



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

ALICE BUARQUE VIEIRA DE MELLO

**A CONSTRUÇÃO DE UM NOVO PROCESSO AUTORREGRESSIVO DE  
PRIMEIRA ORDEM COM DISTRIBUIÇÃO MARGINAL GAMA-LINDLEY**

Recife

2021

ALICE BUARQUE VIEIRA DE MELLO

**A CONSTRUÇÃO DE UM NOVO PROCESSO AUTORREGRESSIVO DE PRIMEIRA ORDEM COM DISTRIBUIÇÃO MARGINAL GAMA-LINDLEY**

Dissertação apresentada 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 Matemática

**Orientadora:** Profa. Maria do Carmo Soares de Lima

**Coorientador:** Prof. Abraão David Costa do Nascimento

Recife

2021

Catálogo na fonte  
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

M527c Mello, Alice Buarque Vieira de  
A construção de um novo processo autorregressivo de primeira ordem com distribuição marginal Gama-Lindley / Alice Buarque Vieira de Mello. – 2021.  
65 f.: il., fig., tab.

Orientadora: Maria do Carmo Soares de Lima.  
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2021.  
Inclui referências e apêndices.

1. Estatística matemática. 2. Análise de séries temporais. I. Lima, Maria do Carmo Soares de (orientadora). II. Título.

519.5

CDD (23. ed.)

UFPE- CCEN 2021 - 32

**ALICE BUARQUE VIEIRA DE MELLO**

**A CONSTRUÇÃO DE UM NOVO PROCESSO AUTORREGRESSIVO DE  
PRIMEIRA ORDEM COM DISTRIBUIÇÃO MARGINAL GAMA- LINDLEY**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Mestre em Estatística.

Aprovada em: 22 de fevereiro de 2021.

**BANCA EXAMINADORA**

---

Prof.<sup>a</sup> Maria do Carmo Soares de Lima  
UFPE

---

Prof.<sup>o</sup> Getúlio José Amorim do Amaral  
UFPE

---

Prof.<sup>o</sup> Josimar Mendes de Vasconcelos  
UFRPE

## **AGRADECIMENTOS**

À minha família e amigos, por todo apoio e compreensão. Serei eternamente grata.

Aos meus orientadores, Profa. Maria do Carmo e Prof. Abraão, por todo apoio e incentivo que recebi. Gostaria de ressaltar minha admiração por ambos e minha gratidão pela confiança em mim depositada para a realização desse trabalho.

Ao meu colega Jairo que me acompanha desde o começo dessa jornada e que sempre está disposto a me ajudar. Obrigada pelas dúvidas esclarecidas e pelos valiosos conselhos.

A todos os professores e funcionários do Departamento de Estatística da UFPE pelo empenho e dedicação.

À CAPES, pelo apoio financeiro.

## RESUMO

Durante as últimas décadas foi dada grande atenção à proposta de novas distribuições. A necessidade da pesquisa na área de teoria de novas distribuições veio, dentre outros aspectos, do fato de encontrar um melhor ajuste a conjuntos de dados reais, que são modelados por distribuições conhecidas na literatura. Nesse contexto, a ordem natural da pesquisa é desenvolver o gerador de novas distribuições ou escolher um já proposto na literatura. Após essa etapa, em geral, o autor escolhe uma distribuição de base e insere no gerador suas funções densidade de probabilidade e distribuição acumulada. Além da aplicação a dados reais, os autores analisam a distribuição taxa de falha; desenvolvem propriedades matemáticas relevantes como: função quantílica, momentos e função geratriz de momentos; realizam o processo de estimação, no intuito de encontrar expressões para os estimadores em forma fechada, ou seja, expressões em função de funções matemáticas conhecidas na literatura e realizam estudos de simulação para verificar se a condição de consistência dos estimadores é válida. Esse é o processo natural na proposta de novas distribuições e/ou novos geradores. Com a quantidade relativamente grande de trabalhos na área, propor novos geradores/distribuições, tem se tornado complicado, uma vez que os revisores das revistas de publicação não veem mais a inovação no tema. Diante disso, o presente trabalho de dissertação traz um tema pouco trabalhado na literatura de novas distribuições: o seu contexto em séries temporais. A ideia aqui é tomar uma distribuição conhecida na literatura e escolher um processo estacionário a ser utilizado. Neste trabalho, escolhemos a distribuição Gama-Lindley (GaL) e o processo autorregressivo de ordem 1, AR(1). Usando o processo AR(1), escolhemos como saída a distribuição GaL e temos como objetivo encontrar a distribuição do erro. O novo processo é chamado GaLAR(1) e é mostrado que o erro do processo possui distribuição como sendo uma mistura das distribuições Delta de Dirac, Exponencial com parâmetro  $\lambda$ , Gama com vetor de parâmetros  $(2, \lambda)$  e Exponencial de parâmetro  $(\lambda\beta + \beta)/\rho$ . Providenciamos algumas propriedades estatísticas do novo processo: medidas estatísticas condicionais e distribuição conjunta. A estimação para os parâmetros do processo proposto foi realizada via método dos mínimos quadrados condicionais e método gaussiano. Os estimadores foram avaliados considerando seis cenários e tamanhos amostrais diferentes. Além disso, objetivando comparar o desempenho do processo proposto, consideramos seis processos autorregressivos de primeira ordem com distribuição marginal não-gaussiana conhecidos na literatura. As medidas utilizadas na comparação foram: Critério de Informação de Akaike (AIC), Critério de Informação Bayesiano (BIC) e a Raiz do Erro Quadrático Médio

de Previsão (RMSE). O banco de dados usado para ilustrar o desempenho e a competitividade do processo proposto, comparado com os citados anteriormente, é um conjunto de dados histórico. Os resultados obtidos mostram que o processo proposto é bastante competitivo e ganha para todos os outros.

**Palavras-chaves:** Aplicação a dados reais. Gerador de distribuições. Novas distribuições. Processos Estocásticos. Séries temporais.

## ABSTRACT

Over the past few decades, great attention has been paid to the proposal for new distributions. The need for research in the field of new distribution theory came, among other aspects, from the fact of finding a better fit to real data sets, which are modeled by distributions already known in the literature. In this context, the natural order of research is to develop the generator of new distributions or choose one proposed in the literature. After this step, in general, the author chooses a base distribution and inserts the functions of probability density and cumulative distribution into the generator. In addition to the application to real data, the authors also analyze the failure rate function; they develop relevant mathematical properties such as: quantile function, moments and moment generating function; perform the estimation process, in order to find expressions for the estimators in closed form, that is, expressions in function of mathematical functions known in the literature and perform simulation studies to verify if the condition of the estimators' consistency is valid. In short, this is the natural process in proposing new distributions and/or generators. However, with the relatively large amount of work in this area, proposing new generators/distributions, has become complicated, since the reviewers of publishing magazines no longer see innovation in the subject. In view of this, the present dissertation work brings a theme that has been little worked on in the literature of new distributions: its context in time series. The idea here is to take a distribution propose in the literature and choose a stationary process to be used. In this work, we chose the Gamma-Lindley distribution (GaL) and the autoregressive process of order 1, AR(1). In other words, using the AR(1) process, we choose the GaL distribution as an output and we aim to find the error distribution. The new process is called GaLAR(1) and it is shown that the process error has distribution as being a mixture of the Dirac Delta distributions, Exponential with parameter  $\lambda$ , Gamma with parameter vector  $(2, \lambda)$  and Exponential of parameter  $(\lambda\beta + \beta)/\rho$ . In addition, we provide some statistical properties of the new process: conditional statistical measures and joint distribution. The estimation for the proposed process parameters was performed using the conditional least squares method. The estimators were evaluated considering six different scenarios and different sample sizes. In this study, it was shown that the estimators are consistent, as expected. In addition, in order to compare the performance of the proposed process, we considered six first-order autoregressive processes with marginal non-Gaussian distribution known in the literature. The measures used in the comparison were: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and the Root Mean

Square Error (RMSE). The database used to illustrate the performance and competitiveness of the proposed process, compared to those mentioned above, is a water data set. The results obtained show that the proposed process is quite competitive and wins for all others.

**Keywords:** Application to real data. Distribution generator. New distributions. Stochastic Processes. Time series.

## LISTA DE FIGURAS

Figura 1 – Principais etapas do trabalho . . . . .	18
Figura 2 – (A) $\rho = 0.2$ , $\lambda = 1$ e $\beta = 1$ , (B) $\rho = 0.5$ , $\lambda = 1$ e $\beta = 1$ e (C) $\rho = 0.9$ , $\lambda = 1$ e $\beta = 1$ . . . . .	35
Figura 3 – (A) $\rho = 0.2$ , $\lambda = 1$ e $\beta = 1$ , (B) $\rho = 0.5$ , $\lambda = 1$ e $\beta = 1$ e (C) $\rho = 0.9$ , $\lambda = 1$ e $\beta = 1$ . . . . .	35
Figura 4 – Níveis anuais do lago Huron, em pés, entre os anos de 1875 e 1972 . . . . .	53
Figura 5 – FAC para os níveis anuais do lago Huron . . . . .	53
Figura 6 – FACP para os níveis anuais do lago Huron . . . . .	54
Figura 7 – Dados observados e valores preditos . . . . .	56

## LISTA DE TABELAS

Tabela 1 – Principais pacotes e funções utilizadas . . . . .	20
Tabela 2 – Estimativas e EQM para os parâmetros do modelo através do método de mínimos quadrados condicionais e método de aproximação gaussiana . . . .	41
Tabela 3 – Parâmetros estimados, AIC, BIC e RMSE para os dados observados . . . .	55

## LISTA DE ABREVIATURAS E SIGLAS

<b>AIC</b>	Critério de Informação de Akaike
<b>ARMA</b>	Autorregressivo e de Médias Móveis
<b>AR(1)</b>	Autorregressivo de ordem 1
<b>AST</b>	Análise de séries temporais
<b>BIC</b>	Critério de Informação Bayesiano
<b>EAR(1)</b>	Exponencial autorregressivo de ordem 1
<b>EQM</b>	Erro quadrático médio
<b>FAC</b>	Função de autocorrelação
<b>FACP</b>	Função de autocorrelação parcial
<b>fda</b>	Função de distribuição acumulada
<b>GaL</b>	Gama-Lindley
<b>GaLAR(1)</b>	Gama-Lindley autorregressivo de ordem 1
<b>GAR(1)</b>	Gama autorregressivo de ordem 1
<b>GLAR(1)</b>	Laplace generalizada autorregressivo de ordem 1
<b>IDE</b>	Integrated Development Environment
<b>i.i.d.</b>	Independentes e identicamente distribuídas
<b>INGAR(1)</b>	Inversa gaussiana autorregressivo de ordem 1
<b>L</b>	Lindley
<b>LAR(1)</b>	Lindley autorregressivo de ordem 1
<b>MG</b>	Método gaussiano
<b>MQC</b>	Mínimos quadrados condicionais
<b>NLAR(1)</b>	normal-Laplace autorregressivo de ordem 1

**RMSE** Raiz do erro quadrático médio de previsão

**SAR** Synthetic Aperture Radar

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>15</b>
1.1	DISCUSSÃO INICIAL E REVISÃO DA LITERATURA	15
1.2	PROPOSTA E OBJETIVOS DA DISSERTAÇÃO	17
1.3	ORGANIZAÇÃO DA DISSERTAÇÃO	19
1.4	PLATAFORMA COMPUTACIONAL UTILIZADA	19
<b>2</b>	<b>A DISTRIBUIÇÃO GAMA-LINDLEY</b>	<b>21</b>
2.1	INTRODUÇÃO	21
2.2	A DISTRIBUIÇÃO GAMA-LINDLEY	22
<b>3</b>	<b>A CONSTRUÇÃO DE UM NOVO PROCESSO AR(1)</b>	<b>26</b>
3.1	INTRODUÇÃO	26
3.2	DEFINIÇÕES	27
3.3	A CONSTRUÇÃO DE UM NOVO PROCESSO AR(1)	28
3.4	PROPRIEDADES ESTATÍSTICAS DO PROCESSO GALAR(1)	30
<b>3.4.1</b>	<b>Medidas estatísticas condicionais</b>	<b>30</b>
<b>3.4.2</b>	<b>Distribuição conjunta</b>	<b>32</b>
<b>3.4.3</b>	<b><i>The run probabilities</i></b>	<b>32</b>
3.5	FUNÇÃO DE AUTOCORRELAÇÃO E DENSIDADE ESPECTRAL	33
3.6	PREVISÃO	36
<b>4</b>	<b>ESTIMAÇÃO DO PROCESSO GALAR(1)</b>	<b>37</b>
4.1	INTRODUÇÃO	37
4.2	MÍNIMOS QUADRADOS CONDICIONAIS	37
4.3	MÉTODO DE ESTIMAÇÃO GAUSSIANO	39
4.4	RESULTADOS	40
<b>5</b>	<b>PROCESSOS AUTORREGRESSIVOS DE PRIMEIRA ORDEM COM DISTRIBUIÇÃO MARGINAL NÃO-GAUSSINA: UMA REVISÃO DA LITERATURA</b>	<b>43</b>
5.1	INTRODUÇÃO	43
5.2	PROCESSO EAR(1)	44
5.3	PROCESSO GAR(1)	45
5.4	PROCESSO INGAR(1)	45

5.5	PROCESSO NLAR(1)	46
5.6	PROCESSO GLAR(1)	47
5.7	PROCESSO LAR(1)	48
<b>6</b>	<b>APLICAÇÃO</b>	<b>50</b>
6.1	INTRODUÇÃO	50
6.2	AIC: CRITÉRIO DE INFORMAÇÃO DE AKAIKE	51
6.3	BIC: CRITÉRIO DE INFORMAÇÃO BAYESIANO	52
6.4	RMSE: RAÍZ DO ERRO QUADRÁTICO MÉDIO DE PREVISÃO	52
6.5	BANCO DE DADOS	53
<b>7</b>	<b>CONCLUSÃO</b>	<b>57</b>
7.1	CONSIDERAÇÕES FINAIS	57
7.2	TRABALHOS FUTUROS	57
	<b>REFERÊNCIAS</b>	<b>58</b>
	<b>APÊNDICE A – PROPOSIÇÃO 1</b>	<b>62</b>
	<b>APÊNDICE B – MÉTODO DE ESTIMAÇÃO DE MÍNIMOS QUADRADOS CONDICIONAIS</b>	<b>65</b>

# 1 INTRODUÇÃO

## 1.1 DISCUSSÃO INICIAL E REVISÃO DA LITERATURA

Na teoria de novas distribuições, é comum propor distribuições que modelam bem diversos tipos de conjuntos de dados. Na literatura, ao longo das últimas décadas, muitos autores usaram tal teoria para modelar conjuntos clínicos, de confiabilidade, da área de engenharia, tempo de falha, fadiga, dentre outros. Muitos dos dados usados são sem censura, mas há aqueles que também trabalham com dados censurados, o que faz mais sentido, quando o trabalho é voltado para a área de análise de sobrevivência.

Neste contexto, podemos citar muitos trabalhos como, por exemplo, os mencionados a seguir. CORDEIRO et. al (2020) propuseram um modelo de regressão baseado numa extensão da distribuição normal. Além de comparar o novo modelo com outros já existentes na literatura usando conjuntos de dados de fadiga e de tempo de vida, os autores avaliaram as mudanças na qualidade de vida relacionada à saúde bucal de crianças em idade escolar usando o modelo de regressão proposto e comparando-o com os modelos de regressão normal exponencializado, Lehmann Tipo II normal, normal e Gama normal. Para mais detalhes sobre o conjunto de dados de regressão, ver PAULA et. al (2012).

LIMA (2015) usou o gerador Gama, proposto por ZOGRAFOS e BALAKRISHNAN (2009), para propor cinco novas distribuições em seu trabalho de dissertação. No entanto, o trabalho teve seu foco na parte matemática, trazendo expressões para as propriedades de cada nova distribuição em forma de séries de potência, uma vez que o tratamento em forma fechada foi inviável. Apesar disso, a autora ainda mostrou a flexibilidade dos novos modelos a diversos conjuntos de dados reais (mas sem censura).

Analisando os trabalhos citados acima e muitos outros existentes na literatura, podemos dizer que o trabalho com novas distribuições tem sua estrutura muito bem definida e sem muitas inovações, em sua maioria. Ou seja, escolhe-se ou propõe-se um gerador; escolhe-se a distribuição de base ou baseline a ser inserida no gerador, propondo, assim, a nova distribuição; desenvolve-se as propriedades matemáticas (muitas vezes em função de expansões em séries de potência); faz-se a estimação por método de máxima verossimilhança, poucas vezes encontrando estimadores em forma fechada; realiza-se um estudo de simulação para verificar o critério de consistência dos estimadores; propõe-se um ou mais conjuntos de dados para mostrar que o novo modelo se ajusta bem a ele, quando comparado com outros modelos

conhecidos na literatura.

Apesar do descrito acima, alguns trabalhos trazem uma análise um pouco mais aprofundada ao propor uma nova distribuição. Como exemplo, podemos citar o trabalho realizado por CORDEIRO et. al (2019), em que os autores propõem um modelo de regressão para dados de tempo de vida censurado, chamado de Marshall-Olkin estendido Weibull flexível. Além de propor o modelo de regressão com duas componentes sistemáticas, os autores apresentam uma análise de diagnóstico baseada em influência global e resíduos.

Ainda nesse sentido, um outro trabalho que faz algo diferenciado é o de CORDEIRO et. al (2016). Os autores, além do usual, trazem em seu trabalho uma abordagem bayesiana para a estimação dos parâmetros do modelo proposto, chamado Gama Birnbaum-Saunders.

No entanto, mesmo havendo trabalhos como os dois últimos citados anteriormente, acreditamos que algo mais pode ser feito no que diz respeito ao contexto de novas distribuições. Diante disso, realizou-se uma pesquisa na literatura afim de verificar se haveria algum problema voltado ao contexto de séries temporais e novas distribuições.

TOMÉ e LATORRE (2001) abordaram séries temporais em Epidemiologia ao estudar tendências de mortalidade infantil através de dados históricos. GOÉS et al. (2012) investigaram uma série temporal de casos de leishmaniose visceral em um dado período. Em Medicina, podemos citar a análise da frequência cardíaca e sua variabilidade por MARÃES (2009); em Meteorologia, o estudo feito por GUEDES et al. (2019) analisa os índices pluviométricos anuais em determinada estação e região. Séries temporais podem aparecer também em Economia, como no estudo de séries financeiras, de taxas anuais de desemprego e históricos de vendas de determinado produto.

Os modelos em análise de séries temporais (AST) podem ser escolhidos com base na tratabilidade e na flexibilidade de suas distribuições marginais. Por muito tempo, uma série de dados temporais era assumida ser uma sequência gaussiana (ou normal). No entanto, em muitas situações, deparamo-nos com dados que apresentam comportamentos assimétricos não normais. Para contornar esse problema, nas décadas mais recentes, muitos trabalhos vêm sendo desenvolvidos com o objetivo de introduzir novos modelos não-gaussianos de séries temporais na literatura.

Grandes avanços têm sido feitos para processos gaussianos e para processos não-gaussianos condicionais, como na proposta de KEDEM e FOKIANOS (2005), que propõem processos tendo marginais pertencentes à família exponencial. Contudo, poucos avanços tem sido feitos para processos não condicionais tendo marginais não-gaussianas.

Podemos citar os modelos com distribuições marginais exponenciais estudados por: GAVER e LEWIS (1980) e LAWRENCE e LEWIS (1980). Processos com modelos marginais Gama foram propostos em GAVER e LEWIS (1980) e SIM (1986); POPOVIC e BAKOUCH (2016) propuseram processos com distribuição marginal Lindley.

Assim, poucos foram os trabalhos encontrados que fazem uso de teoria de novas distribuições em combinação com séries temporais. A próxima seção trará a proposta e os objetivos dessa dissertação.

## 1.2 PROPOSTA E OBJETIVOS DA DISSERTAÇÃO

Diante do exposto acima, podemos notar que a aplicabilidade de séries temporais é ampla e em diversas áreas do conhecimento. Em adição a isso, também sabemos que a grande preocupação em teoria de novas distribuições (ao menos na parte prática) é conseguir mostrar que o modelo proposto possui desempenho superior no ajuste a bancos de dados reais, quando comparado a outros modelos já propostos na literatura.

Logo, percebendo a possibilidade de trabalhar com algo nesse contexto, encontramos o trabalho de POPOVIC e BAKOUCH (2016), mencionado anteriormente, que propõe um novo processo autorregressivo de primeira ordem com distribuição marginal Lindley (LINDLEY, 1958). Os autores derivaram algumas propriedades estatísticas, usaram três métodos para estimar os parâmetros do processo e investigaram algumas propriedades dos estimadores através de alguns resultados numéricos. Além disso, através de dois conjuntos de dados, que são séries temporais, os autores mostraram que o processo proposto teve performance superior quando comparado a outros seis processos.

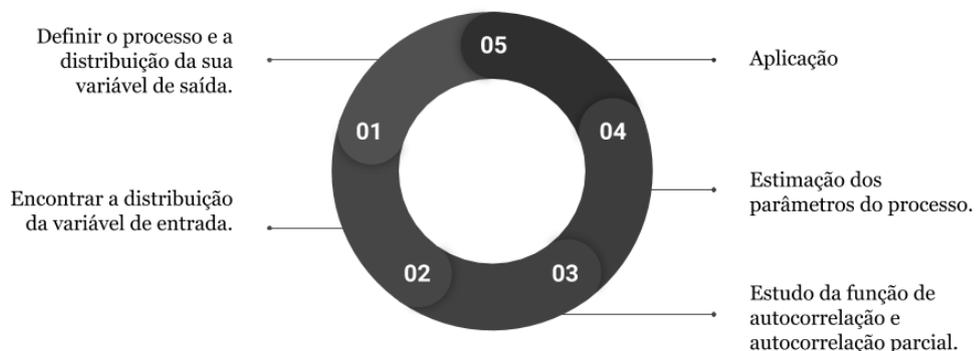
Assim, a partir da ideia proposta pelos autores supracitados e pensando em teoria de novas distribuições, essa dissertação tem como objetivo geral propor um novo processo autorregressivo de primeira ordem com marginal Gama-Lindley, como uma extensão ao processo proposto por POPOVIC e BAKOUCH (2016). Esse novo processo, chamado Gama-Lindley autorregressivo de ordem 1, será denotado por  $GaLAR(1)$ . Com o acréscimo de um parâmetro, espera-se que o modelo se torne mais flexível para a modelagem de dados reais.

Como objetivos específicos, listamos:

- Encontrar a distribuição para a sequência de inovações do processo com a distribuição marginal especificada;

- Encontrar a distribuição para a sequência de inovações como uma mistura de distribuições e seus respectivos pesos;
- Estudar as propriedades estruturais do novo processo;
- Desenvolver medidas estatísticas condicionais, como esperança e variância condicional para analisar o comportamento do modelo proposto;
- Discutir e analisar graficamente as funções de autocorrelação e autocorrelação parcial, com o intuito de ratificar a ordem do modelo apresentado;
- Realizar a estimação dos seus parâmetros através do método de mínimos quadrados condicionais e do método gaussiano;
- Realizar estudos numéricos de simulação, afim de verificar as propriedades assintóticas dos estimadores do processo proposto;
- Fazer uma análise da capacidade preditiva do novo modelo e comparar os resultados com os obtidos usando outros processos existentes na literatura;
- No intuito de mostrar a flexibilidade e desempenho do novo processo, realizar uma aplicação à uma série temporal.

Figura 1 – Principais etapas do trabalho



Fonte: A autora (2021)

### 1.3 ORGANIZAÇÃO DA DISSERTAÇÃO

No capítulo 2, iniciaremos com uma introdução à distribuição Lindley, seguida da definição de mistura de distribuições e revisão das principais propriedades da distribuição Gama-Lindley. No capítulo 3, definiremos alguns conceitos de AST. Construiremos um novo processo autor-regressivo de ordem 1 com marginal Gama-Lindley, denotado de GaLAR(1). Desenvolvemos algumas de suas propriedades estruturais. No capítulo 4, abordaremos dois métodos de estimação: mínimos quadrados condicionais e método de estimação gaussiano. Apresentamos formas semi-fechadas para os estimadores e um estudo de simulação para avaliar seus desempenhos. No capítulo 5, faremos uma revisão da literatura. Apresentaremos alguns processos AR(1) existentes na literatura de modelos não condicionais para séries temporais e discutiremos algumas propriedades estruturais de cada modelo. Analisaremos os possíveis desempenhos desses processos quando aplicados a um conjunto de dados reais de acordo com as características de suas estruturas. No capítulo 6, apresentaremos critérios de desempenho de modelos aplicados a conjunto de observações reais. Discutiremos os valores de AIC, BIC e RMSE dos modelos candidatos e do novo modelo quando aplicados a um conjunto de dados.

### 1.4 PLATAFORMA COMPUTACIONAL UTILIZADA

O sistema de tipografia adotado para preparar essa dissertação foi o  $\text{\LaTeX}$ . Amplamente utilizado na área acadêmica para preparação de documentos científicos, o  $\text{\LaTeX}$  foi inicialmente desenvolvido por Leslie Lamport na década de 80. Muitas informações encontradas em GRÄTZER (2016) foram utilizadas para a elaboração desse trabalho. O Ambiente de Desenvolvimento Integrado (IDE - *Integrated Development Environment*) utilizado para compilar a dissertação foi o TeXstudio (a versão mais recente pode ser encontrada em - <https://www.texstudio.org/>).

Para o desenvolvimento computacional do trabalho, como a estimação dos parâmetros para os processos discutidos, ajuste dos modelos a conjuntos de dados reais e a análise gráfica utilizamos o R. O R é um software livre para computação científica (<https://www.r-project.org/>). O IDE utilizado foi o RStudio (<https://rstudio.com/>). Os principais pacotes e funções utilizadas estão organizadas na Tabela 1.1. Ademais novas funções são implementadas para manipular o novo modelo de séries temporais.

Tabela 1 – Principais pacotes e funções utilizadas

---

Pacotes	Funções
stats	acf(), pacf(), rexp(), rgamma(), optim()
graphics	plot(), par()
base	sum(), lenght(), mean(), var(), sqrt(), summary(), rep(), set.seed(), return()
utils	data()

---

Fonte: A autora (2021)

## 2 A DISTRIBUIÇÃO GAMA-LINDLEY

### Resumo

Neste capítulo, apresentamos a distribuição Gama-Lindley (GaL) proposta por NEDJAR e ZEGHDOUDI (2015). Em seu trabalho, eles desenvolveram propriedades estatísticas estruturais e realizaram um estudo de simulação, além de ilustrar a aplicabilidade da distribuição por intermédio de dois diferentes conjuntos de dados. Aqui, discutimos algumas dessas propriedades que serão necessárias para o andamento dos capítulos seguintes.

*Palavras-chave:* Aplicação a conjunto de dados reais. Distribuição Gama-Lindley. Propriedades estatísticas estruturais.

### Abstract

In this chapter we present the Gamma-Lindley (GaL) distribution proposed by NEDJAR and ZEGHDOUDI (2015). In their work, they developed structural statistical properties and performed a simulation study, in addition to illustrating the applicability of the distribution through two different data sets. Here, we discuss some properties that will be necessary for the progress of the following chapters.

*Keywords:* Application to real data sets Gamma-Lindley distribution. Structural statistical properties.

### 2.1 INTRODUÇÃO

Com a crescente necessidade de se obter novas distribuições de probabilidade, esforços são direcionados para a escolha das *baselines* ou distribuições de base que serão utilizadas para a obtenção das novas distribuições, quando já se tem em mente o gerador a ser utilizado. Algumas características são comumente procuradas, como funções densidade de probabilidade e distribuições acumuladas que não apresentem expressões matemáticas complexas e formas fechadas para o desenvolvimento das principais propriedades estatísticas. Por forma fechada, entendemos expressões que envolvam funções matemáticas conhecidas.

Proposta por LINDLEY (1958), a distribuição Lindley é uniparamétrica e apresenta função densidade de probabilidade e distribuição acumulada dadas, respectivamente, por

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

e

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

em que  $\lambda > 0$  é um parâmetro de escala. A distribuição Lindley, denotada por  $L(\lambda)$ , é uma distribuição de probabilidade contínua com suporte em  $\mathbb{R}^+$ . GHITANY *et al.* (2008) abordaram um conjunto de propriedades para a distribuição Lindley e estudaram os muitos aspectos em que a Lindley é um modelo mais adequado para algumas aplicações quando comparada à distribuição exponencial.

Por ser uma distribuição que não apresenta nenhuma função matemática complicada e que contém apenas um parâmetro em sua forma, a distribuição Lindley está presente na literatura como componente ou distribuição base de outras distribuições: como nos estudos feitos por SANKARAN (1970) ao construir a distribuição Poisson-Lindley; na elaboração da distribuição Lindley através de distribuições Gama, feita nos trabalhos de ZEKERDAZEH e DOLATI (2009). BAKOUCH *et al.* (2012) obtiveram a distribuição Lindley estendida, fornecendo um modelo flexível para dados de vida útil. NEDJAR e ZEGHDOUDI (2015) desenvolveram e discutiram as propriedades da distribuição Gama-Lindley.

## 2.2 A DISTRIBUIÇÃO GAMA-LINDLEY

Algumas distribuições de probabilidade são resultado de combinações lineares de outras distribuições. Quando essas combinações lineares satisfazem determinadas condições, elas passam a ser denominadas de misturas. Esse método é comumente utilizado para a geração de novas distribuições.

**Definição 1** *Dizemos que uma função de distribuição acumulada (fda) é de uma mistura generalizada se ela corresponde a uma soma ponderada de outras fdas, tal que seus respectivos pesos somem 1. Cada distribuição individual é denominada componente da mistura e os seus pesos associados são os pesos da mistura.*

Modelos desse tipo preservam propriedades importantes, por exemplo, no caso de dispormos de componentes da mistura individualmente contínuas, a variável resultante também será contínua. Ainda, as densidades das variáveis de misturas generalizadas satisfazem as condições de uma função densidade de probabilidade.

A intenção é usar distribuições conhecidas na literatura para gerar uma nova, com a finalidade de alcançar uma distribuição mais flexível e adequada para solucionar um maior conjunto de problemas.

Nesse trabalho, optamos pela distribuição Gama-Lindley proposta por NEDJAR e ZEGHDOUDI (2015) para ser a distribuição da variável de saída do novo processo. Uma mistura da distribuição Gama(2, $\lambda$ ) e Lindley( $\lambda$ ) é usada para gerar a distribuição Gama-Lindley escolhida. Os autores estudaram propriedades estatísticas como a função quantílica, curva de Lorenz, densidade de probabilidade das estatísticas de ordem e entropia. Essa distribuição foi aplicada a dois conjuntos de observações referentes a tempos de falha, e se mostrou uma alternativa para descrever dados desse tipo.

Porém, é importante destacar que existe na literatura outra proposta de distribuição Gama-Lindley. LIMA (2015), propôs uma distribuição Gama-Lindley obtida através da família de distribuições gama-G (ZOGRAFOS e BALAKRISHNAN (2009), RISTIC e BALAKRISHNAN (2012)). Essa família gera novas distribuições em função das funções Gama e Gama incompleta. Nesse estudo, a distribuição Lindley uniparamétrica foi escolhida como distribuição base para ser aplicada à família de distribuições Gama-G. Essa proposta se mostrou, com os resultados apresentados, uma boa alternativa para descrever dados de tempo de falha e intensidades em canais de polarização extraídas de imagens SAR.

Pela simplicidade da Gama-Lindley desenvolvida por Nedjar e Zeghdoudi e pela forma de obtenção, mistura de outras distribuições, tema recorrente nesse trabalho, optamos por essa proposta para ser a distribuição da variável de saída do modelo proposto. A distribuição Gama-Lindley é obtida através de uma mistura das distribuições Gama(2, $\lambda$ ) e Lindley( $\lambda$ ), com  $\frac{\beta-1}{\beta}$  e  $\frac{1}{\beta}$ ,  $\beta > 0$ , seus respectivos pesos. A função densidade de probabilidade da GaL( $\lambda, \beta$ ), com  $x \in \mathbb{R}^+$ , é dada por

$$f(x; \lambda, \beta) = \frac{\lambda^2 e^{-x\lambda} [(\beta\lambda + \beta - \lambda)x + 1]}{\beta(1 + \lambda)},$$

em que  $\lambda > 0$  é parâmetro de escala e  $\beta$  é o parâmetro de forma. Note que, se  $\beta = 1$ , retornamos a distribuição Lindley( $\lambda$ ).

A função distribuição acumulada da GaL( $\lambda, \beta$ ) é dada por

$$F(x; \lambda, \beta) = 1 - \frac{e^{-x\lambda} [(\beta\lambda + \beta - \lambda)(\lambda x + 1) + \lambda]}{\beta(1 + \lambda)}.$$

Os autores supracitados também encontraram em seus estudos as expressões para o valor esperado e para a variância da distribuição. Ambos os resultados serão de grande importância para o desenvolvimento de algumas propriedades do nosso modelo.

O valor esperado e a variância de  $X \sim GaL(\lambda, \beta)$  são dados, respectivamente, por

$$\mathbb{E}(X) = \frac{2\beta(1 + \lambda) - \lambda}{\lambda\beta(1 + \lambda)}$$

e

$$\text{Var}(X) = \frac{-(-2\beta\lambda + \lambda)^2 + 2\beta^2(1 + 3\lambda) - 2\beta(\beta\lambda - 3\beta\lambda^2 + 2\lambda^2)}{(\lambda\beta(1 + \lambda))^2}. \quad (2.1)$$

Ademais, a função geradora de probabilidade e a função geradora de momentos são, na devida ordem, dadas por

$$\begin{aligned} P_X(t) = \mathbb{E}[t^X] &= \int_0^\infty t^x f(x) dx, \\ &= \frac{\lambda^2}{(\beta + \beta\lambda)(\lambda - \log(t))} \left( \frac{\beta + \beta\lambda - \lambda}{\lambda - \log(t)} + 1 \right), \quad t \in (0, 1) \end{aligned}$$

e

$$M_X(t) = \frac{\lambda^2}{(\beta + \beta\lambda)(\lambda - t)} \left( \frac{\beta + \beta\lambda - \lambda}{\lambda - t} + 1 \right).$$

A função geradora de momentos pode ser vista como um caso particular da função geradora de probabilidade, observe que  $M_X(t) = P_X(e^t)$ .

Um estudo de propriedades estruturais através das transformadas é realizado com o intuito de tornar mais tratável a solução de diversos problemas matemáticos e estatísticos. A transformada de Laplace é um artifício matemático facilitador para a resolução de problemas, proposta por Pierre-Simon Laplace, nos primeiros anos do século XIX. O método da transformada de Laplace ficou conhecido como um método capaz de simplificar a resolução de equações diferenciais.

A transformada de Laplace é uma transformação de integral largamente aplicada em diversos campos do conhecimento, incluindo fortemente Matemática, Estatística e Física. SUNIL KUMAR *et al.* (2015) propuseram um algoritmo para resolução da equação integral de Abel através da transformada de Laplace; HASAN MODIRSHANECHI e NASERPARI (2003) aplicaram a transformada de Laplace para resolver equações diferenciais não lineares no estudo de comportamentos de sistemas dinâmicos não lineares; METWALLY IA. (1999) utiliza a técnica de transformação de Laplace para modelar a resposta transitória de máquinas elétricas.

Atualmente, em muitos trabalhos, usamos a transformada de Laplace para obter os momentos e a função de uma distribuição de probabilidade. Seja  $\phi_X$  a transformada da variável aleatória  $X$ . Então

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

A transformada de Laplace é uma importante ferramenta para a construção de diversas propriedades do novo processo. Dada a densidade de  $X \sim \text{GaL}(\lambda, \beta)$ , a transformada de Laplace da distribuição é dada por

$$\begin{aligned}
 \phi_X(s) &= \frac{\lambda^2}{\beta(1+\lambda)} \phi_X \left[ x e^{-x\lambda} (\beta\lambda + \beta - \lambda) + e^{-x\lambda} \right], \\
 &= \frac{\lambda^2}{\beta(1+\lambda)} \left[ \frac{\beta\lambda + \beta - \lambda}{(s + \lambda)^2} + \frac{1}{s + \lambda} \right], \\
 &= \frac{\lambda^2(\beta\lambda + \beta + s)}{\beta(1+\lambda)(s + \lambda)^2}.
 \end{aligned} \tag{2.2}$$

A partir das definições apresentadas, será desenvolvido um novo processo autorregressivo de primeira ordem com distribuição marginal Gama-Lindley. Muitas propriedades do modelo proposto advém das propriedades da distribuição marginal geradora do processo.

### 3 A CONSTRUÇÃO DE UM NOVO PROCESSO AR(1)

#### **Resumo**

Motivados pela proposta de um novo modelo autorregressivo com distribuição marginal Lindley, propomos um novo processo autorregressivo com marginal Gama-Lindley. Definições importantes para o estudo de séries temporais são apresentadas. Além disso, desenvolvemos diversas propriedades estatísticas do processo proposto e ilustramos as funções de autocorrelação e autocorrelação parcial.

*Palavras-chave:* Distribuição Gama-Lindley. Distribuição marginal. Propriedades estatísticas. Séries temporais.

#### **Abstract**

Motivated by the proposal for a new autoregressive model with marginal Lindley distribution, we propose a new autoregressive process with marginal Gamma-Lindley. Important definitions for the study of time series are presented. Beside that, we develop several statistical properties of the proposed process and illustrate the autocorrelation and partial autocorrelation functions.

*Keywords:* Gamma-Lindley distribution. Marginal distribution. Statistical properties. Time series.

#### 3.1 INTRODUÇÃO

Uma série temporal é um conjunto de dados resultante da mensuração de uma variável em tempos igualmente espaçados. A partir dessas observações, diferentes investigações devem ser realizadas. A descrição comportamental da série pode ser feita através da construção gráfica dos dados observados, bem como das versões amostrais para as funções de autocorrelação e autocorrelação parcial. A investigação do mecanismo gerador é desenvolvida para se certificar e esclarecer o comportamento observado na análise gráfica, como, por exemplo, a presença de tendências, sazonalidades e ciclos, a identificação de possíveis periodicidades, a previsão e a descrição do comportamento futuro dos dados. Segundo MORETTIN e TOLOI (2006, p.1) "(...) há, basicamente dois enfoques usados na análise de séries temporais. Em ambos, o objetivo é construir modelos para as séries, com propósitos determinados. No primeiro enfoque, a análise é feita no domínio temporal e os modelos propostos são modelos paramétricos (com um número finito de parâmetros). No segundo, a análise é conduzida no domínio de frequências

e os modelos propostos são modelos não-paramétricos."

### 3.2 DEFINIÇÕES

Nesse trabalho, veremos as séries temporais como casos particulares de processos estocásticos. Formalmente, temos a seguinte definição

**Definição 2** *Dado um conjunto  $T$  arbitrário, um processo estocástico é uma família de variáveis aleatórias  $\{X(t), t \in T\}$  definidas no mesmo espaço de probabilidade.*

Com essa definição, podemos tomar as séries temporais como uma sequência de realizações de processos estocásticos.

Ao lidar com séries temporais, um conceito importante é o de estacionaridade (também conhecida com estacionaridade fraca), definido subsequentemente.

**Definição 3** *Uma série temporal  $\{X(t), t \in T\}$  é estacionária de segunda ordem (ou fracamente estacionária) se:*

$$(i) \mathbb{E}[X(t)] = \mu(t) = \mu, \text{ para todo } t \in T,$$

$$(ii) \mathbb{E}[X^2(t)] < \infty, \text{ para todo } t \in T,$$

$$(iii) \gamma_X(t_1, t_2) = \gamma_X(t_1 + \tau, t_2 + \tau), \text{ para todo } t_1, t_2, \tau \in T,$$

em que  $\gamma_X(t_1, t_2)$  representa a função de autocovariância de  $X_t$  nos pontos temporais  $t_1$  e  $t_2$ .

A presença de estacionaridade em séries temporais é algo buscado com grande frequência e é uma suposição presente na maioria das construções de novos modelos. Todavia, na prática, predominam as séries não-estacionárias, como os dados com tendências crescente ou decrescente.

Dentre os modelos estacionários para séries temporais, destacam-se as equações diferença ou processos autorregressivos. Definimos um processo autorregressivo de primeira ordem como um modelo que relaciona as variáveis  $x_t$  com os seus valores conforme a estrutura

$$x_t = \rho x_{t-1} + \epsilon_t, \tag{3.1}$$

em que  $x_t, \rho, \epsilon_t \in \mathbb{R}$ . A equação acima é denotada por  $x_t \sim AR(1) : \rho$ . Dizemos que  $x_t$  é a variável de saída e  $\epsilon_t$  a variável de entrada do modelo.

No estudo com estruturas autorregressivas, estamos interessados em quantificar o impacto recebido pelas variáveis de saída em função de mudanças nas variáveis de entrada. Esse efeito

pode ser quantificado pelo multiplicador dinâmico, definido como  $\psi_i = \partial x_{t+i} / \partial z_t$  (HAMILTON (1994)).

Note que para a Equação (3.1), valem-se as identidades:

$$\begin{cases} x_0 = \rho x_{-1} + \epsilon_0, \\ x_1 = \rho x_0 + \epsilon_1, \\ x_2 = \rho x_1 + \epsilon_2, \\ \vdots \\ x_t = \rho x_{t-1} + \epsilon_t \end{cases}$$

Por substituições recursivas nas equações diferenças acima, segue-se

$$x_t = \rho^{t+1} x_{-1} + \rho^t \epsilon_0 + \rho^{t-1} \epsilon_1 + \cdots + \rho \epsilon_{t-1} + \epsilon_t. \quad (3.2)$$

Note que sabido o valor inicial  $x_{-1}$  e os valores de  $\epsilon_t$ , para todo  $t$ , tem-se o valor da saída  $x_t$  para qualquer instante de tempo. Assim, como  $\epsilon_t$  tem propriedades analíticas mais tratáveis do que as propriedades das variáveis de saída, a Equação (3.2) se torna mais adequada matematicamente.

As próximas seções intencionam construir um processo autorregressivo de primeira ordem tendo a distribuição Gama-Lindley discutida no capítulo anterior como marginal com vistas a descrever fenômenos ederessados pela GaL considerando dependência.

### 3.3 A CONSTRUÇÃO DE UM NOVO PROCESSO AR(1)

Dizemos que  $X_t \sim AR(1)$  tendo um coeficiente autorregressivo  $\rho \in [0, 1)$  se

$$X_t = \rho X_{t-1} + \varepsilon_t, \quad (3.3)$$

em que  $\{\varepsilon_t\}$  é uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) e independentes de  $X_{t-\tau}$  para todo  $\tau \geq 1$ .

A construção de novos modelos não condicionais em Séries Temporais objetiva primeiramente derivar a distribuição de inovação (ou erros aleatórios), dada alguma marginal. Para obter a distribuição dos erros aleatórios, vamos analisar a transformada de Laplace da distribuição, por definição,  $\phi_X(s) = \mathbb{E}[\exp(-sX)]$ . Logo, para o processo autorregressivo em (3.3)

e tomando a transformada de Laplace (2.2), obtemos

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

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

Pela estacionaridade do processo,

$$\phi_{\varepsilon}(s) = \frac{\phi_X(s)}{\phi_X(\rho s)}.$$

Portanto,

$$\phi_{\varepsilon}(s) = \frac{(\beta\lambda + \beta + s)(\rho s + \lambda)^2}{(\beta\lambda + \beta + \rho s)(s + \lambda)^2}. \quad (3.4)$$

A seguinte proposição indica que há uma mistura aproximada de distribuições Gama na Expressão (3.4).

**Proposição 1** *A distribuição aproximada da sequência  $\{\varepsilon_t\}$  é uma mistura de uma distribuição contínua singular e uma absolutamente contínua dada por  $f_{\varepsilon}(x) \approx \rho\delta(x) + (1 - \rho)g(x)$ , em que  $\delta(x)$  é a função Delta de Dirac definida por*

$$\delta(x) = \begin{cases} \infty, & x = 0, \\ 0, & x \neq 0 \end{cases}$$

e  $g(x)$  é dada por

$$g(x) = \frac{A\lambda e^{-\lambda x} + B\lambda^2 x e^{-x\lambda} + C \left(\frac{\lambda\beta + \beta}{\rho}\right) e^{-\frac{x(\lambda\beta + \beta)}{\rho}}}{A + B + C}$$

em que

$$A = \frac{\rho(2\lambda - \beta\lambda - \beta)(2\beta\lambda + 2\beta - \lambda\rho) + \lambda(\beta\lambda + \beta)}{(\beta\lambda + \beta - \lambda\rho)^2}, \quad (3.5)$$

$$B = \frac{\lambda\rho - \lambda\beta\rho - \rho\beta - \lambda + \lambda\beta + \beta}{\beta\lambda + \beta - \lambda\rho} \quad (3.6)$$

e

$$C = \frac{\rho^2[(\beta\lambda + \beta)(2\lambda - \beta\lambda - \beta) - \lambda^2]}{(\beta\lambda + \beta - \lambda\rho)^2}. \quad (3.7)$$

A demonstraç o da Proposiç o 1 pode ser encontrada no Ap ndice A. Portanto, a distribuiç o da vari vel aleat ria  $\{\varepsilon_t\}$  pode ser entendida como uma mistura das distribuiç es Delta de Dirac, Exponencial( $\lambda$ ), Gama(2,  $\lambda$ ) e Exponencial( $\frac{\lambda\beta+\beta}{\rho}$ ).

Como consequ ncia da Proposiç o 1, o processo (3.3) pode ser escrito como segue

$$X_t = \begin{cases} \rho X_{t-1}, & \text{com probabilidade } \rho, \\ \rho X_{t-1} + \varepsilon_t, & \text{com probabilidade } 1 - \rho. \end{cases}$$

Ou equivalentemente,

$$X_t = \rho X_{t-1} + I_t \varepsilon_t, \quad (3.8)$$

em que  $I_t \sim \text{Bernoulli}(1 - \rho)$ . Denotaremos esse processo como  $X_t \sim \text{GaLAR}(1) : \rho$ .

### 3.4 PROPRIEDADES ESTATÍSTICAS DO PROCESSO GALAR(1)

Nessa seç o, vamos fornecer algumas propriedades estruturais capazes de descrever o processo proposto.

#### 3.4.1 Medidas estatísticas condicionais

Aqui, iremos estudar a esperanç a, a vari ncia condicional e a transformada de Laplace no horizonte do processo. Por substituiç o recursiva,  $X_{t+k}$  pode ser reescrita 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,

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

Como  $\{\varepsilon_t\}$    uma sequ ncia de vari veis i.i.d.,

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

Pela estacionaridade do processo (3.3),  $\mathbb{E}(\varepsilon_t) = \mathbb{E}(X_t)(1 - \rho)$ . Portanto,

$$\mathbb{E}(X_{t+k} | X_{t-1} = x_{t-1}) = \rho^{k+1} x_{t-1} + \frac{(1 - \rho^{k+1})(2\beta(1 + \lambda) - \lambda)}{\lambda\beta(1 + \lambda)}. \quad (3.9)$$

Note que quando  $k \rightarrow \infty$ ,  $\mathbb{E}(X_{t+k}|X_{t-1} = x_{t-1}) = \frac{(2\beta(1+\lambda)-\lambda)}{\lambda\beta(1+\lambda)}$ . Ou seja, quando  $k$  tende a valores muito grandes, a esperança condicional do processo  $k + 1$  passos a frente retorna a esperança da distribuição marginal Gama-Lindley.

De maneira semelhante, é possível encontrar a variância condicional. Aplicando a variância e suas propriedades, obtemos

$$\begin{aligned} \text{Var}(X_{t+k}|X_{t-1} = x_{t-1}) &= \rho^{2k}\text{Var}(\varepsilon_t) + \rho^{2(k-1)}\text{Var}(\varepsilon_{t+1}) + \cdots + \rho^2\text{Var}(\varepsilon_{t+k-1}) + \text{Var}(\varepsilon_{t+k}), \\ &= \text{Var}(\varepsilon_t) \sum_{i=0}^k \rho^{2i}, \\ &= \text{Var}(\varepsilon_t) \left( \frac{\rho^{2k+2} - 1}{\rho^2 - 1} \right). \end{aligned}$$

Pela estacionaridade do processo (3.3),  $\text{Var}(\varepsilon_t) = (1 - \rho^2)\text{Var}(X_t)$ . Logo,

$$\text{Var}(X_{t+k}|X_{t-1} = x_{t-1}) = (1 - \rho^{2k+2}) \left( \frac{-(-2\beta\lambda + \lambda)^2 + 2\beta^2(1 + 3\lambda) - 2\beta(\beta\lambda - 3\beta\lambda^2 + 2\lambda^2)}{(\lambda\beta(1 + \lambda))^2} \right) \quad (3.10)$$

Novamente, quando  $k \rightarrow \infty$ ,  $\text{Var}(X_{t+k}|X_{t-1} = x_{t-1}) = \text{Var}(X_t)$ . Além disso,  $\text{Var}(X_{t+k}|X_{t-1} = x_{t-1})$  é um valor constante independente de  $x_{t-1}$ .

Note que o valor esperado da variável aleatória condicional  $X_t$  dado  $X_{t-1} = x_{t-1}$  tem peso de magnitude  $\rho$  em  $\rho x_{t-1}$  e na seguinte, com probabilidade  $(1 - \rho)$ .

As medidas condicionais da variável aleatória  $X_{t+k}$  dado  $X_{t-1} = x$  são medidas assimétricas, ou seja,

$$\mathbb{E}(X_{t-1}|X_{t+k} = x) = \frac{x}{\rho^{k+1}} + \frac{(\rho^{k+1} - 1)(2\beta(1 + \lambda) - \lambda)}{\lambda\beta(1 + \lambda)\rho^{k+1}} \neq \mathbb{E}(X_{t+k}|X_{t-1} = x)$$

e

$$\begin{aligned} \text{Var}(X_{t-1}|X_{t+k} = x) &= \left( \frac{-(-2\beta\lambda + \lambda)^2 + 2\beta^2(1 + 3\lambda) - 2\beta(\beta\lambda - 3\beta\lambda^2 + 2\lambda^2)}{(\lambda\beta(1 + \lambda))^2} \right) \frac{\rho^{2k+2} - 1}{\rho^{2k+2}} \\ &\neq \text{Var}(X_{t+k}|X_{t-1} = x), \end{aligned}$$

agora, vamos calcular a transformada de Laplace condicional para o novo processo proposto:

$$\begin{aligned} \phi_{X_{t+k}|X_{t-1}=x_{t-1}}(s) &= E(\exp(-sX_{t+k})|X_{t-1} = x_{t-1}) \\ &= E \left\{ \exp[-s(\rho^{k+1}X_{t-1} + \sum_{i=0}^k \rho^{k-i}\varepsilon_{t+i})|X_{t-1} = x_{t-1}] \right\} \\ &= \exp(-s\rho^{k+1}x_{t-1}) \prod_{i=0}^k \frac{(\beta\lambda + \beta + s\rho^{k-i})(s\rho^{k-i+1} + \lambda)^2}{(\beta\lambda + \beta + s\rho^{k-i+1})(s\rho^{k-i} + \lambda)^2}. \end{aligned}$$

Portanto,

$$\phi_{X_{t+k}|X_{t-1}=x_{t-1}}(s) = \frac{\exp(-s\rho^{k+1}x_{t-1})(s\rho^{k+1} + \lambda)^2(\beta\lambda + \beta + s)}{(\beta\lambda + \beta + s\rho^{k+1})(s + \lambda)^2}.$$

Como esperado, quando  $k$  tende a um valor muito grande, a transformada de Laplace condicional do processo tende a transformada de Laplace da distribuição marginal Gama-Lindley.

### 3.4.2 Distribuição conjunta

A transformada de Laplace conjunta de  $(X_{t-1}, X_t)$  pode ser obtida através de

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

Pela independência das variáveis aleatórias

$$\phi_{X_{t-1}, X_t}(s_1, s_2) = \phi_{X_{t-1}}(s_1 + \rho s_2) \phi_{\varepsilon_t}(s_2).$$

Finalmente, pela estacionaridade do processo, obtemos

$$\phi_{X_{t-1}, X_t}(s_1, s_2) = \frac{\lambda^2(\beta\lambda + \beta + s_1 + \rho s_2)(\beta\lambda + \beta + s_2)(\rho s_2 + \lambda)^2}{(\beta\lambda + \beta)(s_1 + \rho s_2 + \lambda)^2(\beta\lambda + \beta + \rho s_2)(s_2 + \lambda)^2}.$$

Note que  $\phi_{X_{t-1}, X_t}(0, 0) = 1$  e a distribuição não é simétrica em  $X_t$  e  $X_{t-1}$ , dado que a transformada  $\phi_{X_t, X_{t-1}}(s_1, s_2)$  não é simétrica em  $s_1$  e  $s_2$ . Ou seja, o processo estudado não é reversível no tempo.

### 3.4.3 The run probabilities

Uma maneira simples de avaliar o comportamento de uma amostra do processo é definida por  $P(X_t < X_{t-1})$ . Essa medida é denominada *run probabilities*. Dada (3.8), obtemos

$$\begin{aligned} P(X_t < X_{t-1}) &= \rho P(\rho X_{t-1} < X_t) + (1 - \rho) P(\rho X_{t-1} + \varepsilon_t < X_{t-1}), \\ &= \rho + (1 - \rho) P(\varepsilon_t < (1 - \rho) X_{t-1}). \end{aligned}$$

Note que  $P(\rho X_{t-1} < X_t) = 1$  como efeito de  $\rho \in [0, 1)$ .

Vamos calcular  $P(\varepsilon_t < (1 - \rho)X_{t-1})$ . Primeiro, é sabido que

$$P(\varepsilon_t < (1 - \rho)X_{t-1}) = 1 - \int_0^\infty F_X\left(\frac{y}{1 - \rho}\right) f(y) dy.$$

$$\int_0^\infty F_X\left(\frac{y}{1 - \rho}\right) f(y) dy = \frac{\lambda(1 - \rho)}{A + B + C} \left[ \frac{A}{(\beta + \beta\lambda)(\rho - 2)} \left( \frac{\beta + \beta\lambda - \lambda}{\rho - 2} + \lambda \right) + \frac{B(1 - \rho)}{(\beta + \beta\lambda)(\rho - 2)^2} \left( \frac{2(\lambda - \beta - \beta\lambda)}{\rho - 2} + \lambda \right) + \frac{C\lambda(\rho\beta + \rho\beta\lambda - \rho\lambda + 1)}{\lambda\rho + (1 - \rho)(\beta\lambda + \beta)} \right].$$

Portanto, *the run probabilities* é dada por

$$P(X_t < X_{t-1}) = 1 - \frac{\lambda(1 - \rho)^2}{A + B + C} \left[ \frac{A}{(\beta + \beta\lambda)(\rho - 2)} \left( \frac{\beta + \beta\lambda - \lambda}{\rho - 2} + \lambda \right) + \frac{B(1 - \rho)}{(\beta + \beta\lambda)(\rho - 2)^2} \left( \frac{2(\lambda - \beta - \beta\lambda)}{\rho - 2} + \lambda \right) + \frac{C\lambda(\rho\beta + \rho\beta\lambda - \rho\lambda + 1)}{\lambda\rho + (1 - \rho)(\beta\lambda + \beta)} \right].$$

Em que, A, B e C são dados, respectivamente, por (3.5), (3.6) e (3.7).

### 3.5 FUNÇÃO DE AUTOCORRELAÇÃO E DENSIDADE ESPECTRAL

Uma importante característica presente na maioria das séries temporais é a dependência das observações próximas no tempo. Dada uma coleção de dados observados, podemos analisar a afinidade entre eles.

A função de autocovariância descreve a covariância da variável aleatória do processo com seus valores históricos, ou seja, fornece a covariância entre seu valor presente e seus valores defasados no tempo, enquanto que a função de autocorrelação (FAC) mede a correlação entre esses dados, isto é, determina como o valor atual da série estudada é influenciado pelos valores anteriores. A segunda função é muito utilizada na identificação de um modelo que seja o melhor para o ajuste e a modelagem de uma determinada série investigada.

Para modelos autorregressivos de primeira ordem, a função de autocorrelação é dada por

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

em que  $l$  representa o atraso ou defasagem. Dessa forma, a função de autocovariância é obtida através de

$$\gamma_{X_t}(l) = \rho^{|l|} \text{Var}(X_t). \quad (3.11)$$

Como consequência da estacionaridade do modelo, a correlação e a covariância entre dois valores do processo GaLAR(1) dependem apenas da defasagem.

A Figura 2 apresenta os gráficos da função de autocorrelação para amostras de tamanho 100 do processo GaLAR(1) para diferentes valores de  $\rho$ .

A FAC é uma função de autocorrelação completa, isto é, descreve como o valor presente da série está relacionado com seus valores passados. Em um processo autorregressivo de primeira ordem, a FAC apresenta um comportamento de decrescimento exponencial para zero ao passo que o atraso, o *lag*, aumenta. De fato, quanto mais distante uma observação está do valor atual, menos significativo é o seu impacto no referido. É evidente que quanto mais acentuado é o processo AR(1), ou seja, quanto maior é o valor de  $\rho$ , mais evidente fica esse comportamento.

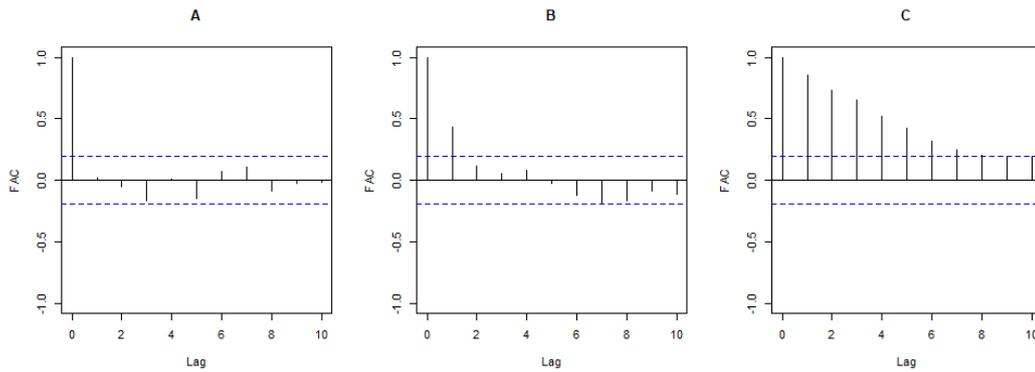
Outra ferramenta importante para a interpretação de uma série temporal é a função de autocorrelação parcial (FACP). Essa função descreve a correlação do valor atual  $X_t$  com o valor  $k$  passos à frente  $X_{t+k}$ , removendo as dependências das variáveis intermediárias,  $X_{t+1}, X_{t+2}, \dots, X_{t+k-1}$ .

A FACP pode ser vista como a correlação condicional  $\text{Cor}(X_{t+k}, X_t | X_{t+1}, \dots, X_{t+k-1})$ .

A Figura 3 apresenta os gráficos para função de autocorrelação parcial para amostras de tamanho 100 do processo GaLAR(1) nos mesmos cenários que estudamos a função de autocorrelação.

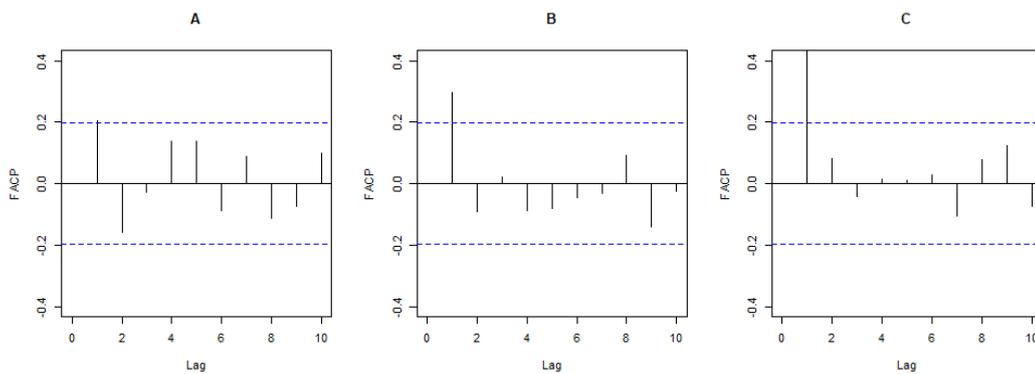
A FACP é muito utilizada na identificação da ordem de um modelo autorregressivo. Os valores de pico de uma FACP indicam em qual atraso se encontra o maior impacto para os demais valores da série. O comportamento ilustrado na Figura 3, nos três gráficos plotados, nos confirma que estamos trabalhando com um processo AR(1). Note que quanto mais próximo de 1 o valor de  $\rho$  está, mais evidente fica o comportamento de processo autorregressivo de primeira ordem.

Figura 2 – (A)  $\rho = 0.2$ ,  $\lambda = 1$  e  $\beta = 1$ , (B)  $\rho = 0.5$ ,  $\lambda = 1$  e  $\beta = 1$  e (C)  $\rho = 0.9$ ,  $\lambda = 1$  e  $\beta = 1$



Fonte: A autora (2021)

Figura 3 – (A)  $\rho = 0.2$ ,  $\lambda = 1$  e  $\beta = 1$ , (B)  $\rho = 0.5$ ,  $\lambda = 1$  e  $\beta = 1$  e (C)  $\rho = 0.9$ ,  $\lambda = 1$  e  $\beta = 1$



Fonte: A autora (2021)

Podemos estudar uma série temporal a partir de suas variações periódicas. O *power spectrum* (espectro de potência) ou densidade espectral de uma distribuição é um instrumento estatístico que contribui para a identificação de frequências, ciclos e periodicidades. Em outras palavras, a densidade espectral é uma ferramenta fundamental para o estudo de séries temporais do domínio das frequências. A densidade espectral pode ser definida através da função de autocovariância, como segue

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

De (3.11) e (2.1), obtemos

$$f_{XX}(w) = \frac{-(-2\beta\lambda + \lambda)^2 + 2\beta^2(1 + 3\lambda) - 2\beta(\beta\lambda - 3\beta\lambda^2 + 2\lambda^2)(1 - \rho^2)}{2\pi(\lambda\beta(1 + \lambda))^2(1 + \rho^2 - 2\rho \cos(w))}.$$

A densidade espectral decompõe a série em componentes periódicos. Dado que a densidade espectral pode ser vista como uma transformação da função de autocovariância, existe uma equivalência entre as mesmas, ou seja, ambas as estruturas contém as mesmas informações. O que as difere é a forma como escolhemos analisar a série estudada: a função de autocovariância apresenta resultados em termos de defasagens, ao passo que a distribuição espectral expressa o mesmo resultado em termos de ciclos. Essa escolha decorre do tipo de problema enfrentado, tomar a decisão de se trabalhar com o domínio do tempo ou das frequências. Isso pode ser determinado de acordo com as exigências do problema e as informações disponíveis.

### 3.6 PREVISÃO

O estudo de séries temporais pode ser dividido em duas ênfases. Podemos estudar dados no tempo a fim de entender e descrever o comportamento da série de observações ou podemos estudar o conjunto de dados com o objetivo de fazer previsões. É evidente que quanto mais compreendermos o comportamento dos dados estudados, mais seremos capazes de fazer boas previsões.

O estudo de previsão de uma série temporal consiste em analisar os possíveis acontecimentos futuros dos dados observados. Dado que conhecemos o conjunto de valores observados até o tempo  $t$ ,  $\{X_1, X_2, \dots, X_t\}$ , temos como objetivo fazer previsões para os valores de  $X_{t+h}$ , com  $h \geq 1$ . Dizemos que  $h$  é o horizonte de previsão e  $t$  é a origem das previsões. É importante destacar que a previsão de dados futuros é estimada a partir do que já foi observado.

Denotamos a melhor aproximação para o valor de  $X_{t+h}$  que minimiza o erro quadrático médio de  $\hat{X}_{t+h}$ . Ou seja, dizemos que  $\hat{X}_{t+h}$  é o preditor de erro quadrático médio mínimo para o valor do processo  $h$  passos a frente.

Temos que a melhor solução que minimiza  $\mathbb{E}[X_{t+h} - \hat{X}_{t+h}]^2$  é dada por

$$\hat{X}_{t+h} = \mathbb{E}[X_{t+h} | X_1, X_2, \dots, X_t].$$

Para um modelo autorregressivo de primeira ordem,  $AR(1)$ , segundo BOX & JENKINS (1970), o preditor para o processo  $h$  passos a frente é dado por

$$\hat{X}_{t+h} = \mu + \rho^h(X_t - \mu),$$

em que  $\mu$  é o valor esperado do processo e  $\rho$  o parâmetro autorregressivo.

## 4 ESTIMAÇÃO DO PROCESSO GALAR(1)

### Resumo

Nesse capítulo, apresentamos dois métodos de estimação para os parâmetros do processo proposto: mínimos quadrados condicionais e estimação por aproximação gaussiana. Para investigar o desempenho dos estimadores foi realizado um estudo de simulação com 10.000 réplicas de Monte Carlo.

*Palavras-chave:* Estimação. Método de Estimação gaussiano. Método dos Mínimos Quadrados Condicionais. Réplicas de Monte Carlo. Simulação.

### Abstract

In this chapter, we present two estimations methods for the parameters of the proposed process: conditional least squares and estimation by Gaussian approximation. To investigate the performance of the estimators a simulation study was performed with 10.000 Monte Carlo replicas.

*Keywords:* Estimation. Gaussian Estimation Method. Method of Conditional Least Squares. Replicas of Monte Carlo. Simulation.

### 4.1 INTRODUÇÃO

Visto que nos capítulos anteriores construímos algumas das propriedades estruturais de um novo processo autorregressivo de primeira ordem, nesse capítulo estudaremos os métodos de estimação mais adequados para o novo modelo proposto.

Abordaremos, então, os seguintes métodos de estimação: mínimos quadrados condicionais e estimação por aproximação gaussiana. Um estudo de simulação de Monte Carlo foi realizado para avaliar a qualidade dos estimadores.

### 4.2 MÍNIMOS QUADRADOS CONDICIONAIS

Dado um conjunto de observações  $X_t, t = 1, 2, \dots, n$ , podemos estimar os parâmetros do modelo proposto (3.3) pelo método de mínimos quadrados condicionais (KLIMKO e NELSON, 1978). Esse método de estimação é baseado na minimização da soma de desvios quadrados

sobre os valores esperados condicionais dos dados observados.

$$Q_n(\rho, \lambda, \beta) = \sum_{t=2}^n (X_t - E(X_t|X_{t-1}))^2. \quad (4.1)$$

Essa metodologia também foi utilizada na estimação dos parâmetros do processo autorregressivo Binomial de primeira ordem (ALWASEL *et al.*, 1998), do processo Lindley autorregressivo de ordem um (BAKOUCH e POPOVIC, 2016) e de um processo autorregressivo de primeira ordem com distribuição marginal Poisson-Lindley (MOHAMMADPOUR *et al.*, 2018). Estes estimadores se mostram fortemente consistentes e assintoticamente normais (KLIMKO e NELSON, 1978).

Portanto, a partir de (4.1), os estimadores de mínimos quadrados condicionais para os parâmetros  $\rho$ ,  $\lambda$  e  $\beta$  do processo GaLAR(1) são obtidos minimizando a função dada por

$$\begin{aligned} Q_n(\rho, \lambda, \beta) &= \sum_{t=2}^n (X_t - \rho X_{t-1} - (1 - \rho)\mathbb{E}(X_t))^2, \\ &= \sum_{t=2}^n \left( X_t - \rho X_{t-1} + \frac{(\rho - 1)(2\beta(1 + \lambda) - \lambda)}{\lambda\beta(1 + \lambda)} \right)^2. \end{aligned}$$

Os estimadores são obtidos com a resolução do sistema de equações  $\left( \frac{\partial Q_n}{\partial \rho}, \frac{\partial Q_n}{\partial \lambda}, \frac{\partial Q_n}{\partial \beta} \right) = (0, 0, 0)$ . Os valores de  $\frac{\partial Q_n}{\partial \rho}$ ,  $\frac{\partial Q_n}{\partial \lambda}$ ,  $\frac{\partial Q_n}{\partial \beta}$  podem ser encontrados no Apêndice B.

Dessa forma, obtivemos os seguintes resultados

$$\hat{\rho} = \frac{\sum_{t=2}^n \left\{ X_t X_{t-1} + \frac{2\hat{\beta}(\hat{\lambda}+1) - \hat{\lambda}}{\hat{\lambda}\hat{\beta}(\hat{\lambda}+1)} \left( -X_t - X_{t-1} + \frac{2\hat{\beta}(\hat{\lambda}+1) - \hat{\lambda}}{\hat{\lambda}\hat{\beta}(\hat{\lambda}+1)} \right) \right\}}{\sum_{t=2}^n \left\{ X_{t-1} - \frac{2\hat{\beta}(\hat{\lambda}+1) - \hat{\lambda}}{\hat{\lambda}\hat{\beta}(\hat{\lambda}+1)} \right\}^2},$$

$$\hat{\beta} = \frac{(1 - \hat{\rho})(n - 1)}{(1 + \hat{\lambda}) \sum_{t=2}^n \left( -X_t + \hat{\rho} X_{t-1} + \frac{2(1 - \hat{\rho})}{\hat{\lambda}} \right)}$$

e

$$2 \sum_{t=2}^n \left\{ X_t \frac{2(1 - \hat{\rho})}{\hat{\lambda}^2} + X_t \frac{\hat{\rho} - 1}{\hat{\beta}(1 + \hat{\lambda})^2} + X_{t-1} \frac{2\hat{\rho}(\hat{\rho} - 1)}{\hat{\lambda}^2} + X_{t-1} \frac{\hat{\rho}(1 - \hat{\rho})}{\hat{\beta}(1 + \hat{\lambda})^2} - \frac{4(1 - \hat{\rho})^2}{\hat{\lambda}^3} + \frac{2(\hat{\rho} - 1)^2}{\hat{\lambda}\hat{\beta}(1 + \hat{\lambda})^2} + \frac{2(1 - \hat{\rho})^2}{\hat{\lambda}^2\hat{\beta}(1 + \hat{\lambda})} - \frac{(1 - \hat{\rho})^2}{\hat{\beta}^2(1 + \hat{\lambda})^3} \right\} = 0.$$

Note que substituindo a segunda e a terceira identidade na primeira, obtêm-se uma equação não linear em termos de  $\hat{\rho}$ . Então, dada uma série, pode-se obter dela uma estimativa para  $\rho$ . Subsequentemente, aplicando-se este valor nas duas últimas equações, as estimativas  $\hat{\beta}$  e  $\hat{\lambda}$  são soluções do sistema 2 por 2 resultante. É interessante observar que esta estimação reduz o processo de estimação no processo GaLAR(1) à resolução de uma equação não linear em termos de  $\hat{\rho}$ .

### 4.3 MÉTODO DE ESTIMAÇÃO GAUSSIANO

O método de estimação gaussiano é uma ferramenta de aproximação para encontrar os estimadores dos parâmetros desconhecidos de um processo. Esse procedimento é usualmente recorrido quando existem dificuldades em avaliar a verossimilhança real do modelo. Esse método foi proposto por WHITTLE (1961). Em seu estudo, o autor considerou a verossimilhança da distribuição gaussiana como critério para as estimativas, sem atribuir a normalidade nos dados. Para tal, é necessário apenas considerar o valor esperado e a variância condicionais do modelo em estudo.

Essa metodologia foi empregada nos estudos de estimação do processo autorregressivo generalizado de Poisson de primeira ordem no trabalho de AL-NACHAWATI *et al.* (1997). ALWASEL *et al.* (1998) aplicou a estimação gaussiana para os parâmetros do processo autorregressivo binomial de ordem um. Esse método também foi usado para estimar os parâmetros do processo Lindley autorregressivo de primeira ordem proposto por POPOVIC e BAKOUCH (2016).

A função de verossimilhança condicional do modelo proposto nesse trabalho é dada por

$$L(\rho, \lambda, \beta) = \prod_{t=2}^n f(x_t|x_{t-1}),$$

em que  $f(x_t|x_{t-1})$  é a função de probabilidade condicional de  $X_t$  dado  $X_{t-1} = x_{t-1}$ . Substituindo  $f(x_t|x_{t-1})$  pela função densidade da distribuição normal, obtemos:

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

Assim,

$$\ell(\rho, \lambda, \beta) = \log L(\rho, \lambda, \beta) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \left[ \left( \frac{x_t - \mu_{x_t}}{\sigma_{x_t}} \right)^2 + \log(\sigma_{x_t}^2) \right],$$

em que  $\mu_{x_t} = \mathbb{E}(X_t|X_{t-1} = x_{t-1})$  e  $\sigma_{x_t}^2 = \text{Var}(X_t|X_{t-1} = x_{t-1})$  são definidas nas Equações (3.9) e (3.10), respectivamente, para  $k = 0$ .

Portanto,

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

Os estimadores podem ser obtidos resolvendo o sistema de equações  $\left( \frac{\partial \ell}{\partial \rho}, \frac{\partial \ell}{\partial \lambda}, \frac{\partial \ell}{\partial \beta} \right) = (0, 0, 0)$ . Devido a complexidade para as expressões das derivadas parciais correspondentes as entradas do sistema, as mesmas foram omitidas.

#### 4.4 RESULTADOS

Para avaliar a qualidade e a performance dos estimadores, foram realizadas simulações com 10.000 réplicas de Monte Carlo para diferentes tamanhos de amostras e configurações paramétricas.

Os estimadores foram avaliados em seis diferentes cenários, cada um com tamanhos de amostras 50, 100 e 500. Foram computados os valores da média, viés, viés relativo e erro quadrático médio (EQM) dos estimadores.

A Tabela 2 fornece a média e o erro quadrático médio (em parênteses) dos estimadores para os cenários  $\rho = \{0, 2; 0, 5; 0, 7; 0, 9\}$ ,  $\lambda = \{0, 4; 1, 0; 1, 6; 2, 0\}$  e  $\beta = \{0, 5; 1, 0\}$ . Aqui,  $\hat{\rho}_{MQC}$ ,  $\hat{\lambda}_{MQC}$  e  $\hat{\beta}_{MQC}$  representam as estimativas para os parâmetros  $\rho$ ,  $\lambda$  e  $\beta$  fornecidas pelo método de mínimos quadrados condicionais e  $\hat{\rho}_{MG}$ ,  $\hat{\lambda}_{MG}$  e  $\hat{\beta}_{MG}$  representam os valores estimados pelo método gaussiano, respectivamente.

Após algumas simulações, concluímos que os estimadores de mínimos quadrados condicionais são os mais eficientes para estimar os parâmetros do processo proposto com relação ao tempo de execução e por apresentar resultados numéricos mais satisfatórios (comportamento consistente dos estimadores) quando comparados aos estimadores obtidos pelo método gaussiano.

Tabela 2 – Estimativas e EQM para os parâmetros do modelo através do método de mínimos quadrados condicionais e método de aproximação gaussiana

Cenários	n	$\hat{\rho}_{MQC}$	$\hat{\lambda}_{MQC}$	$\hat{\beta}_{MQC}$	$\hat{\rho}_{MG}$	$\hat{\lambda}_{MG}$	$\hat{\beta}_{MG}$
(0,5; 1; 0,5)	50	0,4473	1,1181	0,4753	0,5503	1,2144	0,4361
		(0,0181)	(0,1763)	(0,2129)	(0,0034)	(0,0462)	(0,0040)
	100	0,4769	0,9802	0,4770	0,5469	1,2150	0,4365
		(0,0081)	(0,0369)	(0,0683)	(0,0024)	(0,0466)	(0,0040)
	500	0,4951	0,9828	0,4844	0,5452	1,2162	0,4366
		(0,0014)	(0,0164)	(0,0196)	(0,0020)	(0,0467)	(0,0040)
(0,7; 1; 0,5)	50	0,6368	0,9928	0,4597	0,5549	1,2175	0,4368
		(0,0163)	(0,0478)	(0,1763)	(0,0218)	(0,0473)	(0,0039)
	100	0,6684	0,9938	0,4944	0,5559	1,2178	0,4369
		(0,0066)	(0,0204)	(0,0131)	(0,0213)	(0,0474)	(0,0039)
	500	0,6941	0,9751	0,4817	0,9063	1,1022	0,4100
		(0,0010)	(0,0031)	(0,0022)	(0,0426)	(0,0100)	(0,0080)
(0,9; 1; 0,5)	50	0,8255	1,0637	0,5149	-	-	-
		(0,0123)	(0,4350)	(0,1958)	-	-	-
	100	0,8628	1,0062	0,5042	-	-	-
		(0,0040)	(0,0193)	(0,0145)	-	-	-
	500	0,8930	0,9959	0,4973	-	-	-
		(0,0004)	(0,0078)	(0,0080)	-	-	-
(0,2; 0,4; 1)	50	0,1673	0,4315	1,0203	0,9745	0,4348	0,8672
		(0,0194)	(0,0055)	(0,1908)	(0,6500)	(0,0109)	(0,5737)
	100	0,1829	0,4356	1,0750	0,9981	0,4237	0,7714
		(0,0098)	(0,0025)	(0,0436)	(0,6406)	(0,0087)	(0,4814)
	500	0,1970	0,4342	1,0585	-	-	-
		(0,0018)	(0,0014)	(0,0096)	-	-	-
(0,2; 1,6; 1)	50	0,1672	1,6994	1,0907	0,4816	1,8486	0,5436
		(0,0199)	(0,0459)	(0,0327)	(0,0823)	(0,0678)	(0,2083)
	100	0,1819	1,5768	0,9067	0,4901	1,8287	0,5410
		(0,0098)	(0,0290)	(0,0286)	(0,0851)	(0,0551)	(0,2106)
	500	0,1971	1,5864	0,9297	0,5003	1,8005	0,5373
		(0,0019)	(0,0052)	(0,0131)	(0,0906)	(0,0411)	(0,2140)
(0,2; 2; 1)	50	0,1673	1,9471	0,9847	0,4614	2,2359	0,5872
		(0,0196)	(0,0561)	(0,0445)	(0,0697)	(0,0560)	(0,1703)
	100	0,1830	2,0194	1,1243	0,4712	2,2382	0,5874
		(0,0097)	(0,0107)	(0,0442)	(0,0747)	(0,0568)	(0,1702)
	500	0,1973	2,0279	1,1353	-	-	-
		(0,0019)	(0,0020)	(0,0320)	-	-	-

Fonte: A autora (2021)

---

Os valores representados por - correspondem as estimativas que não conseguiram ser obtidas pelos método gaussiano.

A partir da Tabela 2 é possível concluir que os estimadores de mínimos quadrados condicionais possuem a melhor performance dado que para a maioria dos cenários analisados o EQM dos estimadores são menores quando comparados com os resultados fornecidos pelo método de estimação gaussiano. Ademais, para os cenários especificados, os estimadores de mínimos quadrados condicionais são consistentes: os valores médios das estimativas convergem para os valores verdadeiros dos parâmetros a medida que aumentamos o tamanho das amostras, e, da mesma maneira, os erros quadráticos médios tendem a zero quando o valor das amostras aumenta.

Em contrapartida, as estimativas obtidas pelo método gaussiano não apresentam o comportamento assintótico esperado: os estimadores não se mostraram consistentes nos cenários estudados. A escolha dos valores verdadeiros dos parâmetros pode influenciar diretamente na performance dos estimadores. Para a configuração paramétrica  $\rho = 0,9$ ,  $\lambda = 1$  e  $\beta = 0,5$ , o método gaussiano não gerou estimativas (mesmo para  $n = 50$  o tempo de execução foi muito elevado, não concluindo os resultados), da mesma maneira que não conseguiu realizar a estimação para  $n=500$  nos cenários:  $\rho = 0,2$ ,  $\lambda = 0,4$  e  $\beta = 1$  e  $\rho = 0,2$ ,  $\lambda = 2$  e  $\beta = 1$ . Uma investigação mais detalhada do espaço paramétrico precisa ser realizada para maior compreensão do comportamento dos estimadores.

## 5 PROCESSOS AUTORREGRESSIVOS DE PRIMEIRA ORDEM COM DISTRIBUIÇÃO MARGINAL NÃO-GAUSSINA: UMA REVISÃO DA LITERATURA

### Resumo

Nesse capítulo, vamos apresentar seis processos autorregressivos de primeira ordem com distribuição marginal não-gaussiana existentes na literatura de AST. Algumas propriedades estruturais como as funções densidade de probabilidade das distribuições marginais serão ilustradas. Os processos EAR(1), GAR(1), INGAR(1), NLAR(1), GLAR(1) e LAR(1) serão apresentados com o objetivo de comparar suas respectivas estruturas com a estrutura do novo processo proposto GaLAR(1).

*Palavras-chave:* Distribuições marginais. EAR(1). GAR(1). GLAR(1). INGAR(1). LAR(1). Literatura de AST. NLAR(1).

### Abstract

In this chapter, we will present six first-order autoregressive processes with marginal non-Gaussian distribution existing in the TSA literature. Some structural properties such as the probability density functions of the marginal distributions will be illustrated. The processes EAR (1), GAR (1), INGAR (1), NLAR (1), GLAR (1) and LAR (1) will be presented in with the aim to compare their respective structures with the structure of the proposed new process GaLAR (1).

*Keywords:* EAR(1). GAR(1). GLAR(1). INGAR(1). LAR(1). Marginal distributions. NLAR(1). TSA literature.

### 5.1 INTRODUÇÃO

O recurso tradicionalmente utilizado para a modelagem de séries temporais é baseado na suposição de que os dados observados são gaussianos, no entanto, a análise de dados reais não é limitada apenas a dados normais. Nas últimas décadas, processos autorregressivos com distribuições marginais não gaussianas têm sido desenvolvidos e acrescentados na literatura. O objetivo é ser capaz de modelar de forma eficiente uma série temporal sem precisar supor a normalidade das observações.

Nesse capítulo, estudaremos alguns processos autorregressivos de primeira ordem com marginais não-gaussianas existentes na literatura. Faremos uma análise das diferenças e semelhanças relevantes para a aplicabilidade dos modelos quando comparados com o processo

proposto, GaLAR(1). Os modelos EAR(1), GAR(1), INGAR(1), NLAR(1), GLAR(1) e LAR(1) serão discutidos e suas aplicabilidades ilustradas no capítulo 6.

## 5.2 PROCESSO EAR(1)

GAVER e LEWIS (1980) propuseram o modelo autorregressivo de primeira ordem com distribuição marginal exponencial (EAR(1)). Dada uma sequência estacionária  $\{X_t\}$  de variáveis aleatórias exponenciais com função densidade de probabilidade dada por

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0,$$

em que  $\lambda > 0$  é parâmetro de taxa, o processo EAR(1) é definido como

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

em que  $\{\varepsilon_t\}$  é uma sequência i.i.d. conhecida como sequência de inovação e  $\rho \in [0, 1)$  é o parâmetro autorregressivo do modelo.

A equação diferença para esse processo pode ser escrita da forma

$$X_t = \rho X_{t-1} + I_t E_t, \tag{5.1}$$

em que  $\{I_t\}$  é uma sequência i.i.d. de variáveis aleatórias *Bernoulli*( $1 - \rho$ ) e  $\{E_t\}$  é uma sequência de variáveis aleatórias exponenciais com parâmetro  $\lambda$  independentes. A utilidade da equação (5.1) é a forma como  $\{X_t\}$  é representada como uma combinação linear aleatória de sequências exponenciais independentes, facilitando o desenvolvimento computacional do processo.

O processo EAR(1) é formado por dois parâmetros: um da distribuição marginal e outro da estrutura autorregressiva. Se comparado ao modelo GaLAR(1) proposto nesse trabalho, o processo EAR(1) é composto por um número menor de parâmetros, o que influencia positivamente os resultados de possíveis ajustes a conjunto de dados observados. Devido a escolha dos critérios de desempenho que serão detalhados no capítulo 6, um número menor de parâmetros significa uma menor penalidade nos resultados obtidos.

Para aplicar o modelo a um conjunto de dados reais, estimamos os parâmetros do processo EAR(1) através da função de verossimilhança gaussiana dada por

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

### 5.3 PROCESSO GAR(1)

O processo Gama autorregressivo de primeira ordem, GAR(1), foi proposto por GAVER e LEWIS (1980). Esse modelo foi construído baseado na distribuição marginal Gama com função densidade de probabilidade dada por

$$f(x) = \frac{\lambda^\kappa x^{\kappa-1} e^{-\lambda x}}{\Gamma(\kappa)}, \quad x > 0,$$

em que  $\lambda \geq 0$  e  $\kappa \geq 0$  são parâmetros de taxa e forma, respectivamente, e  $\Gamma(\cdot)$  é a função Gama completa.

O processo GAR(1) é obtido através de

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

em que  $\{X_t\}$  é um processo estacionário com distribuição marginal Gama,  $\{\varepsilon_t\}$  é uma sequência de variáveis aleatórias i.i.d. e  $\rho \in [0, 1)$  é o parâmetro autorregressivo do modelo.

O processo GAR(1) e o processo proposto GaLAR(1) têm três parâmetros em sua estrutura. Essa característica estrutural em comum fornece uma comparação mais imparcial se pensarmos em ajuste a conjunto de observações reais, dado que toda vantagem ou desvantagem causada pela quantidade de parâmetros será a mesma para ambos os processos.

A função de verossimilhança gaussiana do processo GAR(1) é dada por

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

### 5.4 PROCESSO INGAR(1)

O processo autorregressivo de primeira ordem com distribuição marginal inversa gaussiana (INGAR(1)) foi introduzido por ABRAHAM e BALAKRISHNA (1998).

Seja  $X$  uma variável aleatória com distribuição inversa gaussiana, a função densidade de probabilidade de  $X$  é dada por

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp \left\{ -\frac{\lambda}{2\mu^2 x} (x - \mu)^2 \right\},$$

para  $x > 0$ ,  $\mu > 0$  média da distribuição e  $\lambda > 0$  parâmetro de forma.

Seja  $\{\varepsilon_t\}$  uma sequência i.i.d. de variáveis aleatórias. O desenvolvimento de novos modelos de séries temporais não gaussianas aborda encontrar a distribuição de  $\varepsilon_t$  tal que

$$X_t = \rho X_{t-1} + \varepsilon_t \tag{5.2}$$

represente um processo AR(1) para distribuição marginal especificada, com  $\rho \in [0, 1)$  parâmetro autorregressivo.

Considerando  $\{X_t\}$  uma sequência estacionária de variáveis aleatórias com distribuição marginal inversa gaussiana, o processo (5.2) é denominado INGAR(1).

A função de verossimilhança gaussiana para esse modelo é dada por

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

Definida a sua estrutura, assim como o processo GaLAR(1), o processo INGAR(1) é formado por três parâmetros, um da configuração autorregressiva e os demais fornecidos pela distribuição marginal. Discutiremos no capítulo 6 a capacidade de ajuste e predição do novo modelo proposto GaLAR(1) quando comparada as capacidades dos demais modelos apresentados nesse capítulo. Alguns dos critérios de desempenho levam em consideração a quantidade de parâmetros presente na estrutura do processo. Nesse quesito, o processo INGAR(1) e o processo GaLAR(1) serão beneficiados ou penalizados de maneira igual.

## 5.5 PROCESSO NLAR(1)

O processo autorregressivo de primeira ordem introduzido por JOSÉ *et al.* (2008) é baseado na distribuição marginal normal-Laplace.

Seja  $X$  uma variável aleatória normal-Laplace (REED, 2004), denotamos  $X \sim NL(v, \tau^2, \alpha, \beta)$ , a função densidade de probabilidade de  $X$  é dada por

$$f(x) = \frac{\alpha\beta}{\alpha + \beta} \phi \left( \frac{x-v}{\tau} \right) \left[ R \left( \alpha\tau - \frac{x-v}{\tau} \right) + R \left( \beta\tau + \frac{x-v}{\tau} \right) \right], \quad x \in \mathbb{R},$$

em que  $v \in \mathbb{R}$  é parâmetro de locação,  $\tau > 0$  parâmetro de escala e  $\alpha$  e  $\beta$  são parâmetros de forma, ambos positivos.  $\phi$  é a função densidade de probabilidade de uma variável aleatória normal padrão e  $R$  é a razão de Mills definida por

$$R(z) = \frac{1 - \Phi(z)}{\phi(z)},$$

em que  $\Phi$  é a função de distribuição acumulada de uma variável aleatória normal padrão.

O processo normal-Laplace autorregressivo de primeira ordem, NLAR(1), é definido por

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

em que  $\rho \in (0, 1)$  é o parâmetro autorregressivo,  $\{X_t\}$  processo estacionário com  $X_t \sim NL(v, \tau^2, \alpha, \beta)$  e  $\{\varepsilon_t\}$  sequência i.i.d. de variáveis aleatórias.

O modelo NLAR(1) é formado por cinco parâmetros. Comparado ao GaLAR(1), o processo com distribuição marginal normal-Lindley é um processo de maior complexidade. Isso significa que o desempenho de ajuste a conjunto de dados reais e a capacidade preditiva do processo tendem a apresentar melhores resultados devido a maior flexibilidade causada pela quantidade de parâmetros. Todavia, dependendo do critério de avaliação de desempenho utilizado, o número de parâmetros pode gerar uma penalidade nos resultados do processo.

Para calcular os critérios de desempenho do modelo, estimamos os parâmetros de máxima verossimilhança gaussiana. A função de verossimilhança do processo NLAR(1) é dada por

$$\begin{aligned} \ell(\rho, v, \tau^2, \alpha, \beta) &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{t=2}^n \frac{\left(x_t - \rho x_{t-1} - (1-\rho)\left(v + \frac{1}{\alpha} - \frac{1}{\beta}\right)\right)^2}{(1-\rho^2)\left(\tau^2 + \frac{1}{\alpha^2} + \frac{1}{\beta^2}\right)} \\ &+ \frac{1-n}{2} \log \left[ \left(\tau^2 + \frac{1}{\alpha^2} + \frac{1}{\beta^2}\right) (1-\rho^2) \right]. \end{aligned}$$

## 5.6 PROCESSO GLAR(1)

O processo autorregressivo de primeiro ordem com distribuição marginal Laplace generalizada foi introduzido por JOSÉ e THOMAS (2011).

O estudo tem como objetivo inicial encontrar a distribuição de  $\{\varepsilon_t\}$  sequência de inovação de variáveis aleatórias tal que o processo dado por

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

represente uma sequência estacionária com distribuição marginal Laplace generalizada, em que  $\rho \in (0, 1)$  é o parâmetro autorregressivo. Esse processo é denominado de processo Laplace generalizado autorregressivo de primeira ordem (GLAR(1)).

Dizemos que uma variável aleatória  $X$  segue a distribuição Laplace generalizada se  $X = \theta + \frac{\sigma}{\sqrt{2}} \left( \frac{1}{\kappa} G_1 - \kappa G_2 \right)$ , com  $G_i$ ,  $i = 1, 2$ , variáveis aleatórias com distribuição Gama uniparamétricas i.i.d. e função densidade de probabilidade dada por

$$f_{G_i} = \frac{1}{\Gamma(\tau)} e^{-x} x^{\tau-1}, \quad x > 0,$$

em que  $\tau > 0$  parâmetro de forma.

A variável aleatória  $X$  pode ser similarmente representada por  $X = \theta + \mu\omega + \sigma Z \sqrt{\Omega}$ , em que  $\Omega$  e  $Z$  são variáveis aleatórias com distribuição Gama( $\tau$ , 1) e normal padrão, respectivamente. Denotamos  $X \sim GL(\theta, \kappa, \sigma, \tau)$ .

A função de verossimilhança gaussiana do processo GLAR(1) definido é dada por

$$\begin{aligned} \ell(\rho, \theta, \kappa, \sigma, \tau) &= -\frac{n}{2} \log(2\pi) \\ &- \frac{1}{2} \sum_{t=2}^n \left[ \frac{\left( x_t - \rho x_{t-1} - (1-\rho) \left( \theta + \tau \frac{\sigma}{\sqrt{2}} \left( \frac{1}{\kappa} - \kappa \right) \right) \right)^2}{(1-\rho^2)\tau\sigma^2 \left( \frac{1}{2} \left( \frac{1}{\kappa} - \kappa \right)^2 + 1 \right)} + \log \left( (1-\rho^2)\tau\sigma^2 \left( \frac{1}{2} \left( \frac{1}{\kappa} - \kappa \right)^2 + 1 \right) \right) \right]. \end{aligned}$$

Assim como o modelo NLAR(1), o processo GLAR(1) apresenta cinco parâmetros em sua estrutura: o parâmetro autorregressivo do modelo e quatro parâmetros da distribuição marginal.

## 5.7 PROCESSO LAR(1)

O modelo Lindley autorregressivo de primeira ordem (LAR(1)) é um processo estacionário proposto por POPOVIC e BAKOUCH (2016). A construção do modelo LAR(1) é baseada na distribuição marginal Lindley, distribuição uniparamétrica com suporte  $\mathbb{R}_+^*$ .

Seja  $X$  uma variável aleatória Lindley( $\lambda$ ). A função densidade de probabilidade e a distribuição acumulada de  $X$  são dadas, respectivamente, por

$$f(x) = \frac{\lambda^2}{\lambda + 1} (1 + x) e^{-\lambda x}, \quad x > 0$$

e

$$F(x) = 1 - \frac{1 + \lambda(1 + x)}{\lambda + 1} (1 + x) e^{-\lambda x},$$

em que  $\lambda > 0$  é o parâmetro de escala.

Seja  $\{X_t\}$  um processo estacionário com distribuição marginal Lindley. O modelo LAR(1) é definido por

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

em que  $\rho \in [0, 1)$  é o parâmetro autorregressivo do processo e  $\{\varepsilon_t\}$  é uma sequência de variáveis aleatórias i.i.d. e independentes de  $X_{t-\tau}$ , para todo  $\tau \geq 1$ .

Portanto, o modelo LAR(1) é um processo que tem em sua estrutura dois parâmetros: um parâmetro referente a distribuição marginal Lindley e o parâmetro autorregressivo. Como veremos no próximo capítulo, podemos dizer que o modelo LAR(1) é um modelo menos complexo quando comparado ao processo GaLAR(1), que tem em sua estrutura um número maior de parâmetros.

Para estimar os parâmetros e ilustrar a aplicabilidade dos modelos candidatos, utilizamos o método de estimação gaussiano. A função de verossimilhança gaussiana para o modelo LAR(1) é dada por

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

## 6 APLICAÇÃO

### Resumo

Esse capítulo é voltado para a análise de desempenho do processo proposto quando aplicado a um conjunto de dados real. Introduzimos três critérios de avaliação para medir a capacidade de ajuste e predição do modelo GaLAR(1): Critério de Informação de Akaike (AIC), Critério de Informação Bayesiano (BIC) e Raiz do Erro Quadrático Médio de Previsão (RMSE). Os resultados obtidos serão comparados com os resultados fornecidos pelos seis modelos competidores introduzidos no capítulo 5, apresentados e debatidos. O conjunto de dados reais utilizado para avaliar o desempenho dos modelos será estudado e algumas propriedades serão discutidas.

*Palavras-chave:* Conjunto de dados reais. Critérios de avaliação. Critério de Informação de Akaike (AIC). Critério de Informação Bayesiano (BIC). Raiz do Erro Quadrático Médio de Previsão (RMSE).

### Abstract

This chapter is dedicated to performance analysis of the proposed process when applied to a real data set. We will introduce three evaluation criteria to measure the fit and prediction capacity of the GaLAR model (1): Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Root Mean Square Error (RMSE). The results obtained will be compared with the results provided by the six competing models introduced in chapter 5, presented and discussed. The set of real data used to evaluate the performance of the models will be studied and some properties will be discussed.

*Keywords:* Akaike Information Criterion (AIC). Bayesian Information Criterion (BIC). Evaluation criteria. Root Mean Square Error (RMSE). Real data set.

### 6.1 INTRODUÇÃO

Nesse capítulo, estudamos o desempenho do processo proposto, GaLAR(1), quando aplicado a um conjunto de dados reais. A performance do novo modelo é comparada com a performance de outros processos com características semelhantes existentes na literatura. Denominamos esses processos de modelos competidores ou modelos candidatos. Visto que temos como objetivo a construção de um novo processo, é importante avaliarmos a competitividade do novo modelo com outros modelos existentes na literatura, de forma que possamos garantir

que a nossa proposta seja relevante para a literatura de modelos em Séries Temporais.

Estamos interessados em avaliar o processo proposto dentro de um conjunto de modelos candidatos. Essa avaliação é feita através de algumas pontuações e levamos diversos fatores em consideração, como qualidade de ajuste, complexidade do modelo e capacidade preditiva. Entendemos por complexidade de modelo a quantidade de parâmetros em sua estrutura. Portanto, um modelo menos complexo é aquele com um menor número de parâmetros.

Para pontuar um modelo quanto ao seu desempenho em relação a um conjunto de dados, ponderamos a capacidade de ajuste e a estrutura do modelo. Em outras palavras, levamos em consideração o encaixe do modelo nos dados observados e a quantidade de parâmetros em sua forma. Estamos interessados em uma boa qualidade de ajuste realizada por modelos menos complexos. A capacidade preditiva pode ser vista como uma qualidade de ajuste dos dados.

Para a avaliação de desempenho dos modelos candidatos, calculamos dois critérios de informação estatísticos: AIC (Critério de Informação de Akaike) e BIC (Critério de Informação Bayesiano), e medimos a capacidade de predição dos competidores calculando a raiz do erro quadrático médio de previsão (RMSE).

## 6.2 AIC: CRITÉRIO DE INFORMAÇÃO DE AKAIKE

O Critério de Informação de Akaike (AIC) é uma medida de qualidade para modelos probabilísticos. O AIC é utilizado para pontuar a performance de um modelo aplicado a um conjunto de dados. Este critério é obtido através de

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

em que  $L(\cdot)$  é a função de verossimilhança do processo,  $\hat{B}$  é o vetor de parâmetros que maximiza  $L(\cdot)$  e  $p$  é o número de parâmetros do modelo.

Ao avaliarmos o AIC, esperamos obter o menor valor possível. O termo  $2p$  é visto como uma penalidade para os modelos que apresentam muitos parâmetros em sua estrutura. O AIC é uma medida de qualidade que valoriza um bom ajuste acompanhado de um modelo com o menor número de parâmetros possível.

### 6.3 BIC: CRITÉRIO DE INFORMAÇÃO BAYESIANO

O Critério de Informação Bayesiano (BIC) é uma medida de desempenho de modelos e atua de maneira semelhante ao AIC. O BIC é dado por

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

em que  $n$  é o número de observações do conjunto de dados ajustado.

Assim como vimos no AIC, o BIC penaliza modelos mais complexos, o que significa que também leva em consideração a qualidade do ajuste acompanhada de modelos formados por um menor número de parâmetros.

### 6.4 RMSE: RAÍZ DO ERRO QUADRÁTICO MÉDIO DE PREVISÃO

A previsão de dados de uma série temporal está diretamente relacionada ao ajuste de modelos baseado no próprio conjunto de observações. Procuramos descrever os dados da melhor maneira possível para obter as previsões mais adequadas para a série estudada.

Uma maneira de avaliar o desempenho de um modelo em relação a um conjunto de dados reais é quantificar a sua capacidade preditiva, ou seja, o quão próximo dos valores reais observados são os valores preditos pelo modelo.

Definimos o erro de previsão como a diferença entre o valor observado e o valor predito. Podemos calcular essa diferença para cada observação do conjunto estudado. O erro de previsão pode ser positivo ou negativo, mas o que estamos interessados em encontrar é uma média dos erros de previsão próxima de zero. Quanto mais próximo de zero for esse valor, mais próximas dos valores observados estão sendo as previsões, conseqüentemente melhor é a performance do modelo para os dados estudados.

Denotamos de raiz quadrada do erro quadrático médio de previsão (RMSE), para um conjunto de observações  $\{X_1, X_2, \dots, X_n\}$ , o valor dado por

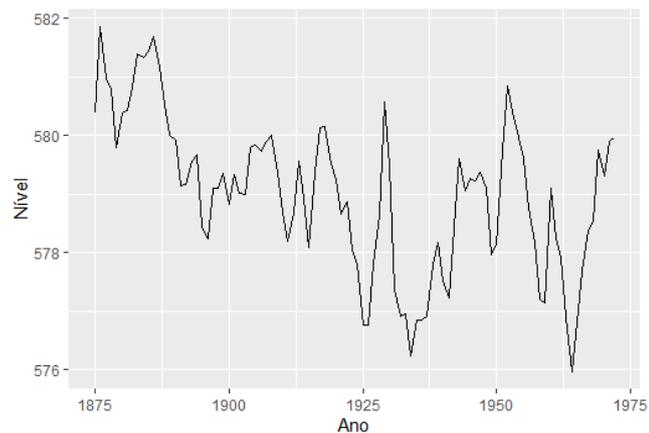
$$RMSE = \sqrt{\frac{\sum_{t=1}^n (x_t - \hat{x}_t)^2}{n}},$$

em que  $x_t$  e  $\hat{x}_t$  são, respectivamente, o valor observado e o valor predito no instante de tempo  $t$ .

## 6.5 BANCO DE DADOS

Para ilustrar o desempenho e a competitividade do modelo proposto, utilizamos um conjunto de dados referente às medições anuais do nível do lago Huron (em pés) localizado na América do Norte entre os anos de 1875 e 1972. Esse banco de dados foi amplamente estudado em BROCKWELL e DAVIS (1991).

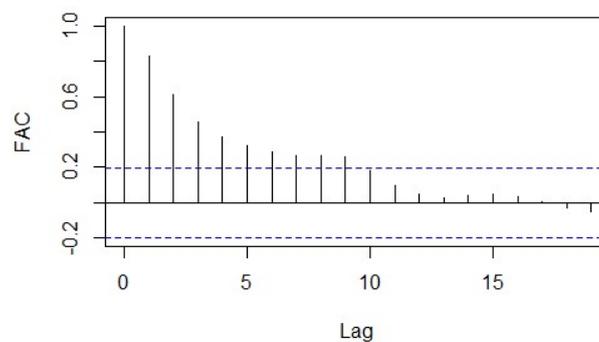
Figura 4 – Níveis anuais do lago Huron, em pés, entre os anos de 1875 e 1972



Fonte: A autora (2021)

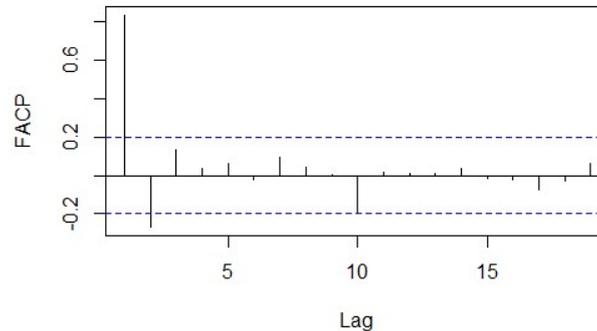
As Figuras 5 e 6 ilustram, respectivamente, a função de autocorrelação e autocorrelação parcial das observações. O decrescimento exponencial para zero a medida que o *lag* aumenta na FAC e o corte significativo após o primeiro *lag* na FACP indicam que modelos AR(1) podem ajustar bem os dados observados.

Figura 5 – FAC para os níveis anuais do lago Huron



Fonte: A autora (2021)

Figura 6 – FACP para os níveis anuais do lago Huron



Fonte: A autora (2021)

Para cada modelo candidato, calculamos os estimadores de máxima verossimilhança por aproximação gaussiana para obter os valores de AIC, BIC e RMSE. A Tabela 3 ilustra os resultados obtidos.

Podemos observar que os modelos selecionados se mostram bastante competitivos em relação ao ajuste e a capacidade preditiva para o conjunto de dados escolhido, não havendo grande diferença nos resultados obtidos. Todavia, o modelo proposto GaLAR(1) obteve os menores valores para AIC e BIC. Assim, ao ponderamos a qualidade de ajuste e a simplicidade do modelo, o novo processo gera os melhores resultados.

Ao comparar a capacidade preditiva dos modelos candidatos observamos uma maior competitividade. O desempenho de predição do modelo GaLAR(1) supera o desempenho da maioria dos candidatos, ficando atrás apenas da capacidade preditiva dos processos GLAR(1) e NLAR(1). Porém, o novo modelo tem em sua estrutura três parâmetros, dois a menos que os modelos GLAR(1) e NLAR(1). Portanto, pela maior simplicidade do novo modelo proposto, em termos de estimação, consideramos o modelo GaLAR(1) como uma boa alternativa para descrever nível de água em rios e barragens.

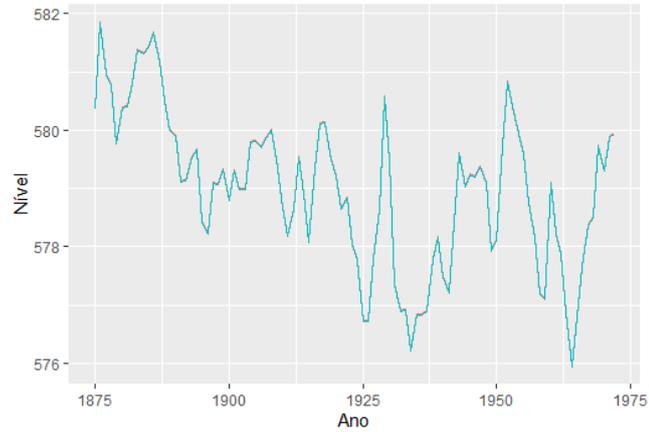
Tabela 3 – Parâmetros estimados, AIC, BIC e RMSE para os dados observados

Modelo	$\hat{\rho}$	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\kappa}$	$\hat{\mu}$	$\hat{\theta}$	$\hat{\alpha}$	$\hat{v}$	$\hat{\sigma}$	$\hat{\tau}$	AIC	BIC	RMSE
GaLAR	0,9999	0,0180	0,7544	-	-	-	-	-	-	-	<b>226,0966</b>	<b>233,8515</b>	<b>0,7453</b>
EAR	0,9986	0,0493	-	-	-	-	-	-	-	-	291,4317	296,6017	1,0546
GAR	0,9997	0,0298	-	0,9726	-	-	-	-	-	-	229,4350	237,1899	0,7547
INGAR	0,9989	0,0070	-	-	1,7566	-	-	-	-	-	289,5155	297,2704	0,9661
NLAR	0,9999	-	0,0080	-	-	-	1,3055	0,0522	-	1,6201	230,0609	242,9857	<b>0,7452</b>
GLAR	0,9999	-	-	0,0490	-	2,2960	-	-	4,6318	2,4264	230,0566	242,9815	<b>0,7451</b>
LAR	0,9996	0,0471	-	-	-	-	-	-	-	-	232,6822	237,8521	0,7735

Fonte: A autora (2021)

Podemos analisar a capacidade preditiva do processo GaLAR(1) a partir do Figura 7, em que o gráfico plotado em azul representa os dados observados e o gráfico plotado em vermelho representa os valores preditos pelo modelo proposto. Podemos concluir que o processo GaLAR(1) tem uma capacidade preditiva elevada para esse banco de dados estudados, ou seja, os valores previstos pelo processo foram muito próximos dos valores observados.

Figura 7 – Dados observados e valores preditos



Fonte: A autora (2021)

## 7 CONCLUSÃO

### 7.1 CONSIDERAÇÕES FINAIS

Nessa dissertação, apresentamos um novo processo autorregressivo de primeira ordem baseado na distribuição marginal Gama-Lindley. No primeiro momento, introduzimos a distribuição Gama-Lindley como uma mistura das distribuições  $\text{Gama}(2, \lambda)$  e  $\text{Lindley}(\lambda)$ , bem como algumas de suas propriedades foram desenvolvidas: expressões para valor esperado, variância, função geradora de probabilidade, função geradora de momentos e transformada de Laplace. Subsequentemente, propôs-se um novo processo de primeira ordem denominado por GaLAR(1). Uma distribuição para a sequência de inovações é proposta com base na transformada de Laplace. Algumas das propriedades matemáticas do processo GaLAR(1) foram derivadas: média e variância condicionais, distribuição conjunta para um par aleatório do processo, Run probability, função de autocorrelação e densidade espectral. Adicionalmente, as equações de previsão são discutidos. Foram propostos dois métodos de estimação, sendo um (baseado na mínimos quadrados condicionais) deles reduzido a resolução de apenas uma equação não linear. A performance dos estimadores é quantificada por um estudo de simulação Monte Carlo.

Finalmente, aplicamos o processo proposto a um conjunto de 98 observações referentes às medições anuais do nível do lago Huron em determinado período de tempo. Com base em três medidas de comparação, o modelo proposto apresentou uma melhor performance comparativamente a outros seis modelos utilizados na literatura.

### 7.2 TRABALHOS FUTUROS

Alguns pontos abordados nessa dissertação requerem estudos mais detalhados. Os tópicos a serem aprofundados são:

1. Otimizar o método de estimação gaussiano para atingir um tempo de execução satisfatório;
2. Investigar as configurações paramétrica que otimizem os estimadores;
3. Propor outros métodos de estimação para o modelo GaLAR(1);
4. Investigar outras aproximações para a sequencia de inovações;
5. Proposta do processo ARMA não condicional para a distribuição GaL.

## REFERÊNCIAS

- [1] Paula, J. S. de, de Oliveira, M., Soares, M. R. S. P., Chaves, M. das G. A. M., & Mialhe, F. L., "Perfil epidemiológico dos pacientes atendidos no Pronto Atendimento da Faculdade de Odontologia da Universidade Federal de Juiz de Fora," *Arquivos Em Odontologia*, 48(4), v. 48 n. 4, 2012
- [2] Zografos, K.; Balakrishnan, N. "On Families of beta- and Generalized Gamma-generated Distributions and Associated Inference," *Statistical Methodology*, v. 6, p. 344–362, 2009
- [3] Lima, Maria C.S.; Cordeiro, Gauss M. ; Ortega, Edwin M.M. ; Nascimento, Abraão D.C. "A new extended normal regression model: simulations and applications," *Journal of Statistical Distributions and Applications*, v. 6, p. 7, 2019.
- [4] Cordeiro, Gauss M. ; Pratavieira, Fábio ; Do Carmo S. Lima, Maria ; Ortega, Edwin M.M. "The Marshall-Olkin extended flexible Weibull regression model for censored lifetime data," *Model Assisted Statistics and Applications*, v. 14, p. 1-17, 2019.
- [5] Cordeiro, Gauss M. ; Lima, Maria do Carmo S. ; Cysneiros, Audrey H.M.A. ; Pascoa, Marcelino A. R. ; Pescim, Rodrigo R. ; Ortega, Edwin M. M. "An extended Birnbaum-Saunders distribution: Theory, estimation, and applications." *Communications in Statistics. Theory and Methods*, v. 45, p. 2268-2297, 2016.
- [6] E. A. Tomé and M. do Rosário Dias de Oliveira Latorre, "Tendências da mortalidade infantil no município de guarulhos: análise do período de 1971 a 1998," *Revista Brasileira de Epidemiologia*, vol. 4, pp. 153-167, nov 2001
- [7] M. A. de Oliveira Góes, C. M. de Melo, and V. de Lourdes Sierpe Jeraldo, "Série temporal da leishmaniose visceral em Aracaju, estado de sergipe, brasil (1999 a 2008): aspectos humanos e caninos," *Revista Brasileira de Epidemiologia*, vol. 15, pp. 298-307, jun 2012.
- [8] V. Marães, "Frequência cardíaca e sua variabilidade: análises e aplicações," *Revista an-*

---

*daluz de medicina del deporte, ISSN 1888-7546, N.º. 1, 2010, pags. 33-42, 2009.*

[9] H. A. S. Guedes, P. dos Santos Priebe, and E. B. Manke, "Tendências em séries temporais de precipitação no norte do estado do rio grande do sul, brasil," *Revista Brasileira de Meteorologia*, vol. 34, pp. 283-291, jun 2019

[10] F. Kedem, *Time Series Analysis*. John Wiley Sons, 2002.

[11] D. P. Gaver and P. A. W. Lewis, "First-order autoregressive gamma sequences and point processes," *Advances in Applied Probability*, vol. 12, pp. 727-745, sep 1980.

[12] A. J. Lawrance and P. A. W. Lewis, "A new autoregressive time series model in exponential variables (NEAR(1))," *Advances in Applied Probability*, vol. 13, pp. 826-845, dec 1981.

[13] C. H. Sim, "First-order autoregressive models for gamma and exponential processes," *Journal of Applied Probability*, vol. 27, pp. 325-332, jun 1990.

[14] H. S. Bakouch and B. V. Popovic, "Lindley first-order autoregressive model with applications," *Communications in Statistics - Theory and Methods*, vol. 45, pp. 4988-5006, jul 2016.

[15] G. Grätzer, *More Math Into LaTeX*. Springer International Publishing, 2016.

[16] D. V. Lindley, "Fiducial distributions and bayes' theorem," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1958.

[17] H. Zeghdoudi and S. Nedjar, "Gamma-lindley distribution and its application," *Journal of Applied Probability and Statistics*, 2014.

[18] M. Ghitany, B. Atieh, and S. Nadarajah, "Lindley distribution and its application," *Mathematics and Computers in Simulation*, vol. 78, pp. 493-506, aug 2008.

[19] M. Sankaran, "275. note: The discrete poisson-lindley distribution," *Biometrics*, vol. 26,

---

p. 145, mar 1970.

[20] H. Zakerzadeh and A. Dolati, "Generalized lindley distribution," *Journal of Mathematical Extension*, 2009.

[21] H. S. Bakouch, B. M. Al-Zahrani, A. A. Al-Shomrani, V. A. Marchi, and F. Louzada, "An extended lindley distribution," *Journal of the Korean Statistical Society*, vol. 41, pp. 75-85, mar 2012.

[22] Kumar, S.; Kumar, A.; Kumar, D.; Singh, J. & Singh, A. Analytical solution of Abel integral equation arising in astrophysics via Laplace transform, *Journal of the Egyptian Mathematical Society*, Springer Science and Business Media LLC, 2015

[23] Shanechi, H.; Pariz, N. & Vaahedi, E. General nonlinear modal representation of large scale power systems, *IEEE Transactions on Power Systems, Institute of Electrical and Electronics Engineers (IEEE)*, 2003.

[24] Metwally, I, Simulation of the impulse response of electrical machines, *IEEE Transactions on Energy Conversion, Institute of Electrical and Electronics Engineers (IEEE)*, 1999.

[25] P. Morettin, *Analise de series temporais*. Sao Paulo: Edgard Blucher, 2006.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

[26] G. E. P. B. Box and G. M. J. Gregory C. Reinsel, *Time Series Analysis 5e*. John Wiley Sons, 2015.

[27] L. A. Klimko and P. I. Nelson, "On conditional least squares estimation for stochastic processes," *The Annals of Statistics*, vol. 6, pp. 629-642, may 1978.

[28] I. Alwasel, A. Alzaid, and H. Al-Nachawati, "Estimating the parameters of the binomial autoregressive process of order one," *Applied Mathematics and Computation*, vol. 95, pp.

---

193-204, sep 1998.

[29] M. Mohammadpour, H. S. Bakouch, and M. Shirozhan, "Poisson-lindley INAR(1) model with applications," *Brazilian Journal of Probability and Statistics*, vol. 32, pp. 262-280, may 2018.

[30] P. Whittle, "Gaussian estimation in stationary time series," *Bulletin of the International Statistical Institute*, 1961.

[31] H. Al-Nachawati, I. Alwasel, and A. Alzaid, "Estimating the parameters of the generalized poisson AR(1) process," *Journal of Statistical Computation and Simulation*, vol. 56, pp. 337-352, mar 1997.

[32] B. Abraham and N. Balakrishna, "Inverse gaussian autoregressive models," *Journal of Time Series Analysis*, vol. 20, pp. 605-618, nov 1999.

[33] K. Jose, L. Tomy, and J. Sreekumar, "Autoregressive processes with normal-laplace marginals," *Statistics & Probability Letters*, vol. 78, pp. 2456-2462, oct 2008.

[34] Reed, W. J. "The Normal-Laplace Distribution and Its Relatives", *Statistics for Industry and Technology*, Birkhäuser Boston, , 61-74

[35] K. K. Jose and M. M. Thomas, "Generalized laplacian distributions and autoregressive processes," *Communications in Statistics - Theory and Methods*, vol. 40, pp. 4263-4277, dec 2011.

[36] P. Brockwell, *Introduction to time series and forecasting*. New York: Springer, 2002.

[37] J. Hamilton-Paterson, *Time Series Analysis*. Princeton Univers. Press.

## APÊNDICE A – PROPOSIÇÃO 1

Nesse apêndice vamos detalhar a obtenção do resultado exposto no Capítulo 3.

Dada a transformada de Laplace para o erro aleatório  $\varepsilon_t$ ,

$$\phi_\varepsilon(s) = \frac{(\beta\lambda + \beta + s)(\rho s + \lambda)^2}{(\beta\lambda + \beta + \rho s)(s + \lambda)^2}.$$

Podemos reescrevê-la da seguinte maneira,

$$\begin{aligned} \phi_\varepsilon(s) &= \frac{(\beta\lambda + \beta + s)(\lambda^2 + 2\lambda\rho s + \rho^2 s^2)}{(\beta\lambda + \beta + \rho s)(\lambda^2 + 2\lambda s + s^2)}, \\ &= \frac{\beta\lambda^2(\lambda + 1) + 2\beta\lambda\rho s(\lambda + 1) + \beta\rho^2 s^2(\lambda + 1) + s\lambda^2 + 2\lambda\rho s^2 + \rho^2 s^3}{\beta\lambda^2(\lambda + 1) + 2\lambda\beta s(\lambda + 1) + \beta s^2(\lambda + 1) + \rho\lambda^2 s + 2\rho\lambda s^2 + \rho s^3}. \end{aligned}$$

Manipulando nossa expressão encontrada acima,

$$\begin{aligned} \phi_\varepsilon(s) &= \frac{\rho^2 s^3 + 2\lambda\rho s^2(1 - \rho + \rho) + s\lambda^2(1 - \rho^2 + \rho^2) + \beta\rho^2 s^2(\lambda + 1)(1 - 1/\rho + 1/\rho)}{\beta\lambda^2(\lambda + 1) + 2\lambda\beta s(\lambda + 1) + \beta s^2(\lambda + 1) + \rho\lambda^2 s + 2\rho\lambda s^2 + \rho s^3} \\ &+ \frac{2\lambda\beta\rho s(\lambda + 1) + \beta\lambda^2(1 - \rho + \rho)(\lambda + 1)}{\beta\lambda^2(\lambda + 1) + 2\lambda\beta s(\lambda + 1) + \beta s^2(\lambda + 1) + \rho\lambda^2 s + 2\rho\lambda s^2 + \rho s^3}. \end{aligned}$$

Simplificando, obtemos

$$\phi_\varepsilon(s) = \rho + \frac{2\lambda\rho s^2(1 - \rho) + s\lambda^2(1 - \rho)(1 + \rho) + s^2\beta\rho^2(\lambda + 1)(1 - 1/\rho) + \beta\lambda^2(1 - \rho)(\lambda + 1)}{\beta\lambda^2(\lambda + 1) + 2\lambda\beta s(\lambda + 1) + \beta s^2(\lambda + 1) + \rho\lambda^2 s + 2\rho\lambda s^2 + \rho s^3}.$$

Após algumas manipulações algébricas chegamos no seguinte resultado:

$$\phi_\varepsilon(s) = \rho + (1 - \rho) \left\{ \frac{s^2[2\lambda\rho - \beta\rho(\lambda + 1)] + s\lambda^2(1 + \rho) + \beta\lambda^2(\lambda + 1)}{(s + \lambda)^2(\beta\lambda + \beta + s\rho)} \right\}.$$

Para obter a função que representa a distribuição do erro aleatório, aplicaremos na expressão acima a transformada inversa de Laplace. Para facilitar a notação, chamaremos de  $G(s)$  a expressão que acompanha o peso  $(1 - \rho)$ . Dessa maneira, denotamos  $g(x)$  a transformada inversa de Laplace correspondente. Com isso,

$$f_\varepsilon(x) = \rho\delta(x) + (1 - \rho)g(x).$$

Para encontrar a expressão para  $g(x)$ , escrevemos  $G(s)$  de forma que possamos decompor a expressão em frações parciais.

$$G(s) = \frac{\rho s^2(2\lambda - \beta\lambda - \beta)}{(s + \lambda)^2(\beta\lambda + \beta + s\rho)} + \frac{s\lambda^2(1 + \rho)}{(s + \lambda)^2(\beta\lambda + \beta + s\rho)} + \frac{\lambda^2\beta(\lambda + 1)}{(s + \lambda)^2(\beta\lambda + \beta + s\rho)}$$

Chamaremos os termos de  $G(s)$  de termos 1, 2 e 3, respectivamente. Seja:

## Termo 1

O mecanismo das frações parciais é muito utilizado nos métodos de integrações e nas transformadas inversas de Laplace. Esse método consiste em decompor uma função racional em termos mais acessíveis. Nosso problema se torna encontrar os coeficientes, A, B e C, como se segue:

$$\frac{\rho s^2(2\lambda - \beta\lambda - \beta)}{(s + \lambda)^2(\beta\lambda + \beta + s\rho)} = \frac{A^*}{s + \lambda} + \frac{B^*}{(s + \lambda)^2} + \frac{C^*}{\beta\lambda + \beta + s\rho}$$

Para  $s = -\lambda$ , encontramos

$$B^* = \frac{\lambda^2 \rho (2\lambda - \beta\lambda - \beta)}{\beta\lambda + \beta - \lambda\rho}.$$

Sendo  $s = \frac{-(\beta\lambda + \beta)}{\rho}$ ,

$$C^* = \frac{\rho(\beta\lambda + \beta)^2(2\lambda - \beta\lambda - \beta)}{(\beta\lambda + \beta - \lambda\rho)^2}.$$

Para  $s = 0$  e usando os coeficientes B e C encontrados,

$$A^* = \frac{-\rho\lambda(2\beta\lambda + 2\beta - \lambda\rho)(2\lambda - \beta\lambda - \beta)}{(\beta\lambda + \beta - \lambda\rho)^2}.$$

## Termo 2

De maneira análoga encontraremos os coeficientes da decomposição em frações parciais para o segundo termo da função  $G(s)$ .

$$\frac{s\lambda^2(1 + \rho)}{(s + \lambda)^2(\beta\lambda + \beta + s\rho)} = \frac{D}{s + \lambda} + \frac{E}{(s + \lambda)^2} + \frac{F}{\beta\lambda + \beta + s\rho}$$

Para  $s = -\lambda$ ,

$$E = \frac{-\lambda^3(1 + \rho)}{\beta\lambda + \beta - \lambda\rho}.$$

Dado  $s = \frac{-(\beta\lambda + \beta)}{\rho}$ ,

$$F = \frac{-\lambda^2 \rho (\beta\lambda + \beta)(1 + \rho)}{(\beta\lambda + \beta - \lambda\rho)^2}.$$

E para  $s = 0$ ,

$$D = \frac{\lambda^2(1 + \rho)(\beta\lambda + \beta)}{(\beta\lambda + \beta - \lambda\rho)^2}.$$

### Termo 3

Com o mesmo raciocínio vamos desenvolver o último termo de  $G(s)$ ,

$$\frac{\lambda^2 \beta (\lambda + 1)}{(s + \lambda)^2 (\beta \lambda + \beta + s \rho)} = \frac{G}{s + \lambda} + \frac{H}{(s + \lambda)^2} + \frac{I}{\beta \lambda + \beta + s \rho}.$$

Seja  $s = -\lambda$ ,

$$H = \frac{\beta \lambda^2 (\lambda + 1)}{\beta \lambda + \beta - \lambda \rho}.$$

Fazendo  $s = \frac{-(\beta \lambda + \beta)}{\rho}$ ,

$$I = \frac{\rho^2 \lambda^2 (\beta \lambda + \beta)}{(\beta \lambda + \beta - \lambda \rho)^2}.$$

E para  $s = 0$ ,

$$G = \frac{-\rho \lambda^2 \beta (\lambda + 1)}{(\beta \lambda + \beta - \lambda \rho)^2}.$$

Observe que a função  $G(s)$  pode ser escrita da seguinte maneira:

$$G(s) = \frac{A^* + D + G}{s + \lambda} + \frac{B^* + E + H}{(s + \lambda)^2} + \frac{C^* + F + I}{\beta \lambda + \beta + \rho s}$$

Note que,

$$A\lambda = A^* + D + G = \frac{\lambda \rho (2\lambda - \beta \lambda - \beta) (2\beta \lambda + 2\beta - \lambda \rho) + \lambda^2 (\beta \lambda + \beta)}{(\beta \lambda + \beta - \lambda \rho)^2},$$

$$B\lambda^2 = B^* + E + H = \frac{\lambda^3 \rho - \lambda^3 \beta \rho - \rho \beta \lambda^2 - \lambda^3 + \lambda^3 \beta + \beta \lambda^2}{\beta \lambda + \beta - \lambda \rho},$$

e

$$C \left( \frac{\beta \lambda + \beta}{\rho} \right) = C^* + F + I = \frac{\rho (\beta \lambda + \beta) [(\beta \lambda + \beta) (2\lambda - \beta \lambda - \beta) - \lambda^2]}{(\beta \lambda + \beta - \lambda \rho)^2}$$

Portanto,

$$G(s) = A \frac{\lambda}{s + \lambda} + B \frac{\lambda^2}{(s + \lambda)^2} + C \frac{\left( \frac{\beta \lambda + \beta}{\rho} \right)}{\beta \lambda + \beta - \lambda \rho}.$$

Dessa maneira, somos capazes de aplicar a transformada inversa de Laplace, dado que obtivemos funções conhecidas. Seja  $g(x)$  a transformada inversa de Laplace de  $G(s)$ , temos:

$$g(x) \approx \frac{A \lambda e^{-\lambda x} + B \lambda^2 x e^{-x \lambda} + C \left( \frac{\lambda \beta + \beta}{\rho} \right) e^{-\frac{x(\lambda \beta + \beta)}{\rho}}}{A + B + C}.$$

Note que para  $g(x)$  ser uma mistura das distribuições Exponenciais de parâmetros  $\lambda$  e  $\left( \frac{\lambda \beta + \beta}{\rho} \right)$  e Gama(2,  $\lambda$ ), é necessário que os pesos correspondentes a cada distribuição somem 1.

## APÊNDICE B – MÉTODO DE ESTIMAÇÃO DE MÍNIMOS QUADRADOS CONDICIONAIS

Minimizaremos a função  $Q_n(\rho, \lambda, \beta)$  resolvendo o sistema  $\left(\frac{\partial Q_n}{\partial \rho}, \frac{\partial Q_n}{\partial \lambda}, \frac{\partial Q_n}{\partial \beta}\right) = (0, 0, 0)$ .

Seja

$$Q_n(\rho, \lambda, \beta) = \sum_{t=2}^n \left( X_t - \rho X_{t-1} + \frac{(\rho - 1)(2\beta(1 + \lambda) - \lambda)}{\lambda\beta(1 + \lambda)} \right)^2.$$

As derivadas parciais são dadas por

$$\frac{\partial Q_n}{\partial \rho} = 2 \sum_{t=2}^n \left[ \left( X_t - \rho X_{t-1} + \frac{(\rho - 1)(2\beta(1 + \lambda) - \lambda)}{\lambda\beta(1 + \lambda)} \right) \left( -X_{t-1} + \frac{2\beta(1 + \lambda) - \lambda}{\lambda\beta(1 + \lambda)} \right) \right],$$

$$\frac{\partial Q_n}{\partial \beta} = \frac{2(\rho - 1)}{\beta^2(\lambda + 1)} \sum_{t=2}^n \left[ X_t - \rho X_{t-1} + \frac{2(1 - \rho)}{\lambda} + \frac{1 - \rho}{\beta(1 + \lambda)} \right]$$

e

$$\frac{\partial Q_n}{\partial \lambda} = \left( \frac{4(1 - \rho)}{\lambda^2} + \frac{2(\rho - 1)}{\beta(\lambda + 1)^2} \right) \sum_{t=2}^n \left[ X_t - \rho X_{t-1} + \frac{2(\rho - 1)}{\lambda} + \frac{1 - \rho}{\beta(1 + \lambda)} \right].$$

Resolvendo a primeira equação do sistema, obtemos

$$\begin{aligned} \frac{\partial Q_n}{\partial \rho} &= 2 \sum_{t=2}^n \left( -X_t X_{t-1} + X_t \frac{2\beta(1 + \lambda) - \lambda}{\beta\lambda(\lambda + 1)} + X_{t-1} \frac{2\beta(1 + \lambda) - \lambda}{\beta\lambda(\lambda + 1)} - \frac{(2\beta(1 + \lambda) - \lambda)^2}{\beta^2\lambda^2(\lambda + 1)^2} \right) \\ &+ 2\rho \sum_{t=2}^n \left( X_{t-1}^2 - 2X_{t-1} \frac{2\beta(1 + \lambda) - \lambda}{\beta\lambda(\lambda + 1)} + \frac{(2\beta(1 + \lambda) - \lambda)^2}{\beta^2\lambda^2(\lambda + 1)^2} \right) \end{aligned}$$

Portanto,  $\frac{\partial Q_n}{\partial \rho} = 0$  implica que

$$\hat{\rho} = \frac{\sum_{t=2}^n \left\{ X_t X_{t-1} + \frac{2\hat{\beta}(\hat{\lambda} + 1) - \hat{\lambda}}{\hat{\lambda}\hat{\beta}(\hat{\lambda} + 1)} \left( -X_t - X_{t-1} + \frac{2\hat{\beta}(\hat{\lambda} + 1) - \hat{\lambda}}{\hat{\lambda}\hat{\beta}(\hat{\lambda} + 1)} \right) \right\}}{\sum_{t=2}^n \left\{ X_{t-1} - \frac{2\hat{\beta}(\hat{\lambda} + 1) - \hat{\lambda}}{\hat{\lambda}\hat{\beta}(\hat{\lambda} + 1)} \right\}^2}.$$

Resolvendo a segunda equação do sistema, obtemos

$$\frac{\partial Q_n}{\partial \beta} = -\frac{2(1 - \rho)}{\beta^2(1 + \lambda)} \sum_{t=2}^n \left( X_t - \rho X_{t-1} + \frac{2(\rho - 1)}{\lambda} \right) - \frac{2(1 - \rho)^2(n - 1)}{(1 + \lambda)^2\beta^3}$$

Assim,  $\frac{\partial Q_n}{\partial \beta} = 0$  implica

$$\hat{\beta} = \frac{(1 - \hat{\rho})(n - 1)}{(1 + \hat{\lambda}) \sum_{t=2}^n \left( -X_t + \hat{\rho} X_{t-1} + \frac{2(1 - \hat{\rho})}{\hat{\lambda}} \right)}.$$

Finalmente, para a terceira equação do sistema

$$\frac{\partial Q_n}{\partial \lambda} = 2 \sum_{t=2}^n \left[ X_t - \rho X_{t-1} + \frac{2(\rho - 1)}{\lambda} + \frac{1 - \rho}{\beta(1 + \lambda)} \right] \left( \frac{2(1 - \rho)}{\lambda^2} + \frac{\rho - 1}{\beta(1 + \lambda)^2} \right).$$

Note que para  $\frac{\partial Q_n}{\partial \lambda} = 0$ , não é possível isolar  $\hat{\lambda}$  em função dos demais parâmetros.