X)]}{\mathbb{E}[\exp(-s\rho X)]} \\ &= \frac{\phi_X(s)}{\phi_X(\rho s)}, \\ &= \frac{(\theta + \gamma + s)\theta^{\alpha+1}}{(\theta + \gamma)(\theta + s)^{\alpha+1}} \times \frac{(\theta + \gamma)(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)\theta^{\alpha+1}} \\ &= \left(\frac{\theta + \gamma + s}{\theta + \gamma + \rho s}\right) \left(\frac{\theta + \rho s}{\theta + s}\right)^{\alpha+1}.\end{aligned}$$

□

A seguir, será apresentada a função densidade de probabilidade dos erros.

Proposição 3. *A sequência de erros aleatórios $\{\varepsilon_t\}$ possui função densidade de probabilidade dada por $f_{\varepsilon_t}(x)$, a qual é uma função absolutamente contínua obtida através da mistura de três distribuições conhecidas, a saber, Exponencial $\left(\frac{\theta + \gamma}{\rho}\right)$, Exponencial (θ) e Gama $(\alpha + 1, \theta)$. Dessa forma, a densidade dos erros aleatórios é dada por*

$$f_{\varepsilon_t}(x) = A \frac{(\theta + \gamma)e^{-\frac{(\theta + \gamma)x}{\rho}}}{\rho} + B\theta e^{-\theta x} + C \frac{\theta^{\alpha+1} x^{\alpha} e^{-\theta x}}{\Gamma(\alpha + 1)},$$

em que

$$A = \frac{(-1)^{\alpha+1} \gamma^{\alpha+1} \rho^\alpha (\rho - 1)}{(\theta \rho - \theta - \gamma)^{\alpha+1}},$$

$$C = \frac{\gamma(1 - \rho)^{\alpha+1}}{\theta + \gamma - \theta \rho},$$

e

$$B = 1 - A - C$$

A demonstração da Proposição 3 encontra-se no Apêndice A e o processo em questão será denotado por $X_t \sim \text{LGAR}(1)$.

3.3 Propriedades estatísticas do processo

Nesta seção, algumas propriedades estatísticas do processo $\text{LGAR}(1)$ são estudadas, a saber: esperança, variância e transformada de Laplace condicionais, distribuição conjunta, *run probabilities*, funções de autocorrelação, autocovariância e autocorrelação parcial, densidade espectral e a previsão do processo.

3.3.1 Medidas estatísticas condicionais

Com o intuito de obter expressões para a esperança, variância e transformada de Laplace condicionais de $X_{t+k}|X_{t-1} = x_{t-1}$, pode-se reescrever X_{t+k} através de substituição recursiva como

$$X_{t+k} = \rho^{k+1} X_{t-1} + \rho^k \varepsilon_t + \rho^{k-1} \varepsilon_{t+1} + \cdots + \rho \varepsilon_{t+k-1} + \varepsilon_{t-k}.$$

Aplicando o valor esperado na equação acima, obtem-se

$$\mathbb{E}(X_{t+k}|X_{t-1} = x_{t-1}) = \rho^{k+1} x_{t-1} + \rho^k \mathbb{E}(\varepsilon_t) + \rho^{k-1} \mathbb{E}(\varepsilon_{t+1}) + \cdots + \rho \mathbb{E}(\varepsilon_{t+k-1}) + \mathbb{E}(\varepsilon_{t-k}).$$

Mas, como $\{\varepsilon_t\}$ é uma sequência de variáveis aleatórias i.i.d.,

$$\begin{aligned} \mathbb{E}(X_{t+k}|X_{t-1} = x_{t-1}) &= \rho^{k+1} x_{t-1} + \mathbb{E}(\varepsilon_t) \sum_{i=0}^k \rho^i \\ &= \rho^{k+1} x_{t-1} + \mathbb{E}(\varepsilon_t) \left[\frac{\rho^{k+1} - 1}{\rho - 1} \right]. \end{aligned}$$

Assim, pela estacionariedade do processo,

$$\begin{aligned} \mathbb{E}(X_{t+k}|X_{t-1} = x_{t-1}) &= \rho^{k+1} x_{t-1} + \mathbb{E}(X_t)(1 - \rho) \left[\frac{\rho^{k+1} - 1}{\rho - 1} \right] \\ &= \rho^{k+1} x_{t-1} + \frac{(1 - \rho^{k+1})(\theta \alpha + \alpha \gamma + \gamma)}{\theta(\theta + \gamma)}. \end{aligned} \quad (3.4)$$

Note que, como $\rho \in [0, 1)$, então $\lim_{k \rightarrow \infty} \rho^{k+1} = 0$. Portanto, quando $k \rightarrow \infty$ tem-se que

$$\mathbb{E}(X_{t+k}|X_{t-1} = x_{t-1}) = \frac{\theta\alpha + \alpha\gamma + \gamma}{\theta(\theta + \gamma)} = \mathbb{E}(X_t).$$

Dessa forma, quando k tende a valores muito altos a esperança condicional do processo $k + 1$ passos à frente regressa à esperança da distribuição marginal. Ou seja, conforme k aumenta, a esperança condicional depende cada vez menos do valor observado $k + 1$ passos anteriores.

Analogamente ao que foi feito para a esperança, a variância condicional de $X_{t+k}|X_{t-1} = x_{t-1}$ é calculada como mostra a seguir

$$\begin{aligned} \text{Var}(X_{t+k}|X_{t-1} = x_{t-1}) &= \text{Var}\left(\rho^{k+1}x_{t-1} + \rho^k\varepsilon_t + \rho^{k-1}\varepsilon_{t+1} + \cdots + \rho\varepsilon_{t+k-1} + \varepsilon_{t-k}\right) \\ &\stackrel{(1)}{=} \rho^{2k}\text{Var}(\varepsilon_t) + \rho^{2(k-1)}\text{Var}(\varepsilon_{t+1}) + \cdots + \rho^2\text{Var}(\varepsilon_{t+k-1}) + \text{Var}(\varepsilon_{t-k}) \\ &\stackrel{(2)}{=} \text{Var}(\varepsilon_t) \sum_{i=0}^k \rho^{2i} \\ &= \text{Var}(\varepsilon_t) \left(\frac{\rho^{2k+2} - 1}{\rho^2 - 1}\right) \\ &\stackrel{(3)}{=} \frac{(1 - \rho^{2k+2})[\theta(2\alpha\gamma + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)]}{\theta^2(\theta + \gamma)^2}, \end{aligned} \quad (3.5)$$

em que (1), (2) e (3) indicam, respectivamente, independência dos erros, os ε são identicamente distribuídos e a estacionariedade do processo juntamente com a independência entre X_{t-1} e ε_t o que implica em $\text{Var}(\varepsilon_t) = (1 - \rho^2)\text{Var}(X_t)$.

Além disso, note que quando $k \rightarrow \infty$ temos que

$$\begin{aligned} \text{Var}(X_{t+k}|X_{t-1} = x_{t-1}) &= \frac{\theta(2\alpha\gamma + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)}{\theta^2(\theta + \gamma)^2} \\ &= \text{Var}(X_t). \end{aligned}$$

Ou seja, quando k tende a valores muito altos a variância condicional do processo $k + 1$ passos a frente retorna à variância da distribuição marginal Lindley Generalizada.

Além disso, é possível mostrar que

$$\begin{aligned} \mathbb{E}(X_{t-1}|X_{t+k} = x) &= \frac{x}{\rho^{k+1}} - \frac{(\theta\alpha + \alpha\gamma + \gamma)(\rho^{-(k+1)} - 1)}{\theta(\theta + \gamma)} \\ &\neq \mathbb{E}(X_{t+k}|X_{t-1} = x) \end{aligned}$$

e

$$\begin{aligned} \text{Var}(X_{t-1}|X_{t+k} = x) &= \frac{[\theta(2\alpha\gamma + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)](\rho^{-2(k+1)} - 1)}{\theta^2(\theta + \gamma)^2} \\ &\neq \text{Var}(X_{t+k}|X_{t-1} = x). \end{aligned}$$

Ou seja, a esperança e variância condicionais de $X_{t+k}|X_{t-1} = x$ são assimétricas.

Corolário 1. *A transformada de Laplace condicional de X_{t+k} dado X_{t-1} é dada por:*

$$\phi_{X_{t+k}|X_{t-1}=x_{t-1}}(s) = \frac{\exp(-s\rho^{k+1}x_{t-1})(\gamma + \theta + s)(\theta + s\rho^{k+1})^{\alpha+1}}{(\gamma + \theta + s\rho^{k+1})(\theta + s)^{\alpha+1}}.$$

A demonstração do corolário acima encontra-se no Apêndice C. Ademais, note que quando k assume valores muito altos a transformada de Laplace condicional de $k + 1$ passos a frente se iguala a transformada de Laplace de X .

3.3.2 Distribuição conjunta

É sabido que a transformada de Laplace conjunta de (X_{t-1}, X_t) é dada por

$$\phi_{X_{t-1}, X_t}(s_1, s_2) = \mathbb{E}[\exp(-s_1 X_{t-1} - s_2 X_t)].$$

De (3.1), tem-se que

$$\phi_{X_{t-1}, X_t}(s_1, s_2) = \mathbb{E}[\exp(-s_1 X_{t-1} - s_2 (X_{t-1} + \varepsilon_t))].$$

Mas, pela independência entre X_{t-1} e ε_t

$$\begin{aligned} \phi_{X_{t-1}, X_t}(s_1, s_2) &= \mathbb{E}[\exp(-(s_1 + s_2)X_{t-1})] \mathbb{E}[\exp(-s_2 \varepsilon_t)] \\ &= \phi_X(s_1 + \rho s_2) \phi_\varepsilon(s_2). \end{aligned}$$

Portanto, de (2.5) e (3.3), é fácil mostrar que

$$\phi_{X_{t-1}, X_t}(s_1, s_2) = \frac{(\theta + \gamma + s_1 + \rho s_2)(\theta + \gamma + s_2)[\theta(\theta + \rho s_2)]^{\alpha+1}}{(\theta + \gamma)(\theta + \gamma + \rho s_2)[(\theta + s_1 + \rho s_2)(\theta + s_2)]^{\alpha+1}}.$$

Por fim, note que $\phi_{X_{t-1}, X_t}(0, 0) = 1$ e que a transformada de Laplace conjunta não é simétrica em X_t e X_{t-1} . Consequentemente, tem-se que o processo em questão não é reversível no tempo.

3.4 Funções de autocorrelação e autocorrelação parcial

Ao realizar uma análise de séries temporais, trabalha-se com um conjunto de dados ordenados ao longo do tempo associados a uma determinada variável que usualmente possui dependência temporal. Nesse sentido, dada a necessidade de se estudar as relações de dependência da variável investigada, algumas funções de grande importância na análise de séries temporais são abordadas nesta seção.

A função de autocovariância mensura a covariância entre a variável de saída no instante atual, X_t , e a própria variável até l instantes de atraso, X_{t-l} . Analogamente, a

acf mede a correlação entre X_t e os valores da variável de saída defasados até X_{t-l} , ou seja, a acf estabelece a influência que os l valores anteriores da variável de saída têm sobre a mesma no instante t .

No caso abordado por este trabalho, isto é, processos autorregressivos de primeira ordem, tem-se que a função de autocorrelação para um atraso l é dada por

$$\eta_{X_t}(l) = \rho^{|l|}.$$

Conseqüentemente, e graças à estacionaridade do modelo, a função de autocovariância de um processo AR(1) é dada por

$$\gamma_{X_t}(l) = \rho^{|l|} \text{Var}(X_t). \quad (3.6)$$

Note que tanto a função de autocovariância quanto a função de autocorrelação do modelo não dependem do instante observado, o que já era esperado visto que o processo estudado é estacionário.

Com o intuito de avaliar o comportamento do processo proposto, a Figura 1 apresenta os gráficos da função de autocorrelação do processo LGAR(1) em amostras de tamanho 100 para três diferentes valores de ρ .

Note que conforme o *lag* aumenta, o valor da acf decresce exponencialmente para 0. Ou seja, quanto maior for a distância no tempo entre uma determinada observação e o valor atual, menor será a influência que tal observação tem sobre a observação atual.

Além disso, conforme o valor de ρ aumenta e se aproxima de 1, o decaimento da acf torna-se mais lento. Tais características já eram esperadas, visto que se trata de um processo autorregressivo de ordem 1.

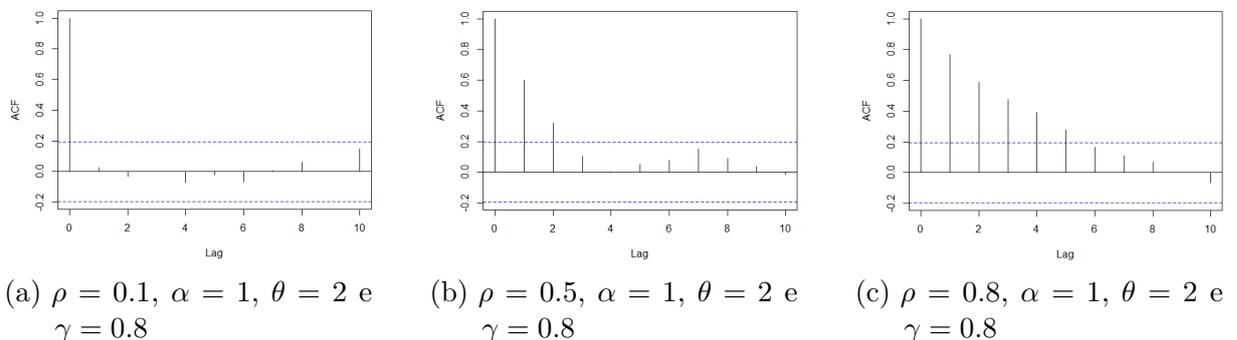


Figura 1 – Função de autocorrelação do processo proposto em três cenários diferentes.

Outra função de grande importância na análise de séries temporais é a função de autocorrelação parcial. Ela mensura a influência direta que uma dada observação, X_t , tem sobre uma observação l passos à frente, X_{t+l} , desconsiderando possíveis interferências exercidas pelas variáveis intermediárias $X_{t+1}, X_{t+2}, \dots, X_{t+l-1}$. Conseqüentemente,

a função de autocorrelação parcial funciona como uma função de correlação condicional: $\text{Cor}(X_{t+l}, X_t | X_{t+1}, X_{t+2}, \dots, X_{t+l-1})$.

Grande parte de sua relevância e utilização reside no fato de que a pacf é amplamente empregada na identificação da ordem de um modelo autorregressivo. Neste sentido, a ordem do modelo é dada pelo atraso associado ao maior valor da pacf. Ou seja, os valores de pico indicam o *lag* de maior influência sobre as demais observações da série.

A Figura 2 apresenta três gráficos da função de autocorrelação parcial do processo LGAR(1) sob as mesmas condições e cenários descritos previamente para a função de autocorrelação.

A figura em questão mostra que em todos os cenários abordados o ponto de pico ocorre no *lag* 1, o que já era esperado, uma vez que o processo explorado é autorregressivo de ordem 1. Além disso, conforme o valor de ρ se aproxima de 1, o pico no *lag* 1 torna-se cada vez mais evidente.

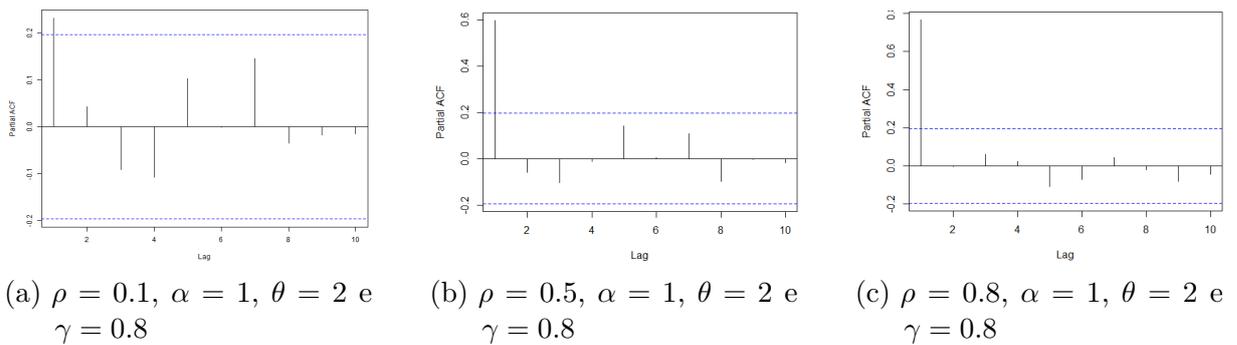


Figura 2 – Função de autocorrelação parcial do processo proposto em três cenários diferentes.

3.5 Densidade espectral

A densidade espectral é uma outra função muito útil na análise de séries temporais. O *power spectrum*, ou densidade espectral, é obtido no domínio das frequências e pode ser visto como análogo ao que a função de autocovariância é para o domínio temporal.

Qualquer processo estacionário possui uma representação no domínio do tempo e uma no domínio das frequências, sendo que cada característica dos dados que pode ser descrita por uma dessas representações pode ser igualmente bem descrita pela outra representação (Hamilton^[26], 1994).

O *power spectrum* pode ser obtido em função da função de autocovariância como

mostrado a seguir

$$\begin{aligned} f_{XX}(w) &= \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \text{Cov}(X_t, X_{t-l}) e^{-iwl} \\ &= \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \gamma_{X_t}(l) e^{-iwl}, \end{aligned}$$

em que $w \in (-\pi, \pi]$.

Proposição 4. *A densidade espectral do processo LGAR(1) é dada por*

$$f_{XX}(w) = \frac{[\theta(2\gamma\alpha + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)](1 - \rho^2)}{2\pi\theta^2(\theta + \gamma)^2(1 + \rho^2 - 2\cos(w))}.$$

Demonstração. Dada a definição da função de densidade espectral e a Equação (3.6), tem-se que

$$\begin{aligned} f_{XX}(w) &= \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \gamma_{X_t}(l) e^{-iwl} \\ &= \frac{1}{2\pi} \sum_{l=-\infty}^{\infty} \rho^{|l|} \text{Var}(X_t) e^{-iwl} \\ &= \frac{\text{Var}(X_t)}{2\pi} \sum_{l=-\infty}^{\infty} \rho^{|l|} e^{-iwl}. \end{aligned}$$

Uma vez que $\rho, w \in \mathbb{R}$, então $\sum_{l=-\infty}^{\infty} \rho^{|l|} e^{-iwl} = \frac{1 - \rho^2}{\rho^2 - 2\cos(w) + 1}$.

Dessa forma, da Equação (2.3), tem-se que a densidade espectral do processo LGAR(1) é dada por

$$\begin{aligned} f_{XX}(w) &= \frac{\text{Var}(X_t)}{2\pi} \left(\frac{1 - \rho^2}{\rho^2 - 2\cos(w) + 1} \right) \\ &= \frac{[\theta(2\gamma\alpha + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)](1 - \rho^2)}{2\pi\theta^2(\theta + \gamma)^2(1 + \rho^2 - 2\cos(w))}. \end{aligned}$$

□

Tendo em vista o processo autorregressivo desenvolvido, estimadores para os seus parâmetros serão estudados no capítulo a seguir tomando como base duas abordagens distintas.

4 Estimação do Processo LGAR(1)

Resumo

Neste capítulo, os estimadores de mínimos quadrados condicionais dos parâmetros do processo autorregressivo de primeira ordem com distribuição marginal Lindley Generalizada são desenvolvidos. Simulações de Monte Carlo são realizadas para avaliar o comportamento de tais estimadores e mostrar a sua consistência.

Palavras-chave: Estimação. Método de mínimos quadrados condicionais. Método de Monte Carlo.

4.1 Introdução

Tendo em vista o processo autorregressivo de primeira ordem proposto neste trabalho, e suas propriedades previamente abordadas, é inegável a necessidade de se estudar os estimadores de seus parâmetros. Para tanto, é preciso não só escolher um método de estimação como principalmente definir uma forma de se avaliar os estimadores encontrados. Nesse sentido, o método de Monte Carlo foi empregado para avaliar o comportamento dos estimadores obtidos através do método de mínimos quadrados condicionais (mqc) e do método Gaussiano.

O mecanismo de Monte Carlo (Metropolis e Ulam^[33], 1949) é um procedimento não determinístico que tem vasta aplicação na avaliação da performance de técnicas inferenciais em amostras finitas. Tal ferramenta foi citada pela primeira vez durante a Segunda Guerra e seu nome vem do Cassino de Monte Carlo, em Mônaco.

O que garante a confiabilidade dos resultados obtidos através das simulações de Monte Carlo é a Lei Fraca dos Grandes Números, que afirma que, para um dado conjunto de dados com variância finita e esperança constante, a média amostral converge em probabilidade para a média populacional.

Dessa forma, o procedimento de Monte Carlo consiste basicamente de se especificar o método de geração dos dados, gerar um número n de amostras, aplicar a ferramenta inferencial sem levar em consideração o mecanismo gerador dos dados e, por fim, avaliar o desempenho de tal ferramenta em relação aos pressupostos determinados inicialmente.

A quantidade escolhida para avaliar as estimativas obtidas nas simulações implementadas foi o erro quadrático médio (EQM). O EQM de um estimador $\hat{\eta}$ de um parâmetro η é a medida dada por $EQM(\hat{\eta}) = \mathbb{E}[(\hat{\eta} - \eta)^2]$, a qual é amplamente utilizada na literatura

na avaliação e comparação de estimadores.

4.2 Método dos Mínimos Quadrados Condicionais

O método de estimação dos mínimos quadrados condicionais para processos estocásticos foi proposto por Klimko e Nelson^[34], em 1978. Os autores mostraram que os estimadores produzidos são consistentes, assintoticamente normais e com taxa de convergência de $\{\log[\log(n)]/n\}^{\frac{1}{2}}$.

Tal método já foi utilizado na estimação dos parâmetros de diversos outros processos autorregressivos, tais como, a classe de modelos autorregressivos para dados de contagem Binomial proposta por WeiB^[35] (2009), o processo autorregressivo de primeira ordem Lindley apresentado por Popovic e Bakouch^[15] (2016), o processo autorregressivo de ordem um Poisson-Lindley (Mohammadpour, Bakouch e Shirozhan, 2018)^[24] e o processo autorregressivo de primeira ordem Gama-Lindley proposto por Mello, Lima e Nascimento^[25] em 2021.

Seja X_t um processo estocástico com $t = 1, 2, \dots$. O procedimento em questão consiste de minimizar a soma do quadrado dos desvios entre o processo no instante t e o valor esperado do processo em t condicionado ao seu valor no instante anterior. Dessa forma, minimiza-se a Equação dada por

$$Q_n(\rho, \alpha, \theta, \gamma) = \sum_{t=2}^n [X_t - \mathbb{E}(X_t|X_{t-1})]^2,$$

com relação ao vetor paramétrico $(\rho, \alpha, \theta, \gamma)^T$.

Ao aplicar a equação acima no processo LGAR(1), obtém-se

$$\begin{aligned} Q_n(\rho, \alpha, \theta, \gamma) &= \sum_{t=2}^n [X_t - \mathbb{E}(\rho X_{t-1} + \varepsilon_t)]^2 \\ &= \sum_{t=2}^n [X_t - \rho X_{t-1} - \mathbb{E}(X_t)(1 - \rho)]^2 \\ &= \sum_{t=2}^n \left[X_t - \rho X_{t-1} + \frac{(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)}{\theta(\theta + \gamma)} \right]^2. \end{aligned}$$

Portanto, os estimadores são encontrados ao se resolver o sistema das seguintes equações:

$$\begin{aligned} \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \rho} &= 0, \\ \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \alpha} &= 0, \\ \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \theta} &= 0 \text{ e} \\ \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \gamma} &= 0, \end{aligned}$$

em que os valores das derivadas parciais encontram-se no Apêndice B.

Dessa forma, resolvendo o sistema de equações acima, tem-se que

$$\hat{\rho} = \frac{\sum_{t=2}^n \left\{ \left[X_t - \frac{(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})}{\hat{\theta}(\hat{\theta} + \hat{\gamma})} \right] \left[\frac{(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})}{\hat{\theta}(\hat{\theta} + \hat{\gamma})} - X_{t-1} \right] \right\}}{\sum_{t=2}^n \left[X_{t-1} - \frac{(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})}{\hat{\theta}(\hat{\theta} + \hat{\gamma})} \right]^2},$$

$$\hat{\alpha} = \frac{(\hat{\rho} - 1) \sum_{t=2}^n (X_t - \hat{\rho} X_{t-1})}{\hat{\theta}(n-1)} + \frac{\hat{\gamma}(\hat{\rho} - 1)^2}{\hat{\theta}^2(\hat{\theta} + \hat{\gamma})},$$

$$\hat{\gamma} = \frac{-\hat{\theta}^2 \sum_{t=2}^n X_t + \hat{\rho}\hat{\theta}^2 \sum_{t=2}^n X_{t-1} - \hat{\theta}\hat{\alpha}(\hat{\rho} - 1)(n-1)}{\hat{\theta} \sum_{t=2}^n X_t - \hat{\rho}\hat{\theta} \sum_{t=2}^n X_{t-1} + (\hat{\rho} - 1)(\hat{\alpha} + 1)(n-1)},$$

e

$$2[\hat{\theta}(\hat{\theta} + \hat{\gamma})]^{-3} \sum_{t=2}^n [\hat{\alpha}\hat{\theta}^2(\hat{\theta} + \hat{\gamma})^2(\hat{\rho} - 1)X_t - \hat{\theta}(\hat{\theta} + \hat{\gamma})(\hat{\rho} - 1)(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})(2\hat{\theta} + \hat{\gamma})X_t$$

$$- \hat{\rho}\hat{\alpha}\hat{\theta}^2(\hat{\theta} + \hat{\gamma})^2(\hat{\rho} - 1)X_{t-1} + \hat{\rho}\hat{\theta}(\hat{\theta} + \hat{\gamma})(\hat{\rho} - 1)(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})(2\hat{\theta} + \hat{\gamma})X_{t-1}$$

$$+ \hat{\alpha}\hat{\theta}(\hat{\theta} + \hat{\gamma})(\hat{\rho} - 1)^2(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma}) - (2\hat{\theta} + \hat{\gamma})(\hat{\rho} - 1)^2(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})^2] = 0.$$

Note que não foi possível isolar o $\hat{\theta}$ na equação acima.

4.3 Método Gaussiano

O método de estimação Gaussiano foi proposto por Whittle em 1961 e é empregado para se obter estimadores aproximados dos parâmetros desconhecidos do modelo estudado. Tal aproximação é realizada através da verossimilhança Gaussiana, mas sem que haja a suposição de normalidade dos dados.

Esse método foi estudado por Crowder^[36] (1985) com aplicação a dados binários correlacionados. Em seu trabalho, o autor mostrou que, até mesmo para observações com distribuição tão distante da curva Gaussiana quanto dados binários, o método Gaussiano apresentou bons resultados.

A ferramenta em questão também já foi utilizada na estimação dos parâmetros desconhecidos de alguns processos autorregressivos de primeira ordem, dentre os quais pode-se destacar o modelo Lindley autorregressivo proposto por Popovic e Bakouch^[15] (2016) e o processo Gama-Lindley autorregressivo desenvolvido por^[25] em 2021.

Dessa forma, com o intuito de se obter estimadores dos parâmetros do processo através do método de estimação Gaussiano, considere a função de verossimilhança condicional do modelo proposto

$$L(\rho, \alpha, \theta, \gamma) = \prod_{t=2}^n f(x_t | x_{t-1}).$$

Substituindo $f(x_t|x_{t-1})$ pela densidade da distribuição normal, tem-se que

$$\begin{aligned} L(\rho, \alpha, \theta, \gamma) &= \prod_{t=2}^n \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma_{x_t}^2}} \exp \left[\frac{-(x_t - \mu_{x_t})^2}{2\sigma_{x_t}^2} \right] \\ &= (2\pi)^{-\frac{n}{2}} \prod_{t=2}^n \exp \left\{ \frac{-1}{2} \left[\frac{(x_t - \mu_{x_t})^2}{\sigma_{x_t}^2} + \log(\sigma_{x_t}^2) \right] \right\}. \end{aligned}$$

Logo, a função de log-verossimilhança será dada por:

$$l(\rho, \alpha, \theta, \gamma) = \frac{-n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[\frac{(x_t - \mu_{x_t})^2}{\sigma_{x_t}^2} + \log(\sigma_{x_t}^2) \right],$$

em que, por definição do método, as quantidades $\mu_{x_t} = \mathbb{E}(X_t|X_{t-1} = x_{t-1})$ e $\sigma_{x_t}^2 = \text{Var}(X_t|X_{t-1} = x_{t-1})$ são respectivamente obtidas a partir das Equações (3.4) e (3.5) ao se tomar $k = 0$.

Dessa forma, os estimadores são obtidos como solução do sistema de equações

$$\begin{aligned} \frac{\partial l(\rho, \alpha, \theta, \gamma)}{\partial \rho} &= 0, \\ \frac{\partial l(\rho, \alpha, \theta, \gamma)}{\partial \alpha} &= 0, \\ \frac{\partial l(\rho, \alpha, \theta, \gamma)}{\partial \theta} &= 0 \text{ e} \\ \frac{\partial l(\rho, \alpha, \theta, \gamma)}{\partial \gamma} &= 0, \end{aligned}$$

em que

$$\begin{aligned} l(\rho, \alpha, \theta, \gamma) &= \frac{-n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[x_t - \rho x_{t-1} - \frac{(1-p)(\theta\alpha + \alpha\gamma + \gamma)}{\theta(\theta + \gamma)} \right]^2 \\ &\quad \times \left[\frac{\theta^2(\theta + \gamma)^2}{(1-\rho^2)[\theta(2\alpha\gamma + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)]} \right] \\ &\quad - \frac{n-1}{2} \log \left\{ \frac{(1-\rho^2)[\theta(2\alpha\gamma + \theta\alpha + 2\gamma) + \gamma^2(\alpha^2 + \alpha + 1)]}{\theta^2(\theta + \gamma)^2} \right\}. \end{aligned}$$

Dada a alta extensão das derivadas parciais presentes no sistema apresentado, as mesmas não serão expostas no presente trabalho.

4.4 Resultados das Estimções

O método de Monte Carlo foi utilizado para avaliar a performance dos estimadores obtidos através dos métodos de mínimos quadrados condicionais e Gaussiano do processo autorregressivo de ordem 1 com distribuição marginal Lindley Generalizada. Para tanto, foram realizadas 1000 réplicas de Monte Carlo, para três diferentes tamanhos amostrais

$n = \{100, 250, 500\}$ e cinco diferentes cenários, em que tanto os tamanhos amostrais quanto os cenários escolhidos foram definidos de maneira arbitrária.

Em cada um dos cenários abordados, os valores dos parâmetros α , θ e γ foram mantidos constantes enquanto ρ assumiu os valores 0.1, 0.3, 0.5, 0.7 e 0.9, para que se possa avaliar a influência da variação de ρ ao manter os demais parâmetros constantes. Os resultados obtidos encontram-se resumidos nas Tabelas 1 e 2.

No que diz respeito às estimativas encontradas através do método de mqc, é notório que o ρ foi subestimado em todos os cenários avaliados, ou seja, as estimativas médias do ρ foram inferiores ao seu valor real. Apesar disso, seu erro quadrático médio foi inferior a 0.01 em todos os casos observados e decai conforme o tamanho amostral aumenta.

Em oposição ao que se verifica em ρ , os demais parâmetros (α , θ e γ) têm viés médio positivo, ou seja, suas estimativas médias são superestimadas em todos os cenários. Além disso, o erro quadrático médio dos parâmetros α , θ e γ tende a diminuir conforme o tamanho amostral aumenta. Em dois casos específicos, o EQM do θ aumenta juntamente com o n , mas é um aumento notado apenas na terceira casa decimal. Algo similar ocorre com o γ , sendo dois leves aumentos na terceira casa decimal do seu erro quadrático médio e um na quarta.

Já com relação aos resultados obtidos por meio do método Gaussiano, tem-se que, similarmente ao que foi observado para o método de mínimos quadrados condicionais, o valor de ρ foi subestimado em todas as simulações desenvolvidas. Ademais, seu erro quadrático médio diminui à medida que o tamanho amostral aumenta, de forma tal que o maior EQM observado (0.0205) ocorre em $n = 100$.

Assim como o que se destacou para ρ , as estimativas médias dos parâmetros α , θ e γ de acordo com o método Gaussiano comportam-se de maneira semelhante àquelas obtidas por mínimos quadrados condicionais. Em que θ e γ são superestimados em todos os cenários observados enquanto α apresenta viés positivo na maioria das simulações realizadas.

Apesar do erro quadrático médio dos três últimos parâmetros apresentar uma tendência a decair ao passo que o tamanho amostral é acrescido de algumas dezenas de unidades, pode-se notar que tal medida cresce em conjunto com n em alguns casos específicos.

Nesse sentido, percebe-se que os erros quadráticos médios de α sofre um aumento na segunda casa decimal, de θ tem um acréscimo na terceira casa decimal e o de γ apresenta quatro pequenos acréscimos. É importante ressaltar que os valores representados por "-" indicam os cenários em que não se foi possível obter as estimativas dos parâmetros através do método Gaussiano.

Tendo em vista tudo o que foi exposto, o método de estimação de mínimos quadrados

condicionais se mostrou mais indicado que o método Gaussiano para a obtenção das estimativas dos parâmetros do modelo aqui estudado. Uma vez que, além do método Gaussiano não apresentar consistência para três dos quatro parâmetros do processo, seu tempo de execução para valores de ρ superiores a 0.5 foi tão elevado que impossibilitou a obtenção de resultados para quase todos os cenários em que $\rho > 0.5$. Apesar disso, é válido ressaltar que ambos os métodos apresentaram alguns erros quadráticos médios um pouco elevados.

Tabela 1 – Estimativa e Erro Quadrático Médio dos parâmetros do modelo LGAR sob o método de Mínimos Quadrados Condicionais

Cenários	n	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\gamma}$
(0.1, 2, 2, 0.5)	100	0.0853	3.3167	2.6260	1.0596
		(0.0094)	(1.7405)	(0.4009)	(0.3134)
	250	0.0935	3.3122	2.6261	1.0589
		(0.0042)	(1.7245)	(0.3960)	(0.3124)
	500	0.096	3.3104	2.6293	1.0585
		(0.002)	(1.7185)	(0.3978)	(0.3119)
(0.3, 2, 2, 0.5)	100	0.2832	2.8850	3.1275	0.9785
		(0.0094)	(0.8002)	(1.2922)	(0.2296)
	250	0.2933	2.8832	3.1321	0.9781
		(0.0036)	(0.7869)	(1.2903)	(0.2289)
	500	0.2977	2.8789	3.1365	0.9773
		(0.0017)	(0.7760)	(1.2959)	(0.2279)
(0.5, 2, 2, 0.5)	100	0.4767	2.4051	3.5929	0.8906
		(0.0076)	(0.2126)	(2.5754)	(0.1541)
	250	0.4919	2.4254	3.5922	0.8938
		(0.0031)	(0.1977)	(2.5500)	(0.1556)
	500	0.4961	2.4184	3.5873	0.8928
		(0.0015)	(0.1845)	(2.5274)	(0.1546)
(0.7, 2, 2, 0.5)	100	0.6684	2.0397	3.9071	0.8243
		(0.005)	(0.1025)	(3.7293)	(0.1085)
	250	0.6849	2.0657	3.8833	0.8292
		(0.0021)	(0.0438)	(3.5833)	(0.1097)
	500	0.6925	2.0702	3.8591	0.8305
		(0.001)	(0.0260)	(3.4710)	(0.1098)
(0.9, 2, 2, 0.5)	100	0.8289	2.1999	3.7778	0.8529
		(0.0066)	(0.2657)	(3.3768)	(0.1322)
	250	0.8599	2.1886	3.7938	0.8508
		(0.0022)	(0.1745)	(3.3474)	(0.1277)
	500	0.8724	2.1837	3.7883	0.8503
		(0.0011)	(0.0948)	(3.2459)	(0.1247)

Tabela 2 – Estimativa e Erro Quadrático Médio dos parâmetros do modelo GLAR sob o método Gaussiano

Cenários	n	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\gamma}$
(0.1, 2, 2, 0.5)	100	0.0869	3.5774	2.7627	0.7746
		(0.0090)	(2.5883)	(0.6085)	(0.1579)
	250	0.0942	3.5596	2.7562	0.7871
		(0.0041)	(2.4867)	(0.5842)	(0.1299)
	500	0.0956	3.5546	2.7626	0.7974
		(0.0018)	(2.4426)	(0.5883)	(0.1102)
(0.3, 2, 2, 0.5)	100	0.2530	2.6157	2.8803	1.2023
		(0.0085)	(0.5751)	(0.8542)	(0.5792)
	250	0.2789	2.5579	2.8644	1.2571
		(0.0034)	(0.3853)	(0.7749)	(0.6072)
	500	0.2892	2.5104	2.8334	1.2896
		(0.0016)	(0.3021)	(0.7123)	(0.6423)
(0.5, 2, 2, 0.5)	100	0.3819	1.8846	3.0321	1.2477
		(0.0205)	(0.2552)	(1.2745)	(0.6423)
	250	0.4241	1.8665	2.9753	1.29678
		(0.0076)	(0.1138)	(1.0396)	(0.6706)
	500	0.4783	1.6572	2.7579	1.4364
		(0.0029)	(0.1659)	(0.6227)	(0.8948)
(0.7, 2, 2, 0.5)	100	0.6094	1.8594	3.4496	1.0432
		(0.0122)	(0.2614)	(2.4323)	(0.3983)
	250	—	—	—	—
		—	—	—	—
	500	—	—	—	—
		—	—	—	—
(0.9, 2, 2, 0.5)	100	—	—	—	—
		—	—	—	—
	250	—	—	—	—
		—	—	—	—
	500	—	—	—	—
		—	—	—	—

5 Aplicação a dados reais

Resumo

Neste capítulo, uma aplicação a dados reais foi realizada para avaliar o desempenho do processo proposto quando comparado a outros seis processos de características semelhantes conhecidos na literatura. Os modelos candidatos e os critérios de comparação são apresentados. Por fim, levando-se em consideração o critério de informação de Akaike e o critério de informação Bayesiano, o modelo aqui proposto obteve resultados superiores aos demais modelos.

Palavras-chave: Aplicação a dados reais, critério de informação de Akaike, critério de informação Bayesiano.

5.1 Introdução

Ao propor um novo modelo, é indispensável avaliar o seu desempenho em aplicações a dados reais com relação a outros modelos previamente estabelecidos na literatura, com o intuito de mostrar a sua relevância na área. Desse modo, três aplicações a dados reais foram realizadas e o ajuste ofertado pelo modelo proposto foi comparado ao ajuste obtido através de outros 7 processos autorregressivos de primeira ordem.

Na escolha do melhor modelo, levou-se em consideração tanto a qualidade dos ajustes quanto a complexidade dos processos avaliados. Por complexidade, entenda a quantidade de parâmetros presente no processo, ou seja, quanto maior o tamanho do vetor paramétrico, maior a complexidade do modelo. Dessa forma, o objetivo é escolher o processo que com a menor complexidade possível ofereça o melhor ajuste. Para tanto, dois critérios de comparação foram escolhidos de forma tal que a qualidade de ajuste do modelo em questão sofre penalidades de acordo com o seu número de parâmetros.

5.2 Modelos candidatos

- Processo EAR(1) proposto por Gaver e Lewis^[21] em 1980 e cuja função de log-verossimilhança gaussiana, necessária para a estimação dos parâmetros na aplicação a dados reais, é dada por

$$l(\rho, \lambda) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[\frac{\lambda^2}{1-\rho^2} \left(x_t - \rho x_{t-1} - \frac{1-\rho}{\lambda} \right)^2 + \log \left(\frac{1-\rho^2}{\lambda^2} \right) \right].$$

- Processo GAR(1) proposto por Gaver e Lewis^[21] em 1980 e cuja função de log-verossimilhança gaussiana, necessária para a estimação dos parâmetros na aplicação a dados reais, é dada por

$$l(\rho, \lambda, k) = -\frac{n}{2}\log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[\frac{\lambda^2}{k(1-\rho^2)} \left(x_t - \rho x_{t-1} - \frac{k(1-\rho)}{\lambda} \right)^2 + \log \left(\frac{k(1-\rho^2)}{\lambda^2} \right) \right].$$

- Processo IGAR(1) proposto por Abraham e Balakrishna^[22] em 1999 e cuja função de log-verossimilhança gaussiana, necessária para a estimação dos parâmetros na aplicação a dados reais, é dada por

$$l(\rho, \mu, \lambda) = -\frac{n}{2}\log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[\frac{\lambda}{\mu^3(1-\rho^2)} (x_t - \rho x_{t-1} - (1-\rho)\mu)^2 + \log \left(\frac{\mu^3(1-\rho^2)}{\lambda} \right) \right].$$

- Processo GLAR(1) proposto por Jose e Thomas^[23] em 2011 e cuja função de log-verossimilhança gaussiana, necessária para a estimação dos parâmetros na aplicação a dados reais, é dada por

$$l(\rho, \theta, k, \sigma, \tau) = -\frac{n}{2}\log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[\frac{(x_t - \rho x_{t-1} - (1-\rho)\tau\nu)^2}{\tau(1-\rho^2)(\nu^2 + \sigma^2)} + \log \left(\tau(1-\rho^2)(\nu^2 + \sigma^2) \right) \right],$$

$$\text{em que } \nu = \frac{\sigma}{\sqrt{2}} \left(\frac{1}{k} - k \right).$$

- Processo LAR(1) proposto por Popovic e Bakouch^[15] em 2016 e cuja função de log-verossimilhança gaussiana, necessária para a estimação dos parâmetros na aplicação a dados reais, é dada por

$$l(\rho, \lambda) = -\frac{n}{2}\log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[\frac{\lambda^2(\lambda+1)^2 \left(x_t - \rho x_{t-1} - \frac{(1-\rho)(\lambda+2)}{\lambda(\lambda+1)} \right)^2}{(1-\rho^2)(\lambda^2 + 4\lambda + 2)} + \log \left(\frac{(1-\rho^2)(\lambda^2 + 4\lambda + 2)}{\lambda^2(\lambda+1)^2} \right) \right].$$

- Processo GaLAR(1) proposto por Mello, Lima e Nascimento^[25] em 2021 e cuja função de log-verossimilhança gaussiana, necessária para a estimação dos parâmetros na aplicação a dados reais, é dada por

$$l(\rho, \lambda, \beta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left(x_t - \rho x_{t-1} - \frac{(1-\rho)(2\beta(1+\lambda) - \lambda)}{\lambda\beta(1+\lambda)} \right)^2 \\ \times \left(\frac{\lambda^2\beta^2(1+\lambda)^2}{[-(-2\beta\lambda + \lambda)^2 + 2\beta^2(1+3\lambda) - 2\beta(\beta\lambda - 3\beta\lambda^2 + 2\lambda^2)](1-\rho^2)} \right) \\ + (n-1)\log \left(\frac{[-(-2\beta\lambda + \lambda)^2 + 2\beta^2(1+3\lambda) - 2\beta(\beta\lambda - 3\beta\lambda^2 + 2\lambda^2)](1-\rho^2)}{\lambda^2\beta^2(1+\lambda)^2} \right).$$

5.3 Critérios de comparação

Ao ajustar diferentes modelos a um mesmo banco de dados, é necessário que haja uma forma previamente definida de se estabelecer qual deles apresentou resultados superiores. Nesse sentido, diversos critérios de comparação foram desenvolvidos ao longo das décadas, em que cada um deles pode levar em consideração diferentes características dos modelos testados.

Comumente, os critérios de comparação mais utilizados favorecem modelos mais parcimoniosos no que diz respeito à qualidade do ajuste e ao seu número de parâmetros. Dito isso, os critérios aqui empregados foram o critério de informação de Akaike (AIC) proposto por Hirotugu Akaike e o critério de informação Bayesiano (BIC) proposto por Gideon Schwarz.

Ambos os métodos citados estimam a quantidade de informação perdida pelo modelo testado através do valor máximo da função de verossimilhança. Como consequência direta, buscamos pelos menores valores do AIC e BIC pois quanto menor for a parcela de informação perdida melhor será o modelo em questão.

Apesar disso, como o aumento da verossimilhança pode ser acarretado apenas pelo acréscimo de parâmetros, os dois critérios mencionados penalizam a quantidade de parâmetros presentes no modelo. Dito isso, o AIC é calculado através de

$$AIC = 2p - 2 \log(\hat{L}),$$

e o BIC é dado por

$$BIC = p \log(n) - 2 \log(\hat{L}),$$

em que p é o número de parâmetros, o \hat{L} representa o valor máximo da função de verossimilhança e n é o tamanho amostral.

5.4 Aplicação a dados reais

O conjunto de dados utilizado na aplicação diz respeito às medições anuais, em pés, do nível do Lago Huron entre os anos de 1875 e 1972. Ele é o segundo maior lago dentre os cinco grandes lagos localizados na América do Norte, entre Canadá e Estados Unidos, sendo que ele tem a maior margem dentre os cinco. Tal lago provê alimentação, recreação e água para consumo pessoal para aproximadamente 3 milhões de pessoas na América do Norte e abriga várias plantas e animais raros e ameaçados de extinção. Além disso, ele é conhecido pela sua água clara e turquesa e por ter um pôr do sol que é dito ser um dos mais bonitos do mundo.

O banco em questão foi extensamente estudado ao longo das últimas décadas, dentre outros artigos, pode-se destacar as abordagens feitas por Brockwell e Davis^[37] em

1991, Fraccaro, Hyndman e Veevers^[38] em 2000, Karaman e Altıok^[39] em 2004, Cheng *et. al.* em 2021 e Mello, Lima e Nascimento^[25] em 2021. Tais observações são apresentadas na Figura 3.

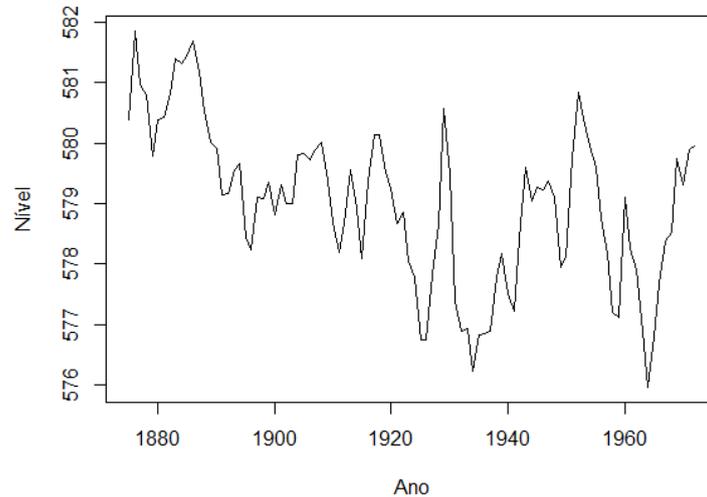


Figura 3 – Nível do Lago Huron entre os anos de 1875 e 1972.

O gráfico presente na Figura 4a mostra que a acf do nível do Lago Huron possui um decaimento lento e exponencial. Já na Figura 4b, tem-se que apesar da pacf apresentar um grande corte após o lag 1, o último lag significativo é o 2. Portanto, há fortes indícios de que um AR(1) ou AR(2) ofereceriam bons ajustes para a série em questão.

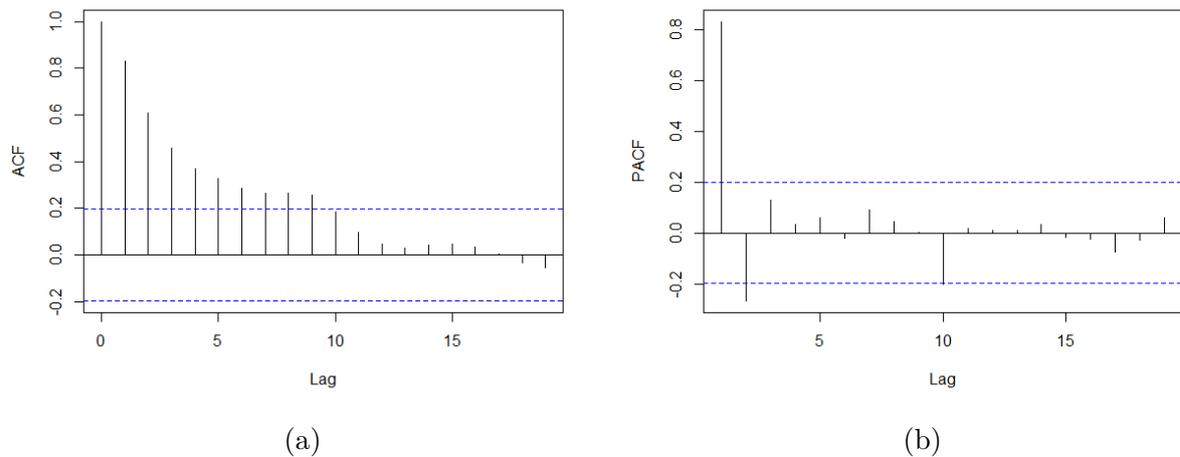


Figura 4 – Funções de autocorrelação (a) e autocorrelação parcial (b) dos dados analisados.

A Tabela 3 contém os resultados dos critérios de comparação escolhidos para cada um dos modelos ajustados. Nela, é possível notar que o modelo aqui proposto apresentou

resultados superiores aos demais concorrentes, uma vez que suas medidas em ambos critérios de informação foram inferiores às medidas dos processos restantes. Além disso, o modelo que obteve desempenho mais próximo do LGAR(1) foi o GaLAR(1), em que a diferença de performance entre o primeiro e o último ocorre na segunda casa decimal tanto do critério de informação de Akaike quanto do critério de informação Bayesiano. Note que, apesar do modelo GaLAR(1) possuir um parâmetro a menos que o processo aqui proposto, o LGAR(1) possui função densidade de probabilidade exata enquanto que a densidade do modelo GaLAR(1) é aproximada, o que pode justificar a escolha do processo LGAR(1) em detrimento do outro.

Modelo	AIC	BIC
LGAR(1)	226.054	233.809
EAR(1)	293.432	301.187
GAR(1)	229.435	237.189
IGAR(1)	289.516	297.270
GLAR(1)	255.687	263.442
LAR(1)	234.682	242.437
GaLAR(1)	226.097	233.851

Tabela 3 – Critérios de comparação para a aplicação 1

Modelo	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\gamma}$	$\hat{\lambda}$	$\hat{\beta}$	\hat{k}	$\hat{\sigma}$	$\hat{\tau}$	$\hat{\mu}$
LGAR	0.9999	0.8236	0.0140	1.0620	-	-	-	-	-	-
EAR	0.9986	-	-	-	0.0494	-	-	-	-	-
GAR	0.9998	-	-	-	0.0298	-	0.9726	-	-	-
IGAR	0.9989	-	-	-	0.0070	-	-	-	-	1.7566
GLAR	0.9994	-	3.7397	-	-	-	5.9185	3.6159	0.1589	-
LAR	0.9996	-	-	-	0.0471	-	-	-	-	-
GaLAR	0.9999	-	-	-	0.0180	0.7545	-	-	-	-

Tabela 4 – Estimativas dos parâmetros para a aplicação 1

6 Conclusão

No presente trabalho, um novo processo autorregressivo de primeira ordem foi proposto, o qual possui como marginal a distribuição Lindley Generalizada. A distribuição marginal escolhida, cuja densidade é obtida a partir da mistura das distribuições $\text{Gama}(\alpha, \theta)$ e $\text{Gama}(\alpha + 1, \theta)$, foi estudada e suas principais características foram apresentadas. Além disso, a transformada de Laplace de uma variável aleatória X com distribuição Lindley Generalizada foi desenvolvida para viabilizar a resolução de outras medidas obtidas ao longo do trabalho.

Em seguida, o processo LGAR(1) foi apresentado e a função densidade de probabilidade para a sequência de erros aleatórios associados a tal modelo foi obtida. Algumas propriedades estatísticas do modelo LGAR(1) foram desenvolvidas. Dessa forma, foram apresentadas fórmulas fechadas para esperança e variância condicionais k passos à frente, distribuição conjunta de um vetor aleatório bidimensional do processo e densidade espectral. Além disso, as funções de autocorrelação e autocorrelação parcial foram graficamente analisadas.

Duas abordagens distintas foram utilizadas na estimação dos parâmetros do modelo estudado. Assim, os parâmetros do processo aqui proposto foram estimados através do método Gaussiano e do método de mínimos quadrados condicionais. Simulações de Monte Carlo foram desenvolvidas para avaliar a qualidade dos estimadores obtidos. O método de mqc apresentou resultados superiores, ao passo que não foi possível obter as estimativas dos parâmetros em todos os cenários avaliados através do método Gaussiano, tendo em vista o alto tempo computacional exigido.

Por fim, uma aplicação a dados reais foi realizada utilizando um conjunto de dados muito utilizado na literatura, a saber, as medições do nível do Lago Huron. Os critérios de comparação escolhidos foram o AIC e o BIC. Com relação a ambos critérios, o modelo proposto ofereceu ajustes superiores a outros seis processos autorregressivos de ordem um.

6.1 Trabalhos Futuros

Dentre os tópicos que não foram abordados no presente trabalho, mas que seria interessante estudar, podem-se elencar:

1. Estudar as *Run Probabilities*;
2. Otimizar o código desenvolvido para as simulações com o intuito de diminuir o tempo de execução e possibilitar a obtenção das estimativas não encontradas no Capítulo 4;

3. Averiguar outros métodos de estimação dos parâmetros do processo LGAR(1);
4. Examinar o comportamento do modelo proposto em outros tipos de dados além dos que foram abordados no Capítulo 5.

Referências

- 1 LINDLEY, D. V. Fiducial distributions and bayes' theorem. **Journal of the Royal Statistical Society**, v. 20, n. 1, p. 102:107, 1958. Citado na página 12.
- 2 GHITANY M. E.; ATIEH, B.; Nadarajah S. Lindley distribution and its application. **ScienceDirect**, Julho 2007. Citado na página 12.
- 3 ZAMANI H.; ISMAIL, N. Negative binomial-lindley distribution and its application. **Journal of Mathematics and Statistics**, p. 4:9, 2010. Citado na página 12.
- 4 MAZUCHELI J.; ACHCAR, J. A. The lindley distribution applied to competing risks lifetime data. **computer methods and programs in biomedicine**, p. 188:192, Março 2011. Citado na página 13.
- 5 BAKOUCH H. S.; AL-ZAHRANI, B. M.; Al-Shomrani A. A.; Marchi V. A. A.; Louzada F. An extended lindley distribution. **Journal of the Korean Statistical Society**, v. 41, p. 75:85, 2012. Citado na página 13.
- 6 GHITANY M. E.; AL-MUTAIRI, D. K.; Balakrishnan-N.; Al-Enezi L. J. Power lindley distribution and associated inference. **Computational Statistics and Data Analysis**, p. 20:33, 2013. Citado 2 vezes nas páginas 13 e 15.
- 7 SHANKER R.; SHARMA, S.; Shanker R. A two-parameter lindley distribution for modeling waiting and survival times data. **Applied Mathematics**, p. 363:368, 2013. Citado na página 13.
- 8 SHANKER R.; MISHRA, A. A quasi lindley distribution. **African Journal of Mathematics and Computer Science Research**, v. 6, n. 4, p. 64:71, Abril 2013. Citado na página 13.
- 9 ÇAKMAKYAPAN S.; KADILAR, G. O. A new customer lifetime duration distribution: The kumaraswamy lindley distribution. **International Journal of Trade, Economics and Finance**, v. 5, n. 5, p. 441:444, 2014. Citado na página 14.
- 10 MEROVCI F.; SHARMA, V. K. The beta-lindley distribution: Properties and applications. **Journal of Applied Mathematics**, 2014. Citado na página 14.
- 11 GUI W.; ZHANG, S.; Lu X. The lindley-poisson distribution in lifetime analysis and its properties. **Hacettepe Journal of Mathematics and Statistics**, v. 43, n. 6, p. 1063:1077, 2014. Citado na página 14.
- 12 BHATI D.; MALIK, M. A.; VAMAN, H. J. Lindley–exponential distribution: properties and applications. **METRON**, 2015. Citado na página 14.
- 13 ZOGRAFOS K.; BALAKRISHNAN, N. On families of beta- and generalized gamma-generated distributions and associated inference. **Statistical Methodology**, v. 6, n. 4, p. 344:362, Julho 2009. Citado 2 vezes nas páginas 14 e 21.

- 14 LIMA, Maria do Carmo Soares de. **Mathematical Properties Of Some Generalized Gamma Models**. Tese (Doctoral thesis), Recife, 2015. Citado na página 14.
- 15 POPOVIĆ B. V.; BAKOUCH, H. S. Lindley first-order autoregressive model with applications. **Communications in Statistics - Theory and Methods**, 2016. Citado 6 vezes nas páginas 15, 17, 18, 34, 35 e 42.
- 16 BARCO K. V. P.; MAZUCHELI, J.; Janeiro V. The inverse power lindley distribution. **Communications in Statistics - Simulation and Computation**, 2016. Citado na página 15.
- 17 SHARMA V. K.; SINGH, S. K.; Singh U.; Agiwal V. The inverse lindley distribution: a stress-strength reliability model with application to head and neck cancer data. **Journal of Industrial and Production Engineering**, 2015. Citado na página 15.
- 18 ÖZEL G.; ÇAKMAKYAPAN, S. A new generalized poisson lindley distribution. **American Institute of Physics**, 2017. Citado na página 15.
- 19 ABEBE B.; SHANKER, R. A discrete lindley distribution with applications in biological sciences. **Biometrics Biostatistics International Journal**, v. 7, n. 1, 2018. Citado na página 15.
- 20 PEÑA-RAMÍREZ F. A.; GUERRA, R. R.; Cordeiro G. M. The nadarajah-haghighi lindley distribution. **Anais da Academia Brasileira de Ciências**, 2019. Citado na página 16.
- 21 GAVER D. P.; LEWIS, P. A. W. First-order autoregressive gamma sequences and point processes. **Advances in Applied Probability**, v. 12, n. 3, p. 727:745, 1980. Citado 3 vezes nas páginas 16, 41 e 42.
- 22 ABRAHAM B.; BALAKRISHNA, N. Inverse gaussian autoregressive models. **Journal of time series analysis**, v. 20, n. 6, 1999. Citado 2 vezes nas páginas 16 e 42.
- 23 JOSE K. K.; THOMAS, M. M. Generalized laplacian distributions and autoregressive processes. **Communications in Statistics—Theory and Methods**, 2011. Citado 2 vezes nas páginas 17 e 42.
- 24 MOHAMMADPOUR M.; BAKOUCH, H. S.; Shirozhan M. Poisson–lindley inar(1) model with applications. **Brazilian Journal of Probability and Statistics**, v. 32, n. 2, p. 262:280, 2018. Citado 2 vezes nas páginas 17 e 34.
- 25 MELLO A. B. V.; LIMA, M. C. S.; Nascimento A. D. C. A notable gamma-lindley first-order autoregressive process: An application to hydrological data. **Environmetrics**, 2021. Citado 6 vezes nas páginas 17, 18, 34, 35, 42 e 44.
- 26 ZAKERZADEH H.; DOLATI, A. Generalized lindley distribution. **Journal of Mathematical Extension**, v. 3, n. 2, p. 13:25, 2009. Citado 5 vezes nas páginas 18, 20, 21, 23 e 31.
- 27 CORDEIRO G. M.; ORTEGA, E. M. M.; da Cunha D. C. C. The exponentiated generalized class of distributions. **Journal of Data Science**, 2013. Citado na página 21.

- 28 KUMARASWAMY, P. A generalized probability density function for doublebounded random processes. **Journal of Hydrology**, v. 46, p. 79:88, 1980. Citado na página 21.
- 29 MARSHALL A. W.; OLKIN, I. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. **Biometrika**, v. 86, n. 3, p. 641:652, Setembro 1997. Citado na página 21.
- 30 EUGENE N.; LEE, C.; Famoye F. Beta-normal distribution and its applications. **Communications in Statistics - Theory and Methods**, 2002. Citado na página 21.
- 31 ALZAATREH A.; LEE, C.; Famoye F. A new method for generating families of continuous distributions. **Metron**, p. 63:79, 2013. Citado na página 21.
- 32 NADARAJAH S.; BAKOUCH, H. S.; Tahmasbi R. A generalized lindley distribution. **Sankhya B**, p. 331:359, Setembro 2011. Citado na página 22.
- 33 METROPOLIS N.; ULAM, S. The monte carlo method. **Journal of the American Statistical Association**, 1949. Citado na página 33.
- 34 KLIMKO L. A.; NELSON, P. I. On conditional least squares estimation for stochastic processes. **The Annals of Statistics**, v. 6, n. 3, p. 629–642. Citado na página 34.
- 35 WEIB, C. H. A new class of autoregressive models for time series of binomial counts. p. 447:460, 2009. Citado na página 34.
- 36 CROWDER, M. Gaussian estimation for correlated binomial data. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 47, n. 2, p. 229:237, 1985. Citado na página 35.
- 37 BROCKWELL P. J.; DAVIS, R. A. **Time Series: Theory and Methods**. [S.l.: s.n.], 1991. Citado na página 43.
- 38 FRACCARO R.; HYNDMAN, R.; Veevers A. Residual diagnostic plots for checking for model mis-specification in time series regression. **Australian New Zealand Journal of Statistics**, 2000. Citado na página 44.
- 39 KARAMAN A. S.; ALTÖK, T. An experimental study on forecasting using tes processes. **Proceedings of the 2004 Winter Simulation Conference**, 2004. Citado na página 44.

APÊNDICE A – Proposição 3

Neste apêndice, consta a demonstração da Proposição 3 apresentada no capítulo 3.

Para se obter uma função que represente a distribuição dos erros aleatórios, aplicou-se a transformada inversa de Laplace na função $\phi_{\varepsilon_t}(s)$ dada pela equação 3.3. Nesse sentido, seja $\mathcal{L}_s^{-1}[f(s)](x)$ a transformada inversa de Laplace de $f(s)$ com relação a s . A função densidade de probabilidade dos erros será dada por

$$\begin{aligned} \mathcal{L}_s^{-1}[\phi_{\varepsilon_t}(s)](x) &= \mathcal{L}_s^{-1}\left[\frac{(\theta + \gamma + s)(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}}\right](x) \\ &= \mathcal{L}_s^{-1}\left[\frac{(\theta + \gamma)(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}} + \frac{s(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}}\right](x) \\ &= \mathcal{L}_s^{-1}\left[\frac{(\theta + \gamma)(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}}\right](x) + \mathcal{L}_s^{-1}\left[\frac{s(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}}\right](x) \\ &= \mathcal{L}_s^{-1}[C_1](x) + \mathcal{L}_s^{-1}[C_2](x) \end{aligned}$$

Em seguida, aplica-se o mecanismo de frações parciais nos termos C_1 e C_2 .

Termo C_1 :

$$\begin{aligned} C_1 &= \frac{(\theta + \gamma)(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}} \\ &= \frac{A^*}{\theta + \gamma + \rho s} + \frac{B^*}{\theta + s} + \frac{C^*}{(\theta + s)^{\alpha+1}} \\ &= \frac{A^*(\theta + s)^{\alpha+1} + B^*(\theta + \gamma + \rho s)(\theta + s)^{\alpha} + C^*(\theta + \gamma + \rho s)}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}} \end{aligned}$$

$$\therefore (\theta + \gamma)(\theta + \rho s)^{\alpha+1} = A^*(\theta + s)^{\alpha+1} + B^*(\theta + \gamma + \rho s)(\theta + s)^{\alpha} + C^*(\theta + \gamma + \rho s).$$

Para $s = -\theta$, encontra-se

$$C^* = \frac{(\theta + \gamma)(\theta - \theta\rho)^{\alpha+1}}{(\theta + \gamma - \theta\rho)}.$$

Se $s = \frac{-(\theta + \gamma)}{\rho}$, tem-se que

$$A^* = \frac{(\theta + \gamma)(-\rho\gamma)^{\alpha+1}}{(\theta\rho - \theta - \gamma)^{\alpha+1}}.$$

Considerando $s = 0$, encontra-se B^* em função de A^* e C^* de tal forma que

$$B^* = \frac{\theta^{\alpha+1}(\theta + \gamma) - A^*\theta^{\alpha+1} - C^*(\theta + \gamma)}{\theta^{\alpha}(\theta + \gamma)}.$$

Termo C_1 :

$$\begin{aligned} C_2 &= \frac{s(\theta + \rho s)^{\alpha+1}}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}} \\ &= \frac{D^*}{\theta + \gamma + \rho s} + \frac{E^*}{\theta + s} + \frac{F^*}{(\theta + s)^{\alpha+1}} \\ &= \frac{D^*(\theta + s)^{\alpha+1} + E^*(\theta + \gamma + \rho s)(\theta + s)^\alpha + F^*(\theta + \gamma + \rho s)}{(\theta + \gamma + \rho s)(\theta + s)^{\alpha+1}}. \end{aligned}$$

$$\therefore s(\theta + \rho s)^{\alpha+1} = D^*(\theta + s)^{\alpha+1} + E^*(\theta + \gamma + \rho s)(\theta + s)^\alpha + F^*(\theta + \gamma + \rho s).$$

Para $s = -\theta$, tem-se que

$$F^* = \frac{-\theta^{\alpha+2}(1 - \rho)^{\alpha+1}}{\theta + \gamma - \theta\rho}.$$

$$\text{Para } s = \frac{-(\theta + \gamma)}{\rho},$$

$$D^* = \frac{(-1)^{\alpha+2}(\theta + \gamma)\gamma^{\alpha+1}\rho^\alpha}{(\theta\rho - \theta - \gamma)^{\alpha+1}}.$$

E para $s = 0$, encontra-se E^* em função de D^* e F^* como mostrado a seguir

$$E^* = \frac{-D^*\theta^{\alpha+1} - F^*(\theta + \gamma)}{\theta^\alpha(\theta + \gamma)}.$$

Dessa forma, $\phi_\varepsilon(s)$ pode ser escrita como

$$\phi_\varepsilon(s) = \frac{A^* + D^*}{\theta + \gamma + \rho s} + \frac{B^* + E^*}{\theta + s} + \frac{C^* + F^*}{(\theta + s)^{\alpha+1}}.$$

Conseqüentemente, retomando a transformada inversa de Laplace, tem-se que

$$\begin{aligned} \mathcal{L}_s^{-1}[\phi_{\varepsilon_t}(s)](x) &= (A^* + D^*)\mathcal{L}_s^{-1}\left[\frac{1}{\theta + \gamma + \rho s}\right](x) + (B^* + E^*)\mathcal{L}_s^{-1}\left[\frac{1}{\theta + s}\right](x) \\ &\quad + (C^* + F^*)\mathcal{L}_s^{-1}\left[\frac{1}{(\theta + s)^{\alpha+1}}\right](x) \\ &= (A^* + D^*)\frac{e^{-\frac{(\theta+\gamma)x}{\rho}}}{\rho} + (B^* + E^*)e^{-\theta x} + (C^* + F^*)\frac{x^\alpha e^{-\theta x}}{\gamma(\alpha + 1)} \\ &= \left(\frac{A^* + D^*}{\theta + \gamma}\right)\frac{(\theta + \gamma)e^{-\frac{(\theta+\gamma)x}{\rho}}}{\rho} + \left(\frac{B^* + E^*}{\theta}\right)\theta e^{-\theta x} + \left(\frac{C^* + F^*}{\theta^{\alpha+1}}\right)\frac{\theta^{\alpha+1}x^\alpha e^{-\theta x}}{\gamma(\alpha + 1)}. \end{aligned}$$

Renomeando os pesos para simplificar a expressão dada acima, obtém-se que a função densidade de probabilidade aproximada dos erros é dada por

$$f_{\varepsilon_t}(x) = A\frac{(\theta + \gamma)e^{-\frac{(\theta+\gamma)x}{\rho}}}{\rho} + B\theta e^{-\theta x} + C\frac{\theta^{\alpha+1}x^\alpha e^{-\theta x}}{\Gamma(\alpha + 1)},$$

em que

$$A = \frac{A^* + D^*}{\theta + \gamma},$$
$$B = \frac{B^* + E^*}{\theta}$$

e

$$C = \frac{C^* + F^*}{\theta^{\alpha+1}}.$$

APÊNDICE B – Método de estimação dos mínimos quadrados condicionais

Para encontrar os estimadores dos parâmetros do processo LGAR(1) pelo método dos mínimos quadrados condicionais, precisa-se solucionar o sistema de equações $\left(\frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \rho}, \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \alpha}, \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \theta}, \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \gamma}\right)' = (0, 0, 0, 0)'$ em que

$$Q_n(\rho, \alpha, \theta, \gamma) = \sum_{t=2}^n \left[X_t - \rho X_{t-1} + \frac{(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)}{\theta(\theta + \gamma)} \right]^2. \quad (\text{B.1})$$

Dessa forma, as derivadas parciais de primeira ordem da equação B.1 são dadas por:

$$\begin{aligned} \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \rho} &= 2 \sum_{t=2}^n \left[X_t - \rho X_{t-1} + \frac{(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)}{\theta(\theta + \gamma)} \right] \\ &\quad \times \left[\frac{\theta\alpha + \alpha\gamma + \gamma}{\theta(\theta + \gamma)} - X_{t-1} \right], \\ \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \alpha} &= 2 \sum_{t=2}^n \left[X_t - \rho X_{t-1} + \frac{(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)}{\theta(\theta + \gamma)} \right] \frac{(\rho - 1)}{\theta}, \\ \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \theta} &= 2 \sum_{t=2}^n \left[X_t - \rho X_{t-1} + \frac{(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)}{\theta(\theta + \gamma)} \right] \\ &\quad \times \left\{ \frac{(\rho - 1)[\alpha(\theta^2 + \theta\gamma) - (\theta\alpha + \alpha\gamma + \gamma)(2\theta + \gamma)]}{\theta^2(\theta + \gamma)^2} \right\} \text{ e} \\ \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \gamma} &= 2 \sum_{t=2}^n \left[X_t - \rho X_{t-1} + \frac{(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)}{\theta(\theta + \gamma)} \right] \\ &\quad \times \left\{ \frac{(\rho - 1)[(\alpha + 1)(\theta^2 + \theta\gamma) - \theta(\theta\alpha + \alpha\gamma + \gamma)]}{\theta^2(\theta + \gamma)^2} \right\}. \end{aligned}$$

Igualando a zero a derivada parcial de primeira ordem de $Q_n(\rho, \alpha, \theta, \gamma)$ com relação a ρ e resolvendo para ρ , tem-se que

$$\hat{\rho} = \frac{\sum_{t=2}^n \left\{ \left[X_t - \frac{(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})}{\hat{\theta}(\hat{\theta} + \hat{\gamma})} \right] \left[\frac{(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})}{\hat{\theta}(\hat{\theta} + \hat{\gamma})} - X_{t-1} \right] \right\}}{\sum_{t=2}^n \left[X_{t-1} - \frac{(\hat{\theta}\hat{\alpha} + \hat{\alpha}\hat{\gamma} + \hat{\gamma})}{\hat{\theta}(\hat{\theta} + \hat{\gamma})} \right]^2}.$$

Fazendo $\frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \alpha} = 0$, obtém-se que

$$\hat{\alpha} = \frac{(\hat{\rho} - 1) \sum_{t=2}^n (X_t - \hat{\rho} X_{t-1})}{\hat{\theta}(n - 1)} + \frac{\hat{\gamma}(\hat{\rho} - 1)^2}{\hat{\theta}^2(\hat{\theta} + \hat{\gamma})}.$$

Ao resolver a quarta equação do sistema, encontra-se

$$\hat{\gamma} = \frac{-\hat{\theta}^2 \sum_{t=2}^n X_t + \hat{\rho}\hat{\theta}^2 \sum_{t=2}^n X_{t-1} - \hat{\theta}\hat{\alpha}(\hat{\rho} - 1)(n - 1)}{\hat{\theta} \sum_{t=2}^n X_t - \hat{\rho}\hat{\theta} \sum_{t=2}^n X_{t-1} + (\hat{\rho} - 1)(\hat{\alpha} + 1)(n - 1)}.$$

Para a terceira equação, tem-se que

$$\begin{aligned} \frac{\partial Q_n(\rho, \alpha, \theta, \gamma)}{\partial \theta} = & 2[\theta(\theta + \gamma)]^{-3} \sum_{t=2}^n [\alpha\theta^2(\theta + \gamma)^2(\rho - 1)X_t \\ & - \theta(\theta + \gamma)(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)(2\theta + \gamma)X_t \\ & - \rho\alpha\theta^2(\theta + \gamma)^2(\rho - 1)X_{t-1} \\ & + \rho\theta(\theta + \gamma)(\rho - 1)(\theta\alpha + \alpha\gamma + \gamma)(2\theta + \gamma)X_{t-1} \\ & + \alpha\theta(\theta + \gamma)(\rho - 1)^2(\theta\alpha + \alpha\gamma + \gamma) \\ & - (2\theta + \gamma)(\rho - 1)^2(\theta\alpha + \alpha\gamma + \gamma)^2], \end{aligned}$$

mas não é possível isolar θ em função dos demais parâmetros.

APÊNDICE C – Transformada de Laplace condicional de X_{t+k} dado X_{t-1}

Neste apêndice, consta a demonstração do Corolário 1 apresentada no capítulo 3.

Primeiramente, sabe-se que

$$\begin{aligned}
 \phi_{X_{t+k}|X_{t-1}=x_{t-1}}(s) &= \mathbb{E}[\exp(-sX_{t+k})|X_{t-1} = x_{t-1}] \\
 &= \mathbb{E} \left[\exp \left(-s \left(\rho^{k+1} X_{t-1} + \sum_{i=0}^k \rho^{k-i} \varepsilon_{t+i} \right) \right) | X_{t-1} = x_{t-1} \right] \\
 &= \mathbb{E} \left[\exp \left(-s \rho^{k+1} x_{t-1} \right) \exp \left(-s \sum_{i=0}^k \rho^{k-i} \varepsilon_{t+i} \right) \right] \\
 &= \exp \left(-s \rho^{k+1} x_{t-1} \right) \prod_{i=0}^k \mathbb{E} \left[\exp \left(-s \rho^{k-i} \varepsilon_{t+i} \right) \right] \\
 &= \exp \left(-s \rho^{k+1} x_{t-1} \right) \prod_{i=0}^k \phi_{\varepsilon}(-s \rho^{k-i}) \\
 &= \exp \left(-s \rho^{k+1} x_{t-1} \right) \prod_{i=0}^k \left[\frac{\theta + \gamma + s \rho^{k-i}}{\theta + \gamma + s \rho^{k-i+1}} \left(\frac{\theta + s \rho^{k-i+1}}{\theta + s \rho^{k-i}} \right)^{\alpha+1} \right] \\
 &= \exp \left(-s \rho^{k+1} x_{t-1} \right) \left(\frac{\gamma + \theta + s}{\gamma + \theta + s \rho^{k+1}} \right) \left(\frac{\theta + s \rho^{k+1}}{\theta + s} \right)^{\alpha+1}.
 \end{aligned}$$

Portanto,

$$\phi_{X_{t+k}|X_{t-1}=x_{t-1}}(s) = \frac{\exp(-s \rho^{k+1} x_{t-1}) (\gamma + \theta + s) (\theta + s \rho^{k+1})^{\alpha+1}}{(\gamma + \theta + s \rho^{k+1}) (\theta + s)^{\alpha+1}}.$$