



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

ERNANDO CAMPOS FERREIRA

**PROCESSOS AUTORREGRESSIVOS DE SEGUNDA ORDEM COM
DISTRIBUIÇÕES MARGINAIS LINDLEY E GAMA-LINDLEY**

Recife

2025

ERNANDO CAMPOS FERREIRA

**PROCESSOS AUTORREGRESSIVOS DE SEGUNDA ORDEM COM
DISTRIBUIÇÕES MARGINAIS LINDLEY E GAMA-LINDLEY**

Tese apresentada ao Programa de Pós-Graduação em Estatística do Centro de Ciências Exatas e da Natureza da Universidade Federal de Pernambuco, como requisito parcial à obtenção do título de doutora em Estatística.

Área de Concentração: Estatística Aplicada

Orientadora: Profa. Dra. Maria do Carmo Soares Lima

Coorientador: Prof. Dr. Abraão D. C. Nascimento

Recife

2025

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Ferreira, Ernando Campos.

Processos autoregressivos de segunda ordem com distribuições marginais Lindley e Gama-Lindley / Ernando Campos Ferreira. - Recife, 2025.

125f.: il.

Tese (Doutorado)- Universidade Federal de Pernambuco, Centro de Ciências Exatas e da Natureza, Programa de Pós-Graduação em Estatística, 2025.

Orientação: Maria do Carmo Soares Lima.

Coorientação: Abraão D. C. Nascimento.

1. Gama-Lindley; 2. Lindley; 3. Modelo autorregressivo; 4. Séries Temporais. I. Lima, Maria do Carmo Soares. II. Nascimento, Abraão D. C. III. Título.

UFPE-Biblioteca Central

ERNANDO CAMPOS FERREIRA

**PROCESSOS AUTORREGRESSIVOS DE SEGUNDA ORDEM COM
DISTRIBUIÇÕES MARGINAIS LINDLEY E GAMA-LINDLEY**

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

Aprovado em: 29 de agosto de 2025.

BANCA EXAMINADORA

Documento assinado digitalmente
MARIA DO CARMO SOARES DE LIMA
Data: 10/09/2025 13:20:32-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Maria do Carmo Soares de Lima
Presidente (Orientadora), UFPE

Documento assinado digitalmente
VINICIUS TEODORO SCHER
Data: 08/09/2025 10:40:13-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Vinícius Teodoro Scher
Examinador(a) Interno, UFPE

Prof. Dr. Josimar Mendes de Vasconcelos
Examinador(a) Externo, UFRPE

Documento assinado digitalmente
JOSIMAR MENDES DE VASCONCELOS
Data: 05/09/2025 05:09:29-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Leonardo Henrique Silva Fernandes
Examinador(a) Externo, UFRPE

Profa. Dra. Daiane Aparecida Zuanetti
Examinador(a) Externo, UFSCar

Documento assinado digitalmente
DAIANE APARECIDA ZUANETTI
Data: 03/09/2025 16:58:42-0300
Verifique em <https://validar.iti.gov.br>

AGRADECIMENTOS

Gostaria de expressar minha gratidão a todos que contribuíram para a realização deste trabalho.

À minha família e amigos, pelo apoio e compreensão durante toda esta jornada. Nos momentos de desafios, a presença direta ou indiretamente de cada um fizeram dessa longa caminhada uma travessia possível.

Aos meus orientadores, Profa. Maria do Carmo e Prof. Abraão, agradeço pela paciência e valiosas orientações. Cada encontro, reuniões, a dedicação e apoio de ambos foram essenciais na construção desse trabalho.

Aos meus colegas de pesquisa, pelo compartilhamento de ideias, discussões enriquecedoras e pelo apoio mútuo ao longo deste percurso.

Aos professores do Departamento de Estatística da UFPE, cujo saber e exemplo, contribuíram para minha formação acadêmica e pessoal.

Por fim, agradeço a Deus pela força nos momentos difíceis, perseverança e luz nos dias de dúvida e privilégio de tornar esse sonho realidade.

RESUMO

Na literatura, proposta de novas distribuições busca desenvolver modelos probabilísticos mais flexíveis tendo como uma das finalidades melhorar o ajuste a conjuntos de dados reais. Com isso, consegue-se uma abordagem mais precisa ao estudar fenômenos complexos, além de facilitar a análise dos parâmetros de um modelo. Já no contexto de séries temporais, pode-se fazer estudos em diversas áreas como, por exemplo, prever séries temporais financeiras, analisar padrões de taxas de juros e modelagem climática. Isso é feito ao obter modelos que identifiquem a dependência entre os valores passados e valores atuais de uma série, podendo, assim, fazer previsões. No entanto, existem poucos trabalhos desenvolvidos que utilizam distribuições atreladas aos modelos autorregressivos não condicionais, principalmente de segunda ordem. Acredita-se que isso ocorra devido à dificuldade de desenvolver expressões matemáticas tão complexas quanto as dos modelos elaborados para processos de primeira ordem. Para processos autorregressivos de primeira ordem, as observações são dependentes apenas de seus valores anteriores, o que não ocorre para processos de ordem 2, pois as análises são realizadas observando tanto os valores anteriores quanto os anteriores ao anterior, demandando métodos mais complexos. Contudo, pensando na lacuna existente na literatura quanto aos processos autorregressivos de segunda ordem, dois novos modelos autorregressivos com distribuição marginal Lindley e Gama-Lindley foram propostos. Por meio de uma série temporal de ordem 2, $AR(2)$, considerou-se, primeiramente, os dois parâmetros autorregressivos iguais. Essa condição foi chamada de *Caso 1*. Após isso, selecionou-se a distribuição Lindley e Gama-Lindley para a saída de cada processo, com o propósito de determinar a distribuição do erro. Os modelos apresentados foram chamados de $LAR_1(2)$ e $GLAR_1(2)$. Considerando a região de estabilidade para cada um deles, a distribuição da inovação de ambos modelos corresponde a uma mistura de duas distribuições exponenciais com parâmetros diferentes e uma Gama com parâmetros $(2, \lambda)$. Ambos modelos autorregressivos de ordem 2 foram introduzidos, apresentando suas propriedades, tais como as funções de autocorrelação e autocorrelação parcial, equação de previsão, densidade espectral, esperança e variância condicional. Para verificar suas propriedades assintóticas, utilizou-se os métodos de estimação Yule-Walker e mínimos quadrados condicionais (MQC), através de simulação de Monte Carlo. Afim de verificar a eficiência dos modelos, aplicou-se a dois conjuntos de dados reais e, por predição, os dois modelos tiveram um bom ajuste, além dos resíduos de cada um possuírem características de um ruído branco. Após construída toda a estrutura para o caso particular, *Caso 1*, generalizou-se os modelos

para o caso geral, *Caso 2*, em que os parâmetros autorregressivos são diferentes. Dessa forma, considerando a região de estabilidade para cada modelo, a distribuição de inovação para o modelo $LAR_2(2)$, corresponde a uma mistura de distribuições exponencial, com parâmetro λ , Gama com parâmetro $(2, \lambda)$ e três exponenciais com parâmetros (ϕ_1) , (ϕ_2) e (ϕ_3) que representam o oposto das raízes de um polinômio de grau 3, e dependem dos parâmetros do processo e da distribuição Lindley. De maneira similar, a distribuição de inovação para o modelo $GLAR_2(2)$ é uma mistura de distribuições exponencial, com parâmetro λ , Gama com parâmetro $(2, \lambda)$ e três exponenciais com parâmetros (φ_1) , (φ_2) e (φ_3) que representam o oposto das raízes de um polinômio de terceiro grau, que é função dos parâmetros do processo e da distribuição Gama-Lindley. Conforme feitas para o *Caso 1*, as simulações também foram realizadas utilizando os métodos Yule-Walker e MQC por meio de simulação de Monte Carlo. Também aplicou-se os modelos a conjuntos de dados reais confirmando um bom ajuste, além do resíduo não possuir dependência. Por fim, os modelos propostos foram comparados com outros na literatura, de modo que os modelos $LAR_1(2)$ e $LAR_2(2)$, mostraram ser um bons competidores em relação ao modelos, $EAR(2)$ $GAR(2)$ e $INGAR(2)$. Já os modelos $GLAR_1(2)$ e $GLAR_2(2)$ apresentaram melhores resultados em relação aos modelos $LAR_1(2)$, $LAR_2(2)$, $EAR(2)$ e $GAR(2)$ e $INGAR(2)$. De modo geral, ambos os modelos destacaram-se como bons concorrentes aos já existentes na literatura.

Palavras-chaves: Gama-Lindley. Lindley. Modelo autorregressivo. Séries Temporais.

ABSTRACT

In the literature, the proposal of new distributions aims to develop more flexible probabilistic models with the purpose of improving the fit to a set of real data. This enables a more precise approach when studying complex phenomena, as well as facilitating the analysis of model parameters. In the context of time series, studies can be conducted in various areas such as predicting financial time series, analyzing interest rate patterns, and climate modeling. This is achieved by obtaining models that identify the dependence between past values and current values of a series, thereby enabling predictions. However, there are few developed works that use distributions linked to autoregressive models, especially of the second order. This is due to the difficulty of developing mathematical expressions as complex as those designed for first-order processes. For first-order autoregressive processes, observations are dependent only on their previous values, which is not the case for second-order processes, as analyses are performed by observing both the previous values and the values before the previous one, requiring more complex methods. However, considering the existing gap in the literature regarding second-order autoregressive processes, two new autoregressive models with Lindley and Gama-Lindley marginal distributions were proposed. Through a second-order time series, $AR(2)$, the two autoregressive parameters were considered equal. This condition was called *Case 1*. After that, the Lindley and Gama-Lindley distributions were selected for the output of each process, with the purpose of determining the error distribution. The presented models were called $LAR_1(2)$ and $GLAR_1(2)$. Considering the stability region for each of them the innovation distribution of both models corresponds to a mixture of two exponential distributions with different parameters and a Gamma with parameters $(2, \lambda)$. Both second-order autoregressive models were introduced, presenting their properties such as autocorrelation and partial autocorrelation functions, prediction equation, spectral density, and conditional mean and variance. To verify their asymptotic properties, the Yule-Walker and conditional least squares estimation methods were used through Monte Carlo simulation. To verify the efficiency of the models, were applied to two real data sets. Through prediction, both models showed a good fit, and the residuals of each exhibited characteristics of white noise. After building the entire structure for these particular cases, *Case 1*, the models were generalized to the general case, *Case 2*, in which the autoregressive parameters are different. Thus, considering the region of stability for each model, the innovation distribution for the $LAR_2(2)$ model corresponds to a mixture of Exponential distributions, with parameter λ , Gamma with parameter $(2, \lambda)$ and three expo-

nentials with parameters (ϕ_1) , (ϕ_2) and (ϕ_3) representing the opposite roots of a polynomial of degree 3, and depend on the process parameters and the Lindley distribution. Similarly, the innovation distribution for the $GLAR_2(2)$ model is a mixture of the Exponential distribution with parameter λ , the Gamma distribution with parameter $(2, \lambda)$ and three exponentials with parameters (φ_1) , (φ_2) and (φ_3) representing the opposite roots of a third-degree polynomial, which is a function of the parameters of the process and the Gama-Lindley distribution. As was done for case 1, the simulations were also carried out using the Yule-Walker and conditional least squares methods by means of Monte Carlo simulation. The models were also applied to real data sets, confirming a good fit and that the residuals were not dependent. Finally, the proposed models were compared with others in the literature, so that $LAR_1(2)$ and $LAR_2(2)$ proved to be a good competitors in relation to models $EAR(2)$, $GAR(2)$, and $INGAR(2)$. The $GLAR_1(2)$ and $GLAR_2(2)$ models showed better results than the $LAR_1(2)$, $LAR_2(2)$, $EAR(2)$, $GAR(2)$ and $INGAR(2)$ models. In general, both models stood out as good competitors to those already existing in the literature.

Keywords: Gama-Lindley. Lindley. Autoregressive model. Time Serie.

LISTA DE FIGURAS

Figura 1 – Triângulo onde as condições de estacionariedade para AR(2) são satisfeitas	27
Figura 2 – Triângulo de estabilidade em um processo AR(2), <i>Caso 1</i>	33
Figura 3 – Exemplo de uma série não estacionária	33
Figura 4 – Espaço paramétrico do parâmetro λ .	34
Figura 5 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 para os valores $\alpha_1 = \alpha_2 = a = 0.2, \lambda = 12$	41
Figura 6 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 para os valores $\alpha_1 = \alpha_2 = a = 0.3, \lambda = 12$	41
Figura 7 – Gráfico da densidade do erro para amostra de tamanho 500, com $\lambda = 12$, variando o parâmetro a .	42
Figura 8 – Fechamento da bolsa de valores	51
Figura 9 – Autocorrelação e autocorrelação parcial para a série	51
Figura 10 – Gráfico para as séries observada e $LAR_1(2)$	52
Figura 11 – Análise dos resíduos para o $LAR_1(2)$	52
Figura 12 – Triângulo de estabilidade para um processo AR(2), <i>Caso 2</i>	54
Figura 13 – Região em que as raízes do polinômio são reais	59
Figura 14 – Espaço paramétrico para os parâmetros a e λ , considerando as raízes r_1, r_2 e r_3 do polinômio.	61
Figura 15 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.6, \alpha_2 = 0.3, \lambda = 100$	63
Figura 16 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.22, \alpha_2 = 0.68, \lambda = 100$	63
Figura 17 – Gráficos da densidade do erro para amostra de tamanho 500, e diferentes valores para os parâmetros.	64
Figura 18 – Gráfico para as séries observadas e $LAR_2(2)$	67
Figura 19 – Autocorrelação e autocorrelação parcial do resíduo	67
Figura 20 – Gráfico para a série	68
Figura 21 – Gráfico para as séries observadas e $LAR_2(2)$	69
Figura 22 – Autocorrelação e autocorrelação parcial do resíduo	69
Figura 23 – Região satisfeita para os parâmetros β e λ .	74

Figura 24 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 com os valores $a = 0.3, \lambda = 1, \beta = 0.5$	76
Figura 25 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 com os valores $a = 0.4, \lambda = 1.5, \beta = 0.6$	76
Figura 26 – Gráfico da densidade do erro para amostra de tamanho 500, com $\lambda = 0.1, \beta = 0.5$, variando o parâmetro a	77
Figura 27 – Gráfico para as séries observadas e $GLAR_1(2)$	81
Figura 28 – Análise dos resíduos para o $GLAR_1(2)$	82
Figura 29 – Região em que as raízes do polinômio de grau 3 são reais	86
Figura 30 – Espaço paramétrico para os parâmetros β e λ	87
Figura 31 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.25, \alpha_2 = 0.65, \beta = 0.85, \lambda = 5$	89
Figura 32 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.32, \alpha_2 = 0.58, \beta = 0.93, \lambda = 18$	89
Figura 33 – Gráficos da densidade do erro para amostra de tamanho 500, com $\lambda = 10, \beta = 0.92$ e diferentes valores para os autorregressivos.	90
Figura 34 – Gráfico para as séries observadas e $GLAR_2(2)$	93
Figura 35 – Análise dos resíduos para o $GALAR(2)$	93
Figura 36 – Gráfico para as séries observadas e $GLAR_2(2)$	94
Figura 37 – Análise dos resíduos do $GLAR_2(2)$	95

LISTA DE TABELAS

Tabela 1 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados	49
Tabela 2 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados Condicionais	66
Tabela 3 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados	80
Tabela 4 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados	92
Tabela 5 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados do mercado de ações - <i>Caso 1</i>	102
Tabela 6 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados do mercado de ações - <i>Caso 2</i>	102
Tabela 7 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados em hidrologia- <i>Caso 1</i>	103
Tabela 8 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados em hidrologia- <i>Caso 2</i>	103
Tabela 9 – Valores de λ para diferentes ϵ	114

SUMÁRIO

1	INTRODUÇÃO GERAL	15
1.1	CONTEXTUALIZAÇÃO	15
1.2	MOTIVAÇÃO E OBJETIVOS	18
1.3	PLATAFORMA COMPUTACIONAL	19
1.4	ORGANIZAÇÃO TEXTUAL	19
2	REVISÃO LITERÁRIA	22
2.1	DISTRIBUIÇÃO LINDLEY	22
2.2	PROCESSO AUTORREGRESSIVO DE SEGUNDA ORDEM - AR(2)	25
2.3	FUNÇÃO DE AUTOCORRELAÇÃO	27
3	A CONSTRUÇÃO DE UM NOVO MODELO LINDLEY AUTOR- REGRESSIVO DE SEGUNDA ORDEM	29
3.1	CASO 1: PROCESSO AR(2) COM DISTRIBUIÇÃO MARGINAL LINDLEY	31
3.1.1	Propriedades Estruturais do novo modelo	36
3.1.2	Propriedades não condicionais	37
3.1.3	Propriedades Condicionais	38
3.1.4	Autocorrelação, Autocorrelação Parcial e Densidade do erro	39
3.1.5	Previsão	42
3.1.6	Estimação do Processo LAR ₁ (2)	44
3.1.6.1	Método Yule-Walker	44
3.1.6.2	Mínimos Quadrados Condicionais	47
3.1.7	Resultados das Simulações	48
3.1.8	Aplicação	49
3.1.9	Considerações Parciais	53
3.2	CASO 2: PROCESSO AR(2) COM DISTRIBUIÇÃO MARGINAL LINDLEY	54
3.2.1	Propriedades Estruturais	62
3.2.2	Método Yule-Waker	64
3.2.3	Mínimos Quadrados Condicionais	65
3.2.4	Resultados das Simulações	66
3.2.5	Aplicação	66
3.2.6	Considerações Parciais	69

4	A CONSTRUÇÃO DE UM NOVO MODELO GAMA-LINDLEY DE SEGUNDA ORDEM	71
4.1	DISTRIBUIÇÃO GAMA-LINDLEY	71
4.2	A CONSTRUÇÃO DE UM NOVO MODELO AUTORREGRESSIVO DE SEGUNDA ORDEM	72
4.2.1	Propriedades Estruturais	75
4.2.1.1	<i>Esperança e Variância Condicional</i>	75
4.2.1.2	<i>Autocorrelação, Autocorrelação Parcial e Densidade do erro</i>	76
4.2.2	Previsão	77
4.2.3	Estimação	78
4.2.3.1	<i>Mínimos Quadrados Condicionais</i>	78
4.2.3.2	<i>Método Yule-Walker</i>	78
4.2.4	Resultado das Simulações	78
4.2.5	Aplicação	80
4.2.6	Considerações Parciais	82
4.3	CASO 2- PROCESSO AR(2) COM DISTRIBUIÇÃO MARGINAL GAMA-LINDLEY	83
4.3.1	Propriedades Estruturais	89
4.3.2	Simulação	90
4.3.2.1	<i>Mínimos Quadrados Condicionais</i>	91
4.3.2.2	<i>Resultados das Simulações</i>	91
4.3.3	Aplicação	93
4.3.4	Considerações Parciais	95
5	COMPARAÇÃO DOS MODELOS	97
5.1	ANÁLISE DOS MODELOS PROPOSTOS	97
5.1.1	Critérios de Seleção	98
5.1.2	Processo EAR(2)	99
5.1.3	Processo GAR(2)	100
5.1.4	INGAR(2)	101
5.2	RESULTADOS	101
6	CONCLUSÃO E TRABALHOS FUTUROS	104
	REFERÊNCIAS	105
	APÊNDICE A –	109

APÊNDICE B –	112
APÊNDICE C –	115
APÊNDICE D –	119
APÊNDICE E –	122
APÊNDICE F –	124

1 INTRODUÇÃO GERAL

1.1 CONTEXTUALIZAÇÃO

Nas últimas décadas, uma infinidade de novas distribuições foram propostas na literatura. Com isso, espera-se que novos modelos se ajustem adequadamente a diversos tipos de conjuntos reais de dados. Ao sugerir novas distribuições ou aprimorar as já existentes, tem-se como objetivo aprofundar suas propriedades estatísticas, bem como demonstrar que esses novos modelos são tão eficientes ou mais que os já vigentes. Ver (LIMA et al., 2019).

Geralmente, esse processo segue um protocolo sistemático. A primeira fase, consiste na elaboração de um mecanismo gerador, paralelamente é estabelecido a formulação analítica da distribuição, a fim de garantir sua validade matemática. A próxima fase, concentra-se na caracterização teórica da nova distribuição, abrangendo a obtenção das suas propriedades fundamentais. Em seguida, são realizados experimentos computacionais afim de analisar o comportamento assintótico dos estimadores, por meio de simulações (CORDEIRO et al., 2019).

Por fim, a última etapa se encerra com uma aplicação na qual compara-se o desempenho da nova distribuição com modelos conhecidos da literatura. Isso é realizado por meio de testes estatísticos e critérios para qualidade do ajuste. Essa etapa torna-se importante, uma vez que tem por objetivo comparar a relevância ou até mesmo a superioridade da nova proposta em relação a modelos existentes na literatura.

No contexto de séries temporais de primeira e segunda ordem, as aplicações podem ser feitas em diversos campos, por exemplo, em finanças e análises de risco, nas quais utilizam séries dessas ordens para prever tendências e comportamentos futuros no mercado. Há também aplicação em hidrologia, em que valem de séries temporais para prever padrões de precipitações. Dessa forma, consegue-se fazer planejamento de irrigação agrícola e até mesmo prever desastres naturais.

Nessa perspectiva, tem-se o trabalho de Callen e Morel (2001) que apresenta uma contribuição significativa na área de contabilidade. O estudo partia do princípio que o valor de mercado representa uma função linear entre os valores contábeis do patrimônio líquido e os lucros acima do mercado, denominados de anormais. Assim, considerou que os lucros assumiam um processo autorregressivo de primeira ordem, $AR(1)$. No entanto, os autores constataram que, quando os lucros acima do normal forem um autorregressivo de segunda ordem, $AR(2)$, a relação entre valor de mercado e variáveis contábeis apresentam uma relação de dois *lags*.

Assim, o modelo AR(2) permite identificar padrões de persistência dos lucros mais flexíveis.

Ainda observando os estudos sobre processos autorregressivos, Hugueny (2006) utilizou processos AR(1) e AR(2) para analisar a sincronia espacial em componentes demográficos e ambientais entre duas populações. Através de um banco de dados reais representando um conjunto regional de populações, notou-se que elas apresentaram comportamentos distintos. O estudo feito, utilizando o processo AR(1), destacou um padrão simples, diferente do AR(2) que demonstrou ter estruturas complexas, não apresentando relações entre a dinâmica populacional e o fator demográfico da sintonia espacial.

No mesmo ano, Russo, Rodrigues e Camargo (2006) desenvolveram um estudo utilizando análise de séries temporais para descrever as alterações nos moldes de peças produzidas por uma indústria. Os autores buscaram modelos que simulassem a realidade dos dados, ao fazer previsões. O banco de dados utilizado representa o número diário do teor da umidade de areia. Os resultados indicaram que o processo AR(2) foi o melhor modelo encontrado.

Com aplicação em medicina, Antunes e Cardoso (2015) abordaram séries temporais em estudos epidemiológicos ao fazer previsão sobre a distribuição de doenças em uma população. Milani (2016) apresenta modelos de séries temporais autorregressivos (AR(1)) e de médias móveis (MA(1)), usando distribuições Normal generalizada e log-Normal generalizada, fazendo aplicação nas áreas de hidrologia, economia e políticas públicas.

Balbinot et al. (2017) utilizaram modelos de séries temporais para descrever valores da radiação solar, concluindo que processos autorregressivos e de médias móveis descrevem bem o comportamento da série pesquisada. Souza (2022) abordou a trajetória da dívida pública federal do Brasil, entre os anos de 2004 e 2020 utilizando modelo autorregressivo de médias móveis (ARMA). Isso possibilitou prever o comportamento da dívida pública para os próximos 20 meses.

No entanto, existem poucos trabalhos que utilizam a teoria de novas distribuições em conjunto com análises de séries temporais. Além disso, é importante destacar que existem dois tipos de abordagens: os modelos condicionais e os não condicionais. Os modelos condicionais consideram que os parâmetros da distribuição variam ao longo do tempo. Assim, tem-se o trabalho Santos (2006) na área do mercado financeiro, uma abordagem condicional que tem ganhado destaque por sua capacidade de incorporar informações de mercado de forma dinâmica. Neste trabalho, foi examinado um método quantitativo avançado para avaliação de risco extremo, comparando-o com técnicas em análises do Ibovespa e ações da Petrobras.

Ainda no contexto da modelagem condicional aplicada ao mercado financeiro, destaca-se o

trabalho de Lima et al. (2021), que propõe o modelo Gompertz-ARMA como uma abordagem inovadora para séries temporais. Nesta estrutura condicional, os parâmetros são dinamicamente recalibrados em função do comportamento histórico da série, superando as limitações dos modelos de parâmetros fixos. A flexibilidade do Gompertz-ARMA permite capturar não apenas variações locais, mas também regimes não-estacionários característicos de séries financeiras voláteis. Essa capacidade adaptativa foi comprovada na aplicação prática aos preços da Petrobras, onde o modelo demonstrou eficácia em acompanhar mudanças abruptas de volatilidade, mesmo em cenários de crise.

Embora os modelos condicionais apresentem vantagens ao incorporar a dinâmica temporal, neste trabalho será adotado a abordagem não-condicional. Tal escolha fundamenta-se na adequação dos modelos não condicionais para análise de séries estacionárias, ou seja, suas propriedades estatísticas não variam ao longo do tempo. Outra vantagem de se utilizar modelos não condicionais, refere-se a simplicidade operacional. Pelo fato de não apresentar mecanismos de atualização paramétrica, faz com que a complexidade computacional seja reduzida, vantagem significativa ao analisar séries longas.

Assim, esta abordagem permite capturar padrões estruturais subjacentes que frequentemente permanecem mascarados pela excessiva sensibilidade a variações de curto prazo, típica dos modelos condicionais. Além disso, essa abordagem apresenta clareza na interpretação, em que os parâmetros estimados caracterizam a distribuição marginal do processo, excluindo interferências e dinâmicas locais.

Nesse contexto, tem-se o trabalho desenvolvido por Bakouch e Popović (2016), que propõe um modelo AR(1) com distribuição marginal Lindley. Depois desse período, poucos estudos com esse enfoque foram desenvolvidos. Mello et al. (2021) apresentaram um novo modelo de série temporal (ST) de ordem 1 com distribuição marginal Gama-Lindley e compararam seu desempenho preditivo com outros modelos existentes na literatura, tomando como aplicação um conjunto de dados em hidrologia.

Desse modo, ainda há uma extensa defasagem na literatura quanto a novas distribuições atrelado ao modelo autorregressivo de ST de segunda ordem. Uma possível justificativa para tal fato, talvez seja a dificuldade de desenvolver cálculos matemáticos complexos e obter uma análise computacional eficaz. Lawrance e Lewis (1980) propõem um modelo autorregressivo exponencial de ordem 2. No entanto, para o caso em questão, a abordagem é simples, visto que apresenta apenas um parâmetro em sua estrutura. Além disso, os autores não apresentam aplicação nem simulação computacionais, o que motiva a construção de novos trabalhos neste

sentido. Com isso, na Subseção 1.2 é exposta as motivações para construção dos modelos e os objetivos a serem alcançados.

1.2 MOTIVAÇÃO E OBJETIVOS

Existem na literatura poucos trabalhos direcionados ao estudo da distribuição Lindley no cenário de Séries Temporais. Só em 2016, 50 anos depois que a distribuição Lindley foi proposta, Bakouch e Popović (2016) desenvolveram o primeiro processo autorregressivo tendo a Lindley como distribuição marginal. A dificuldade se torna maior ao abordar processos autorregressivos de segunda ordem, tendo o trabalho de Lawrance e Lewis (1980) que desenvolveram processos exponenciais autorregressivos de médias móveis - EARMA(p, q) como um dos únicos como referência desenvolvido há mais de 40 anos. Assim, tem-se como objetivo geral propor dois novos modelos AR(2) não condicionais, chamados de LAR(2) e GLAR(2), com distribuição marginal Lindley e Gama-Lindley, respectivamente. Como objetivos específicos pretende-se:

1. Determinar uma distribuição para a inovação de cada modelo;
2. Descobrir os pesos correspondentes a cada distribuição que compõe as inovações;
3. Comprovar que a soma dos pesos da variável de saída, é de fato, uma mistura de distribuições;
4. Desenvolver propriedades estruturais para os modelos AR(2), como a esperança e a variância condicional a fim de analisar o comportamento dos modelos sobre determinadas condições;
5. Construir um gerador para o erro aleatório;
6. Analisar os gráficos da série, autocorrelação e autocorrelação parcial para comprovar que novos modelos provém de um AR(2);
7. Construir a densidade do erro para cada modelo;
8. Realizar estimação dos parâmetros por meio do método Yule-Walker e mínimos quadrados, adaptando para os processos em estudo;
9. Estudar as regiões onde as condições de estacionariedade são satisfeitas, para cada processo AR(2) em construção;

10. Realizar uma aplicação a ST, mostrando a versatilidade dos modelos, ao utilizar bancos de dados reais e analisar previsões.
11. Comparar os modelos propostos com outros bem definidos na literatura.

1.3 PLATAFORMA COMPUTACIONAL

Essa tese será desenvolvida utilizando o $\text{\LaTeX}$, um sistema de formatação de texto bastante difundido na comunidade acadêmica para a elaboração de trabalhos científicos. Para a realização das funções computacionais, utilizou-se os softwares R e Matlab, além da linguagem de programação em Python. Os cálculos das transformadas inversa de Laplace, simplificações e resoluções de expressões mais complexas foram desenvolvidos no Matlab. O Matrix Laboratory ou Matlab é um software com um ambiente interativo e de fácil análise numérica, visualização de dados e modelagem matemática. Para a simulação dos parâmetros, construção das séries temporais, desenvolvimento dos gráficos, afim de estudar os modelos e ajustes a conjuntos de dados reais, utilizou-se o R. Ambos programas se diferem pelo fato do R ser um software gratuito. Já o Matlab é um software comercial e requer licença paga para o uso, apesar de existirem opções de licenciamento acadêmico. Além disso, para análises estatísticas e análise de dados, o R se sobressai por possuir uma infinidade de pacotes e recursos, além de ser mais flexível.

1.4 ORGANIZAÇÃO TEXTUAL

No Capítulo 2, é apresentada a definição da distribuição Lindley com algumas de suas extensões e generalizações, mostrando aplicação em diversas áreas. Na Seção 2.2 é feita uma revisão sobre Análise de Séries Temporais e suas propriedades, apresentando em seguida um modelo $\text{AR}(2)$. Nesse mesmo capítulo, são discutidos critérios para a região de estacionariedade de um processo $\text{AR}(2)$. Conceitos de estacionariedade fraca e autocorrelação também serão definidos.

No Capítulo 3, um novo processo autorregressivo não condicional de ordem 2 com distribuição marginal Lindley, chamado $\text{LAR}(2)$ é construído. Devido a sua complexidade, o processo foi dividido em duas partes, chamados de *Caso 1* e *Caso 2*.

A partir desse ponto, a Seção 3.1 apresenta o modelo $\text{LAR}_1(2)$ considerando o *Caso 1*,

com uma restrição aos parâmetros autorregressivos. Em seguida, são desenvolvidas algumas propriedades estruturais como esperança e variância condicional, autocorrelação, autocorrelação parcial e densidade do erro, que podem ser vistos nas Subseções 3.1.2 e 3.1.4. A Subseção 3.1.5 mostra o processo para fazer previsão do novo modelo apresentado.

Na Subseção 3.1.6 é apresentado o método de estimação Yule-Waker juntamente com a forma fechada para estimar os parâmetros autorregressivos do modelo, por meio das equações de Yule-Walker, bem como o método de estimação por MQC. Os resultados das simulações usando esses métodos para vários cenários são mostrados na Subseção 3.1.7. Na Subseção 3.1.8 o modelo é aplicado a um conjunto de dados reais referentes a um processo $AR(2)$, com o objetivo de verificar se ele se ajusta adequadamente a esses tipos de dados. Em seguida, para finalizar a análise da construção do modelo, considerando o *Caso 1* são feitas as conclusões parciais na Subseção 3.1.9.

Ainda no âmbito das discussões do Capítulo 3, na Seção 3.2 o modelo $LAR_2(2)$ é proposto, considerando a forma geral, chamada de *Caso 2*. Dessa forma, são apresentadas as propriedades estruturais do modelo na Subseção 3.2.1, que foram derivadas das desenvolvidas para o caso particular - *Caso 1*. Por fim, as simulações, aplicação e as considerações são feitas nas Subseções 3.2.4, 3.2.5 e 3.2.6, respectivamente.

No Capítulo 4, é apresentado um modelo $AR(2)$, não condicional, com distribuição marginal Gama-Lindley, chamado de $GLAR(2)$. Na Seção 4.1, é feita uma revisão sobre a distribuição Gama-Lindley e seus vários tipos. Em seguida, na Seção 4.2, é apresentado um novo modelo com distribuição marginal Gama-Lindley. Da mesma forma como no capítulo anterior, a discussão foi apresentada em duas partes: *Caso 1*, em que o modelo $GLAR_1(2)$ foi construído e *Caso 2*, para o $GLAR_2(2)$.

Em seguida, a Subseção 4.2.1 apresenta as propriedades desenvolvidas para o novo modelo. Na Subseção 4.2.2 é discutido o processo para realizar previsão. Já a estimação usando os métodos Yule-Walker e MQC junto com os resultados das simulações são feitas nas Subseções 4.2.3 e 4.2.4, respectivamente. A aplicação a conjunto de dados reais é mostrada na Subseção 4.2.5 e as considerações parciais do capítulo são feitas em na Subseção 4.2.6.

No Capítulo 5, é feita a comparação dos modelos utilizando as bases de dados reais propostas para modelos $AR(2)$. Para verificar a efetividade dos novos modelos, realizou-se uma comparação entre eles, analisando o número de parâmetros existentes e a qualidade do ajuste. Além disso, os modelos propostos serão comparados com outros desenvolvidos na literatura como o Exponencial autorregressivo $EAR(2)$, Gama autorregressivo $GAR(2)$ e Inverso

gaussiano autorregressivo INGAR(2), em que é feita uma descrição sobre cada um deles. Para isso, planeja-se analisar o desempenho dos modelos utilizando os critérios estatísticos como AIC, BIC e RMSE. Os resultados das análises dos modelos são feitos na Seção 5.2. Por fim, na Seção 6 apresenta-se as conclusões da Tese e perspectivas para trabalhos futuros.

2 REVISÃO LITERÁRIA

Neste Capítulo, é feito uma discussão sobre a distribuição Lindley, apontando algumas de suas extensões e aplicações em diversas áreas. Na Seção 2.1, tem-se uma revisão sobre estacionariedade fraca. Após isso, na Seção 2.2 é apresentada a estrutura de um $AR(2)$. Além disso, é realizada uma investigação com o intuito de verificar a região em que as condições de estacionariedade são satisfeitas. Ao final do capítulo, é elaborada uma discussão sobre a autocorrelação. Por meio da autocorrelação, é encontrada uma expressão para variância de um modelo $AR(2)$.

2.1 DISTRIBUIÇÃO LINDLEY

É importante utilizar modelos probabilísticos para descrever com mais precisão dados reais. Dentro desse contexto, em Análise de Sobrevivência, por exemplo, a distribuição exponencial tem sido amplamente utilizada (PERCONTINI et al., 2014). Uma das suas vantagens refere-se a sua simplicidade paramétrica, além de apresentar propriedades matemáticas mais dinâmicas.

Embora a distribuição exponencial ofereça vantagens como praticidade e facilidade de interpretação, tem-se a distribuição Lindley como uma boa competidora. Proposta por Lindley (1958), a distribuição Lindley foi desenvolvida como uma abordagem eficaz sendo uma alternativa vantajosa em relação à distribuições clássicas, como a exponencial. Tal fato decorre porque a distribuição Lindley apresenta uma estrutura matemática que combina características da exponencial e Gama. Além disso, possui maior versatilidade e um melhor ajuste a aplicações práticas, veja Ghitany, Atieh e Nadarajah (2008).

Isso se deve à flexibilidade matemática apresentada pela Lindley. Já que, ao contrário da distribuição exponencial, que possui apenas um parâmetro, a distribuição Lindley possui um parâmetro extra x , além do parâmetro λ (ver Equação (2.1)). Isso permite um ajuste mais versátil com maior controle sobre assimetria e cauda da distribuição.

Além disso, a distribuição Lindley possui uma função de risco não constante, ao contrário da distribuição exponencial, que apresenta taxa de falha fixa. Logo, a distribuição Lindley consegue modelar melhor fenômenos com uma taxa de falha variável ao longo do tempo. Da mesma forma, a distribuição Lindley se adapta melhor às particularidades dos processos estocásticos reais. Tal fato se explica porque, em Análise de Sobrevivência, por exemplo, a

Lindley captura melhor a cauda mais longa em relação a distribuição exponencial que possui função de risco constante e coeficiente de variação fixo. Ver (GHITANY; ATIEH; NADARAJAH, 2008).

Assim, por essas características, a Lindley destaca-se pela versatilidade matemática. Em vista disso, Mazucheli e Achcar (2011) também estudaram a distribuição Lindley, atribuindo esta como uma alternativa compatível às distribuições exponencial e Weibull, em análise de riscos competitivos. Nesse trabalho, os autores escolheram a distribuição Lindley por apresentar uma simplicidade nas funções de risco, sobrevivência e densidade, além de apresentar uma boa flexibilidade ao ajuste de dados reais.

No mesmo período tem-se o trabalho de Krishna e Kumar (2011) que utiliza a Lindley como um modelo de confiabilidade. A aplicação é realizada considerando o tempo de espera de clientes em uma agência bancária. Os resultados mostram que a Lindley apresentou um bom ajuste em comparação com as distribuições Gama e log-normal.

Em seu estudo, Al-Mutairi, Ghitany e Kundu (2013) investigaram as propriedades da distribuição Lindley e sua aplicação ao modelo stress-strength em dados reais de análise de sobrevivência, demonstrando que esse modelo apresenta melhor ajuste comparado às distribuições exponencial e Weibull. Adicionalmente, por meio de estudos de simulação, os autores avaliaram o viés e o erro quadrático médio (EQM) dos estimadores, constatando que a distribuição Lindley fornece resultados inferenciais mais precisos que os modelos exponencial e Weibull.

Considere que X segue uma distribuição Lindley, sua função densidade de probabilidade (fdp) é dada por

$$f(x; \lambda) = \frac{\lambda^2}{1 + \lambda} (1 + x) \exp(-\lambda x), \quad \text{para } x > 0, \quad (2.1)$$

em que $\lambda > 0$ é o parâmetro de forma da distribuição.

Sua função de distribuição acumulada (fda) é dada por

$$F(x; \lambda) = 1 - \left(\frac{1 + \lambda + \lambda x}{1 + \lambda} \right) \exp(-\lambda x), \quad \text{para } x > 0.$$

Para fins de notação, neste trabalho, escrevemos X tem distribuição Lindley com parâmetro λ , será denotado como, $X \sim L(\lambda)$.

O valor esperado e a variância são, respectivamente:

$$\mathbb{E}(X) = \frac{(2 + \lambda)}{\lambda + \lambda^2} \quad \text{e} \quad \text{Var}(X) = \frac{\lambda^2 + 4\lambda + 2}{\lambda^2(1 + \lambda)^2}. \quad (2.2)$$

Nas últimas décadas, diversas extensões da distribuição Lindley foram introduzidas. Sankaran (1970) apresentou a distribuição Poisson-Lindley, para modelar dados de contagem, como o número de acidentes ocorridos por mulheres no ambiente de trabalho. Nadarajah, Bakouch e Tahmasbi (2011) desenvolveram a distribuição Lindley generalizada, uma versão com 2 parâmetros em sua estrutura, com melhor ajuste às distribuições Weibull, gama e log-normal, tendo como aplicação a modelagem de tempos de vida.

Da mesma forma, Bakouch et al. (2012) apresentaram a distribuição Lindley extendida com 3 parâmetros aplicada a um conjunto de dados reais. Essa distribuição mostrou um ajuste superior aos dados quando comparadas a outras distribuições como a Weibull, Weibull modificada e Gama. No ano seguinte, Shanker e Mishra (2013) propõem a distribuição quase Lindley, com 2 parâmetros em sua estrutura. Testes em conjuntos de dados reais foram realizados, mostrando que, para quase todos eles, a distribuição quase Lindley apresentou resultados satisfatórios em relação a Lindley.

Simultaneamente, Ghitany et al. (2013) propuseram a distribuição Lindley potência com dois parâmetros, considerando a potência de uma variável aleatória (v.a) com a distribuição Lindley. Dessa maneira, essa distribuição tem sido aplicada em áreas como análise de risco, análise de sobrevivência e modelagem de dados de contagem em epidemiologia. Em 2015, Sharma et al. (2015) empregaram o inverso da v.a da distribuição Lindley, propondo assim a distribuição Lindley inversa. Como aplicação, utilizaram-na para modelar o tempo de vida de pacientes com uma determinada doença.

No mesmo ano, Lima (2015), desenvolveu uma extensão da distribuição Lindley denominada Gama-Lindley, com dois parâmetros em sua estrutura. O modelo construído apresentou melhor ajuste para dois conjuntos de dados reais em comparação às distribuições Lindley e Exponencial Geométrico Complementar. Atualmente, existem várias distribuições Gama-Lindley, como a desenvolvida por Zeghdoudi e Nedjar (2015), que será objeto de investigação desta tese.

Logo em seguida, Arslan, Acitas e Senoglu (2017) apresentaram a distribuição Lindley generalizada como uma melhor opção na modelagem estatística da velocidade do vento, em relação a distribuição Weibull. No ano seguinte, Joshi e Jose (2018) propõem a distribuição Lindley envolvida apresentando diversas de suas propriedades como a função característica e momentos trigonométricos. O novo modelo foi aplicado a conjuntos de dados reais em ecologia, com o objetivo de fornecer diretrizes precisas para orientar as tartarugas após a postura dos ovos. Os resultados indicam que a distribuição Lindley envolvida ajustou melhor o conjunto

de dados apresentando melhores resultados em comparação à distribuição exponencial.

Nesse cenário, Dey, Nassar e Kumar (2019) desenvolveram uma nova distribuição com dois parâmetros, sendo uma generalização das distribuições Lindley inversa e Weibull inversa. Sua aplicação consiste em representar o tempo de sobrevivência de um grupo de pacientes com câncer tratados com radioterapia. Essa nova classe de distribuição apresentou um bom ajuste, de acordo com os resultados dos critérios de informação. A nova distribuição mostrou-se adequada em relação aos outros modelos tradicionais testados.

No ano seguinte, tem-se o estudo realizado por Mazucheli, Bapat e Menezes (2020), introduzindo a distribuição unit-Lindley, uma extensão da distribuição Lindley projetada para modelar dados limitados ao intervalo $(0, 1]$. Essa distribuição mantém as características principais da Lindley, além de apresentar propriedades relevantes como expressão fechada para os momentos. Além disso, a distribuição unit-Lindley apresentou bom ajuste a conjuntos de dados em finanças, em comparação com o modelo de regressão beta.

No entanto, mesmo com vários trabalhos propondo extensões da distribuição Lindley e para o cenário de independência, tem-se pouca abordagem no contexto de séries temporais. Diante disso, essa lacuna possui uma relevância, visto que, ao apresentar assimetria e caudas pesadas, a Lindley pode modelar de maneira mais adequada séries financeiras e padrões climáticos extremos. A próxima seção versa sobre séries temporais estacionárias e apresenta uma estrutura para processos AR(2), foco desse trabalho.

2.2 PROCESSO AUTORREGRESSIVO DE SEGUNDA ORDEM - AR(2)

Para analisar como os valores da série se comportam ao longo do tempo, é necessário observar se ela atende as condições de estacionariedade, com média, variância e autocorrelação constante, ao longo do tempo. Esse conceito de estacionariedade também pode ser definido como estacionariedade fraca. Assim, uma ST $\{X_t, t \in \mathbb{Z}\}$ é dita estacionária se

- i) $\mathbb{E}(X_t) = k, \quad \forall \quad t \in \mathbb{Z},$
- ii) $\mathbb{E}|X_t|^2 < \infty, \quad \forall \quad t \in \mathbb{Z},$
- iii) $\gamma_X(m, n) = \gamma_X(t + m, t + n), \quad \forall \quad t, m, n, \in \mathbb{Z},$

em que função de autocovariância é definida como:

$$\gamma_X(m, n) = \text{Cov}(X_m, X_n) = \mathbb{E}[(X_m - \mu_m)(X_n - \mu_n)].$$

Feita a definição, vamos analisar a região de estabilidade para um processo AR(2).

Seja $\{X_t, t \in \mathbb{Z}\}$ um processo estacionário AR(2), de média nula,

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t, \quad (2.3)$$

em que $\{\epsilon_t, t \in \mathbb{Z}\}$ representa um ruído branco tal que $\mathbb{E}(\epsilon_t) = 0$ e $\text{Var}(\epsilon_t) = \sigma^2$. Pela Equação (2.3), note que o termo atual da série depende diretamente dos dois termos anteriores, juntamente com o erro associado. Além disso, X_t e ϵ_t são denominados variável de saída e de entrada do modelo, respectivamente.

Da Equação (2.3), se $\alpha_2 = 0$, tem-se um modelo AR(1). Por sua estrutura, neste caso, X_t depende do termo anterior mais um erro associado (ϵ_t). No entanto, processos autorregressivos de qualquer ordem podem ser simplificados, usando o operador de defasagem. Dessa forma, o polinômio autorregressivo de ordem p pode ser descrito na forma:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

em que $B^d X_t = X_{t-d}$ é o operador de defasagem de ordem d e p representa a ordem do modelo. Dessa forma, pode-se reescrever a expressão (2.3):

$$\phi_2(B)X_t = \epsilon_t \iff (1 - \phi_1 B - \phi_2 B^2)X_t = \epsilon_t.$$

Assim, para analisar as condições de estacionariedade de um processo AR(2), é importante investigar o polinômio característico. Tome $\phi_1 = \alpha_1$ e $\phi_2 = \alpha_2$ e $\phi_2(B) = 0$, então

$$\phi_2(B) = 1 - \alpha_1 B - \alpha_2 B^2.$$

As soluções do polinômio característico são:

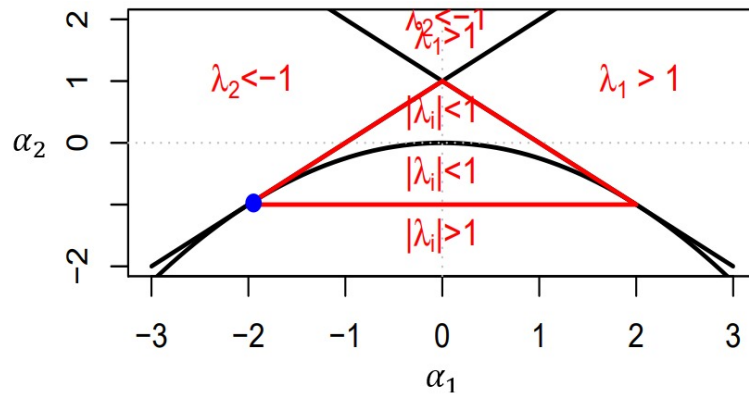
$$\lambda_i = \frac{\alpha_1 \pm \sqrt{(\alpha_1^2 + 4\alpha_2)}}{-2\alpha_2}.$$

Pode-se mostrar que a estacionariedade é obtida se todas as raízes são em módulo maior do que 1. Isto é, todas estão fora do círculo unitário, veja Shumway, Stoffer e Stoffer (2000). Dessa forma, para o AR(2), tem-se

$$\{(\alpha_1, \alpha_2) \in \mathbb{R}^2 : \alpha_1 + \alpha_2 < 1, \alpha_2 - \alpha_1 < 1, |\alpha_2| < 1\}.$$

A Figura 1 ilustra o triângulo que representa a região para estacionariedade de um processo AR(2). A área ilustrada indica os pontos do plano (α_1, α_2) em que as raízes (λ_i) do polinômio característico devem se localizar para garantir a estacionariedade da série.

Figura 1 – Triângulo onde as condições de estacionariedade para AR(2) são satisfeitas



Fonte: Elaboração Própria (2025)

2.3 FUNÇÃO DE AUTOCORRELAÇÃO

A autocorrelação de uma ST avalia como os valores da série estão correlacionados com seus próprios valores em defasagens anteriores. Para o AR(2), o seguinte resultado é válido. Seja X_t uma ST estacionária com média nula (sem perda de generalidade). A sua função de autocovariância é dada por

$$\gamma_k = \alpha_1 \gamma_{k-1} + \alpha_2 \gamma_{k-2}, \quad \text{para } k \geq 1, \quad (2.4)$$

em que $\gamma_k = \text{cov}(X_t, X_{t-k})$ é a autocovariância de ordem k .

Por definição, a autocorrelação de ordem k é dada por $\rho_k = \frac{\gamma_k}{\gamma_0}$, com $\gamma_0 = \text{Var}(X_t)$ para todo k , onde γ_0 é a variância do processo. Assim, dividindo ambos os lados da equação (2.4) por γ_0 , tem-se

$$\frac{\gamma_k}{\gamma_0} = \alpha_1 \left(\frac{\gamma_{k-1}}{\gamma_0} \right) + \alpha_2 \left(\frac{\gamma_{k-2}}{\gamma_0} \right) \quad \therefore \quad \rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2}.$$

Para $k = 1, 2$, as autocorrelações são:

$$\rho_1 = \frac{\alpha_1}{(1 - \alpha_2)} \quad \text{e} \quad \rho_2 = \frac{\alpha_1^2 - \alpha_2^2 + \alpha_2}{(1 - \alpha_2)}.$$

Enquanto a ACF mede a correlação bruta entre X_t e X_{t-k} , a função de autocorrelação parcial (PACF) no lag k , denotada por: $\phi_{k,k}$, quantifica a correlação diretamente atribuída a X_{t-k} , removendo os efeitos dos lag intermediários $X_{t-1}, X_{t-2}, \dots, X_{t-k+1}$.

Para encontrar a variância do modelo, considere a Equação (2.3), reescrita como:

$$X_t - \mu = \alpha_1 (X_{t-1} - \mu) + \alpha_2 (X_{t-2} - \mu) + \epsilon_t,$$

em que μ representa a média incondicional da ST. Multiplicando ambos os lados por $X_t - \mu$ e tomando o valor esperado

$$\begin{aligned}\mathbb{E}[(X_t - \mu)(X_t - \mu)] &= \alpha_1 \mathbb{E}[(X_t - \mu)(X_{t-1} - \mu)] + \alpha_2 \mathbb{E}[(X_t - \mu)(X_{t-2} - \mu)] + \mathbb{E}[\epsilon_t(X_t - \mu)] \\ \iff \mathbb{E}(X_t - \mu)^2 &= \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \mathbb{E}(\epsilon_t^2) \\ \iff \gamma(0) &= \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \mathbb{E}(\epsilon_t^2).\end{aligned}$$

Da expressão:

$$\gamma(0) = \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \mathbb{E}(\epsilon_t^2),$$

tem-se:

$$\gamma(0) = \alpha_1 \left[\frac{\alpha_1 \gamma(0)}{(1 - \alpha_2)} \right] + \alpha_2 \gamma(0) \left[\frac{\alpha_1^2}{1 - \alpha_2} + \alpha_2 \right] + \sigma^2,$$

e, portanto:

$$\gamma(0) = \frac{(1 - \alpha_2)\sigma^2}{(1 + \alpha_2)[1 - (\alpha_1 + \alpha_2)][1 - (\alpha_2 + \alpha_1)]}.$$

No próximo capítulo é apresentado um novo modelo AR(2) com distribuição marginal Lindley.

3 A CONSTRUÇÃO DE UM NOVO MODELO LINDLEY AUTORREGRESSIVO DE SEGUNDA ORDEM

Neste Capítulo é apresentado um novo processo AR(2) com distribuição marginal Lindley. Adicionalmente, diversas propriedades estatísticas e o desenvolvimento das condições de estacionariedade para o novo modelo são apresentados.

Para começarmos a definir o novo processo, é importante estudar a transformada de Laplace como um mecanismo para obtenção da fdp do erro. Uma das vantagens no uso da transformada inversa de Laplace é a capacidade de converter equações diferenciais em forma algébrica, dispensando integração ou métodos mais complexos, veja Metwally (1999).

Seja $X \sim L(\lambda)$, a transformada de Laplace, $\phi_X(s)$, é dada por:

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

As transformadas de Laplace das distribuições marginais são dadas por

$$\phi_X(s) = \mathbb{E}(\exp(-Xs)) \quad \text{e} \quad \phi_\epsilon(s) = \mathbb{E}(\exp(-\epsilon s)).$$

A partir da definição de estacionariedade fraca e utilizando o artifício analítico empregado por Lawrance e Lewis (1980), considere o processo AR(2), definido como:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t,$$

em que $\{\epsilon_t\}$ é um ruído branco. O método consiste em decompor X_t na soma de dois termos:

$$X_t = C + D,$$

em que

$$C = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} \quad \text{e} \quad D = \epsilon_t.$$

Como o ruído D é independente dos termos passados, a função característica de X_t é dada pelo produto das funções características de C e de ϵ_t :

$$\phi_X(s) = \phi_C(s) \times \phi_\epsilon(s).$$

Para determinar $\phi_C(s)$, no caso geral, se $Y = kZ$, então:

$$\phi_Y(s) = \phi_Z(ks).$$

E, se $Y = k_1Z_1 + k_2Z_2$, com Z_1 e Z_2 independentes, então:

$$\phi_Y(s) = \phi_{Z_1}(k_1s) \times \phi_{Z_2}(k_2s).$$

Entretanto, no modelo AR(2), as variáveis X_{t-1} e X_{t-2} não são independentes. Para solucionar isso, Lawrance e Lewis (1980) emprega um artifício baseado na representação do termo C como uma combinação convexa dos estados defasados, ponderada por $(1 - \alpha_2)$ e α_2 . Essa abordagem é possível devido à natureza linear do modelo e à premissa de estacionariedade fraca que garante constância das propriedades estatísticas ao longo do tempo.

Assim, considera-se uma variável auxiliar Z que assume X_{t-1} com probabilidade $(1 - \alpha_2)$ e X_{t-2} com probabilidade α_2 . Sob essa construção, a função característica do termo C é expressa como uma combinação ponderada das funções características de X_{t-1} e X_{t-2} , de modo que:

$$\phi_C(s) = (1 - \alpha_2)\phi_{X_{t-1}}(\alpha_1s) + \alpha_2\phi_{X_{t-2}}(\alpha_2s).$$

Finalmente, combinando com o ruído branco ϵ_t , que é independente dos termos passados, obtém-se a expressão final para a função característica de X_t :

$$\phi_X(s) = \left[(1 - \alpha_2)\phi_{X_{t-1}}(\alpha_1s) + \alpha_2\phi_{X_{t-2}}(\alpha_2s) \right] \phi_\epsilon(s). \quad (3.1)$$

Considerando que a série X_t seja estacionária, reescrevendo a Equação (3.1), é possível obter

$$\phi_X(s) = \left[(1 - \alpha_2)\phi_X(\alpha_1s) + \alpha_2\phi_X(\alpha_2s) \right] \phi_\epsilon(s). \quad (3.2)$$

Da Equação (3.2), tem-se

$$\phi_\epsilon(s) = \frac{\phi_X(s)}{\left[(1 - \alpha_2)\phi_X(\alpha_1s) + \alpha_2\phi_X(\alpha_2s) \right]}. \quad (3.3)$$

Aplicando

$$\phi_X(\alpha_1s) = \frac{\lambda^2(\lambda + \alpha_1s + 1)}{(\lambda + 1)(\lambda + \alpha_1s)^2} \quad \text{e} \quad \phi_X(\alpha_2s) = \frac{\lambda^2(\lambda + \alpha_2s + 1)}{(\lambda + 1)(\lambda + \alpha_2s)^2},$$

em (3.3), tem-se

$$\phi_\epsilon(s) = \frac{(\alpha_2 - 1)(\lambda + \alpha_1s + 1)(\lambda + \alpha_2s)^2}{(\lambda + \alpha_1s)^2\alpha_2(\lambda + \alpha_2s + 1)} + \frac{(\lambda + s + 1)(\lambda + \alpha_2s)^2}{(\lambda + s)^2\alpha_2(\lambda + \alpha_2s + 1)}.$$

Após algumas manipulações analíticas, chega-se a expressão

$$\phi_\epsilon(s) = \frac{(\lambda + s + 1)(\lambda + \alpha_1 s)^2(\lambda + \alpha_2 s)^2}{(\lambda + s)^2[(1 - \alpha_2)(\lambda + \alpha_1 s + 1)(\lambda + \alpha_2 s)^2 + \alpha_2(\lambda + \alpha_2 s + 1)(\lambda + \alpha_1 s)^2]}. \quad (3.4)$$

Por questões de tratabilidade analítica e interpretabilidade, dividiremos a discussão em dois casos:

1. Caso em que os parâmetros autorregressivos são iguais.
2. Caso em que os parâmetros autorregressivos são diferentes.

A partir daí, a condição referente ao *Caso 1* será apresentada a seguir. Para fins de notação, considere:

$$X_t \sim \text{LAR}_1(2).$$

3.1 CASO 1: PROCESSO AR(2) COM DISTRIBUIÇÃO MARGINAL LINDLEY

Para a construção do modelo, será aplicada uma restrição nos parâmetros autorregressivos de modo que se tenha:

$$\alpha_1 = \alpha_2 = a, \quad a \in (0, 1),$$

isto é,

$$X_t = a(X_{t-1} + X_{t-2}) + \epsilon_t.$$

Substituindo essa restrição em (3.4) tem-se:

$$\phi_\epsilon(s) = \frac{(\lambda + s + 1)(\lambda + as)^2(\lambda + as)^2}{(\lambda + s)^2[(1 - a)(\lambda + as + 1)(\lambda + as)^2 + a(\lambda + as + 1)(\lambda + as)^2]},$$

ou equivalentemente,

$$\phi_\epsilon(s) = \frac{(\lambda + s + 1)(\lambda + as)^2}{[(\lambda + s)^2(\lambda + as + 1)]}. \quad (3.5)$$

No que segue, apresenta-se a primeira proposição desta seção. Esta estabelece que a inovação ϵ_t sob a restrição segue uma mistura entre variáveis Gama como enunciado na proposição a seguir.

Proposição 3.1.1 *Sejam $X_t \sim L(\lambda)$ e $\alpha_1 = \alpha_2 = a$. Então, a distribuição de inovação ϵ_t sob a restrição segue uma mistura de variáveis Gama e exponencial.*

$$X_t = a(X_{t-1} + X_{t-2}) + \epsilon_t \quad \text{tal que}$$

$$F_{\epsilon_t}(l) = T_1 F_{\epsilon_{t,1}}(l) + T_2 F_{\epsilon_{t,2}}(l) + T_3 F_{\epsilon_{t,3}}(l),$$

em que $F_{\epsilon_t}(\cdot)$ é a fda de ϵ_t e

$$\epsilon_{t,1} \sim \exp(\lambda), \quad \epsilon_{t,2} \sim \Gamma(2, \lambda) \quad \text{e} \quad \epsilon_{t,3} \sim \exp\left(\frac{1+\lambda}{a}\right).$$

Os pesos, T_1 , T_2 e T_3 são dados por:

$$T_1 = \frac{a^3\lambda + a^2\lambda^2 - 3a^2\lambda - 2a^2 - 2a\lambda^2 + a\lambda + 3a + \lambda^2 + \lambda}{(\lambda - a\lambda + 1)^2},$$

$$T_2 = \frac{(a-1)^2}{(\lambda - a\lambda + 1)} \quad \text{e} \quad T_3 = \frac{-a^2(a-1)}{(\lambda - a\lambda + 1)^2}.$$

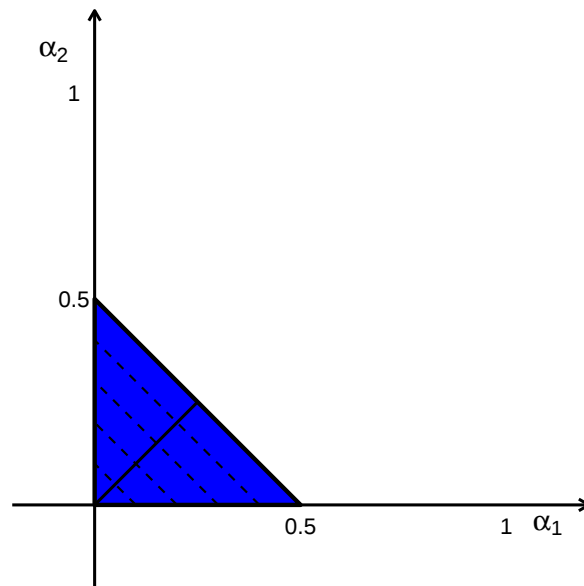
A expressão que representa a soma dos pesos em função dos parâmetros a e λ é determinada por:

$$S(a, \lambda) = \frac{-a^3 + a^2\lambda^2 - 2a\lambda^2 - 2a\lambda + a + \lambda^2 + 2\lambda + 1}{(\lambda - a\lambda + 1)^2}.$$

Com o intuito de obter a distribuição do erro, aplicou-se em $\phi_\epsilon(s)$ a transformada inversa de Laplace, usando o método de frações parciais, técnica essa utilizada em Cálculo Diferencial que facilita a integração de funções racionais (para mais detalhes, ver Peixoto, Renna e Corrêa (2013)). Como a distribuição de ϵ_t é uma mistura, tem-se que a soma dos pesos T_1 , T_2 e T_3 é igual a 1. A demonstração da Proposição 3.1.1 pode ser conferida no Apêndice A.

A Figura 2 representa a região do novo processo onde as condições de estacionariedade são satisfeitas.

Figura 2 – Triângulo de estabilidade em um processo AR(2), Caso 1

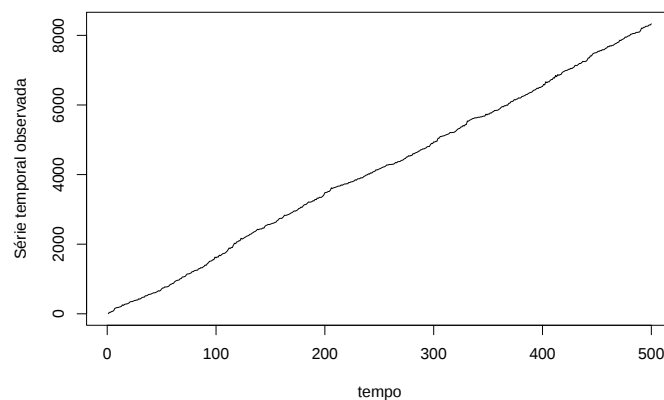


Fonte: Elaboração Própria (2025)

A área em destaque no primeiro quadrante do plano cartesiano representa a região de (α_1, α_2) , com α_1 pertencente ao eixo das abscissas e α_2 ao eixo das ordenadas, na qual todas as raízes estão fora do círculo unitário.

Para o modelo $LAR_1(2)$, uma das condições para o sistema ser estacionário, refere-se aos parâmetros autorregressivos do modelo. Os resultados computacionais via software R indicam que a série exibe um comportamento estacionário quando se tem $\alpha_1 = \alpha_2 = a < 0.49$, com $p\text{-valor} < 0.05$. No entanto, a partir de $a \geq 0.5$ a série perde sua estacionariedade. Sendo assim, a Figura 3 apresenta um modelo de série não estacionária com um $a > 0.5$.

Figura 3 – Exemplo de uma série não estacionária



Fonte: Elaboração Própria (2025)

Assim, para estudar o modelo é imprescindível que as análises feitas estejam coerentes com a estacionariedade, diretamente influenciada pelo parâmetro autorregressivo em função de λ . A proposição a seguir refere-se ao espaço paramétrico para o parâmetro λ em função de a .

Proposição 3.1.2 *Dada a soma dos pesos $T_1 + T_2 + T_3 = 1$ e fixando $\epsilon > 0$, a restrição de λ é obtida através da resolução da expressão:*

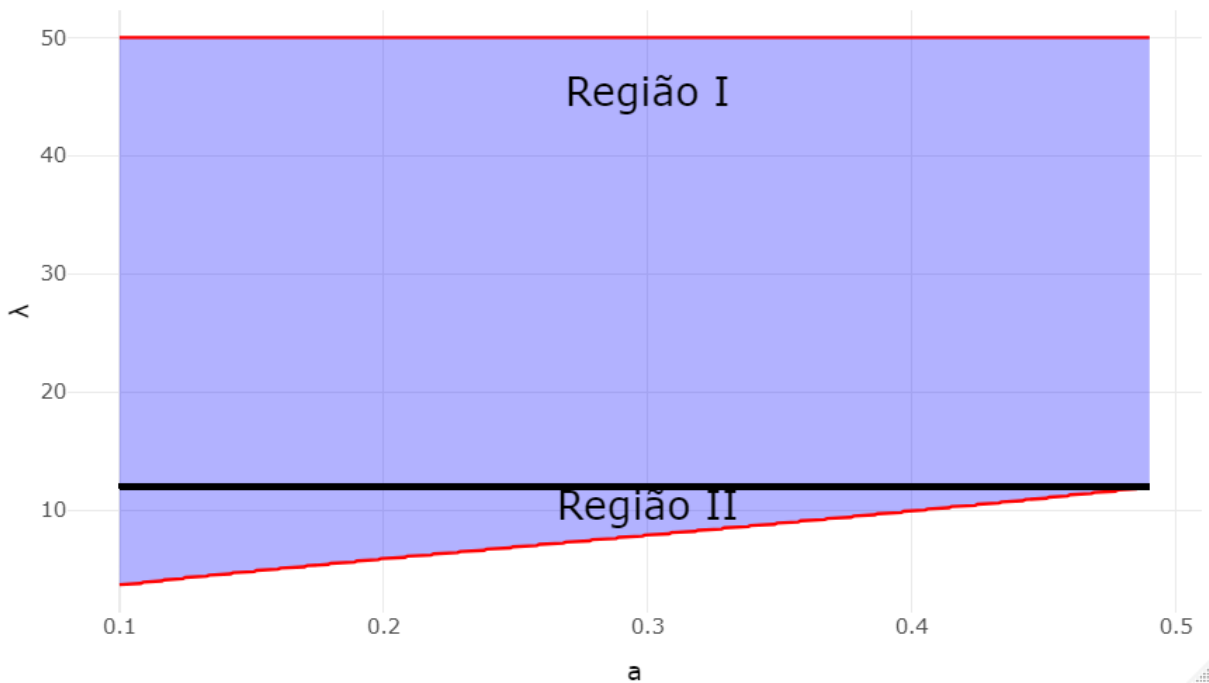
$$S_{\epsilon}(\lambda) > \frac{1}{D\epsilon} \times \left[\sqrt{(A\epsilon - B)^2 - \epsilon(E\epsilon - F) + C} - (A\epsilon - B) \right],$$

em que:

$$\begin{aligned} A &= 1148417904979476480000, & B &= 4399, \\ C &= 1171386263079066009600, & D &= 585693131539533004800, \\ E &= 1125899906842624000000, & F &= 562724773439943475200. \end{aligned}$$

A demonstração da Proposição 3.1.2 pode ser conferida no Apêndice B. A seguir, tem-se a Figura 4 que representa a região de λ que satisfaz essa proposição.

Figura 4 – Espaço paramétrico do parâmetro λ .



Fonte: Elaboração Própria (2025)

Contudo, a região padrão de estacionariedade também sofre o efeito do parâmetro λ da Lindley.

Dessa maneira, pode-se também estimar o espaço paramétrico para o λ , factível com a estacionariedade, por meio de expansão em série de Fourier. Esse procedimento consiste em expandir uma função periódica com uma combinação infinita de senos e cossenos, para mais detalhes, veja Arfken, Weber e Harris (2011). Sua expressão é dada por:

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)),$$

em que

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \quad \text{e} \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx.$$

Para a região II da Figura (4) é necessário definir de forma aproximada a função que a limita inferiormente e será apresentada na proposição a seguir.

Proposição 3.1.3 *A função aproximada que define inferiormente a região II da Figura 4 em relação a λ é dada por:*

$$\begin{aligned} L(a) = & 703.863523692193 \sin(2\pi x) - 799.644435128453 \sin(4\pi x) + 797.199224350749 \sin(6\pi x) \\ & - 392.00206957564 \sin(8\pi x) - 61.6077748356555 \sin(10\pi x) - 446.845128375007 \sin(12\pi x) \\ & - 1363.47210940117 \sin(14\pi x) + 440.732346645609 \sin(16\pi x) - 1245.32812300171 \sin(18\pi x) \\ & - 2248.44810011216 \sin(20\pi x) - 64.2959366539768 \sin(22\pi x) + 719.481922561717 \sin(24\pi x) \\ & + 171.364031399465 \sin(26\pi x) - 41.4166663640313 \sin(28\pi x) - 9.4311075853508 \sin(30\pi x) \\ & + 1838.32686151142 \cos(2\pi x) + 918.422035617075 \cos(4\pi x) + 883.419337345749 \cos(6\pi x) \\ & + 1789.54093010238 \cos(8\pi x) + 782.432193288363 \cos(10\pi x) + 1892.03471892147 \cos(12\pi x) \\ & + 263.447314628281 \cos(14\pi x) + 493.838702930015 \cos(16\pi x) + 1823.73378462976 \cos(18\pi x) \\ & - 650.274127943238 \cos(20\pi x) - 1578.98362034074 \cos(22\pi x) - 270.44043063636 \cos(24\pi x) \\ & + 218.489780868658 \cos(26\pi x) + 57.6826911620093 \cos(28\pi x) - 3.570012347606 \cos(30\pi x) \\ & + 3.78669192483672. \end{aligned}$$

A expressão acima, se refere a expansão de série de Fourier utilizando 15 termos. Desse modo, para o modelo $LAR_1(2)$, a curva limite do espaço paramétrico de λ admitindo a região de estabilidade, para o parâmetro autorregressivo $a = 0.49$, é tal que:

$$\lambda = f(a) = 12.$$

Já para $a = 0.1$ e $f(a) = 3.99$, para esse valor específico do parâmetro autorregressivo, todo valor de λ a partir de 3.99 consegue-se obter uma mistura. A demonstração da Proposição

3.1.3 foi realizada computacionalmente utilizando a linguagem Python. O passo a passo para a construção da função é dado a seguir.

1. Inicialmente, escolheu-se 100 pares ordenados do tipo (a, λ) de acordo com a função soma determinada pela soma dos pesos da mistura.
2. O código pede para o usuário determinar o número de termos da série.
3. Construção da série de Fourier: Foi definido a variável x e a constante a_0 .
4. Para cada termo da série indicada no passo 2, é acrescentado um termo cossenoidal e senoidal do tipo $(a_n \cos \pi n x)$ e $(b_n \sin \pi n x)$, respectivamente. A partir daí, a expressão simbólica da série é gerada e armazenada.
5. Nessa etapa, a expressão simbólica é convertida em função numérica que permite calcular os valores de λ para qualquer valor do parâmetro a .
6. Por meio da função `curve_fit`, a função é ajustada, iniciando os parâmetros com valores nulos, com os ajustes retornando os melhores parâmetros.
7. A penúltima etapa consiste em exibir os parâmetros ajustados.
8. Por fim, com os parâmetros ajustados, calcula-se os valores de λ para cada valor de a dentro do intervalo $[0.1, 0.49]$, que corresponde à região de estabilidade para o modelo.

3.1.1 Propriedades Estruturais do novo modelo

Após a construção do novo modelo AR(2) com distribuição marginal Lindley, é necessário investigar suas propriedades como o gerador do erro da nova distribuição, a densidade do novo processo e determinar a esperança e variância condicional. Através da transformada inversa de Laplace, foi possível encontrar os coeficientes da distribuição da mistura, vistos na seção anterior. Desse modo, utilizou-se o software Matlab para encontrar a transformada inversa de Laplace. No R, foi feita a densidade do processo e o gerador para a inovação. Após isso, construíram-se os gráficos da série, autocorrelação, autocorrelação parcial, para 2 cenários diferentes, e densidade do erro, considerando a região de estabilidade.

3.1.2 Propriedades não condicionais

Nesta Subsessão, são apresentadas algumas propriedades não condicionais do modelo $\text{LAR}_1(2)$.

Pode-se redefinir um $\text{AR}(2)$ usando a equação vetorial:

$$\begin{bmatrix} X_t \\ X_{t-1} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X_{t-1} \\ X_{t-2} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ 0 \end{bmatrix}. \quad (3.6)$$

Utilizando os resultados apresentados em Hamilton (2020), considere: $f_{i,j}^{(k)}$ a entrada (i, j) de F^t , que é a matriz de transição elevada a t , sendo

$$\mathbf{F} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 0 \end{bmatrix}.$$

Com k o número de passos ao longo do tempo t , quando $k = t$ a matriz F^t fornece a dinâmica do sistema após t períodos de tempo. Tomando $f_{i,j}^{(k)}$ a entrada (i, j) de F na Equação (3.6) e adotando $f_{1,1}^{(0)} = 1$

Por substituições recursivas, tem-se:

$$X_t = f_{1,1}^{(t+1)} X_{t-1} + f_{1,2}^{(t+1)} X_{t-2} + \sum_{i=0}^t f_{1,1}^{(i)} \epsilon_{t-i}. \quad (3.7)$$

Da Equação 3.7, $\mathbb{E}|X_{t-k}| < \infty$ pela suposição de estacionariedade, e

$$\lim_{k \rightarrow \infty} f_{1,i}^{(k+1)} = 0,$$

$$\begin{aligned} \mathbb{E}(X_t) &= \lim_{t \rightarrow \infty} \mathbb{E}(X_t) = \mathbb{E}(X_{t-1}) \lim_{t \rightarrow \infty} f_{1,1}^{(t+1)} + \mathbb{E}(X_{t-2}) \lim_{t \rightarrow \infty} f_{1,2}^{(t+1)} + \sum_{i=0}^{\infty} f_{1,1}^{(i)} \mathbb{E}(\epsilon_{t-i}) \\ &= \sum_{i=0}^{\infty} f_{1,1}^{(i)} \mathbb{E}(\epsilon_{t-i}) = \mathbb{E}(\epsilon_t) \left[\sum_{i=0}^{\infty} f_{1,1}^{(i)} \right], \end{aligned}$$

com:

$$E(\epsilon_t) = \sum_{i=1}^3 T_i \mathbb{E}(\epsilon_{N,0i}),$$

cujos os pesos são dados conforme a Proposição 3.1.1.

De acordo com Brockwell e Davis (1991), a variância de X_t para um $\text{AR}(2)$, é dada por:

$$\text{Var}(X_t) = \alpha_1 \gamma_X(1) + \alpha_2 \gamma_X(2) + \text{Var}(\epsilon_t), \quad (3.8)$$

em que a função de autocovariância de X_t , $\gamma_X(h) = \text{Cov}(X_{t+h}, X_t)$, obedece a equação diferencial homogênea:

$$\gamma_X(h) = \alpha_1 \gamma_X(h-1) + \alpha_2 \gamma_X(h-2), \quad (3.9)$$

a partir do qual, após algumas manipulações algébricas,

$$[\gamma_X(1), \gamma_X(2)] = \left[\frac{\alpha_1}{1-\alpha_2}, \alpha_2 + \frac{\alpha_1^2}{1-\alpha_2} \right] \times \text{Var}(X_t).$$

Inserindo a última identidade na Equação (3.8),

$$\text{Var}(X_t) = \frac{(1-\alpha_2)\text{Var}(\epsilon_t)}{(1+\alpha_2)[1-(\alpha_2+\alpha_1)][1-(\alpha_2-\alpha_1)]},$$

que equivale a

$$\text{Var}(\epsilon_t) = \frac{(1+\alpha_2)[1-(\alpha_2+\alpha_1)][1-(\alpha_2-\alpha_1)]\text{Var}(X_t)}{1-\alpha_2}.$$

3.1.3 Propriedades Condicionais

Nesta subsecção vamos determinar o valor esperado condicional e a variância condicional para o modelo AR(2).

Tomando $f_{i,j}^{(k)}$ a entrada a entrada (i, j) de F na Equação (3.6) e adotando $f_{1,1}^{(0)} = 1$, tem-se a partir de substituições recursivas:

$$X_{t+k} = f_{1,1}^{(k)} X_t + f_{1,2}^{(k)} X_{t-1} + \sum_{i=1}^k f_{1,1}^{(k-i)} \epsilon_{t+i}. \quad (3.10)$$

Agora, seja $\{X_t; t = 1, \dots, N\}$ N-pontos de uma série temporal. Tomando o valor esperado condicional $\mathbb{E}(\cdot | X_1, \dots, N)$ em ambos os lados da Equação (3.10) para $t = N$: Brockwell e Davis (1991),

$$\begin{aligned} \mathbb{E}(X_{N+k} | X_N, X_{N-1}) &= f_{1,1}^{(k)} X_N + f_{1,2}^{(k)} X_{N-1} + \sum_{i=1}^k f_{1,1}^{(k-i)} \mathbb{E}(\epsilon_{N+i}) \\ &= f_{1,1}^{(k)} X_N + f_{1,2}^{(k)} X_{N-1} + \mathbb{E}(\epsilon_N) \sum_{i=1}^k f_{1,1}^{(k-i)}, \end{aligned}$$

que é a equação de previsão linear para k passos à frente. Para $k = 1$ (previsão de um passo à frente), tem-se:

$$\mathbb{E}(X_{N+1}|X_N, X_{N-1}) = f_{1,1}^{(1)}X_N + f_{1,2}^{(1)}X_{N-1} + f_{1,1}^{(0)}\mathbb{E}(\epsilon_N) \quad (3.11)$$

$$= \alpha_1 X_N + \alpha_2 X_{N-1} + \mathbb{E}(\epsilon_N). \quad (3.12)$$

Admitindo o processo estacionário tem-se:

$$\mathbb{E}(\epsilon_N) = (1 - \alpha_1 - \alpha_2)\mathbb{E}(X_t). \quad (3.13)$$

Considerando o Caso 1, $\alpha_1 = \alpha_2 = a$, vale-se:

$$\mathbb{E}(\epsilon_N) = (1 - 2a)\mathbb{E}(X_t).$$

Assim, a Equação (3.11) pode ser reescrita da forma:

$$\mathbb{E}(X_{N+1}|X_N, X_{N-1}) = a(X_N + X_{N-1}) + \mathbb{E}(\epsilon_N).$$

De maneira semelhante, é possível encontrar a variância condicional. Aplicando a variância $\text{Var}(\cdot|X_1, \dots, N)$ na Equação (3.10):

$$\text{Var}(X_{N+k} | X_N, X_{N-1}) = \sum_{i=1}^k [f_{1,1}^{(k-i)}]^2 \text{Var}(\epsilon_{N+i}) = \text{Var}(\epsilon_N) \sum_{i=1}^k [f_{1,1}^{(k-i)}]^2$$

e para $k = 1$, $\text{Var}(X_{N+k} | X_N, X_{N-1}) = \text{Var}(\epsilon_N)$.

3.1.4 Autocorrelação, Autocorrelação Parcial e Densidade do erro

Conforme definido anteriormente, a ACF para um processo AR(2) é dada por:

$$\rho_k = \alpha_1 \rho_{k-1} + \alpha_2 \rho_{k-2}, \quad \text{para } k > 0, \quad (3.14)$$

com:

$$\rho_1 = \frac{\alpha_1}{1 - \alpha_2} \quad \text{e} \quad \rho_2 = \frac{\alpha_1^2 - \alpha_2^2 + \alpha_2}{1 - \alpha_2}.$$

Para a ACF, valem as seguintes relações:

$$\phi_{1,1} = \rho_1 = \frac{\alpha_1}{1 - \alpha_2} \neq 0, \quad \phi_{2,2} = \alpha_2 \neq 0, \quad \phi_{k,k} = 0, \quad \forall k \geq 3.$$

Note que a função de autocorrelação parcial (PACF), para um AR(2) estacionário, é diferente de 0 até a ordem 2 e igual a 0 para ordens superiores a 2. Desse modo, ao interpretar

os gráficos das ACF e PACF consegue-se identificar a ordem do modelo da ST, pela estrutura de dependência temporal dos dados.

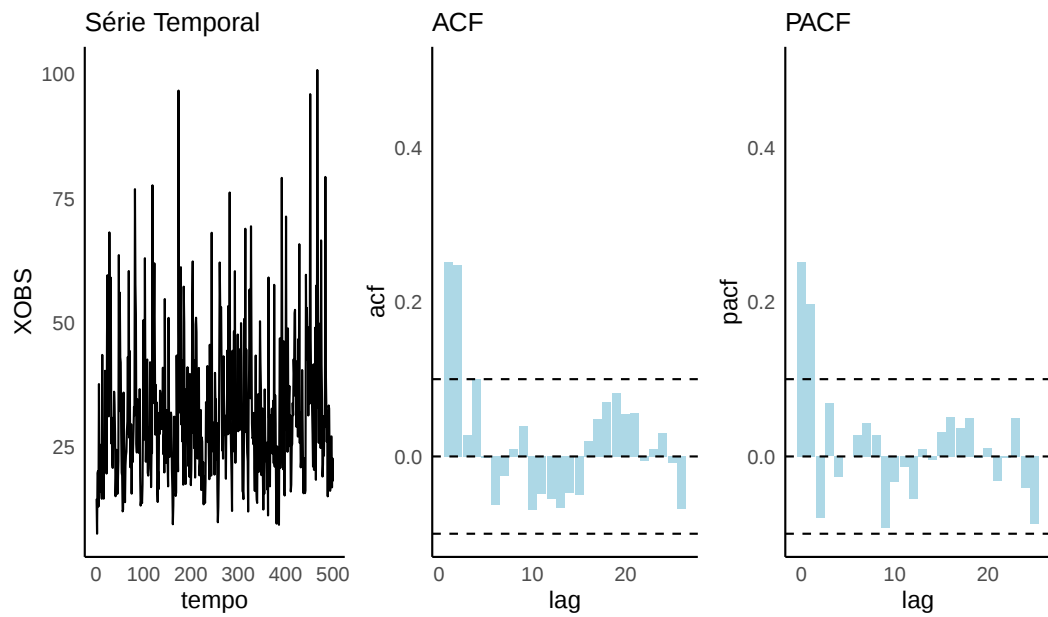
Em vista disso, para um modelo $AR(2)$, o gráfico da ACF apresenta um pico no $lag\ 2$, além de ter um pico significativo, também no $lag\ 2$, para o gráfico de autocorrelação parcial. Como visto anteriormente, nota-se que, para o modelo $LAR_1(2)$, a série é estacionária no intervalo $0 < a < 0.49$. A partir de $a = 0.5$ “estoura” a não estacionariedade. Além disso, é necessário analisar o espaço paramétrico para o parâmetro λ , realizado anteriormente, gerando as Proposições 3.1.2 e 3.1.3.

Desse modo, foi construído números aleatórios para produzir a ST que corresponde ao novo modelo. Assim, uma função foi implementada com os parâmetros do $LAR_1(2)$, dependente dos coeficientes da mistura, de acordo com a Proposição 3.1.1. As Figuras 5 e 6 representam os gráficos da série, da ACF e PACF, nessa ordem, de 2 cenários para séries de tamanho $N = 500$.

Diante disso, nota-se que o gráfico da PACF apresenta significância estatística em dois atrasos. Ou seja, tem-se dois primeiros traços verticais fora do intervalo, indicando que há evidências de que se trata de um processo $AR(2)$. Já os gráficos gerados para a ACF, apresentam um decaimento exponencial e comportamento semelhante aos das funções seno e cosseno.

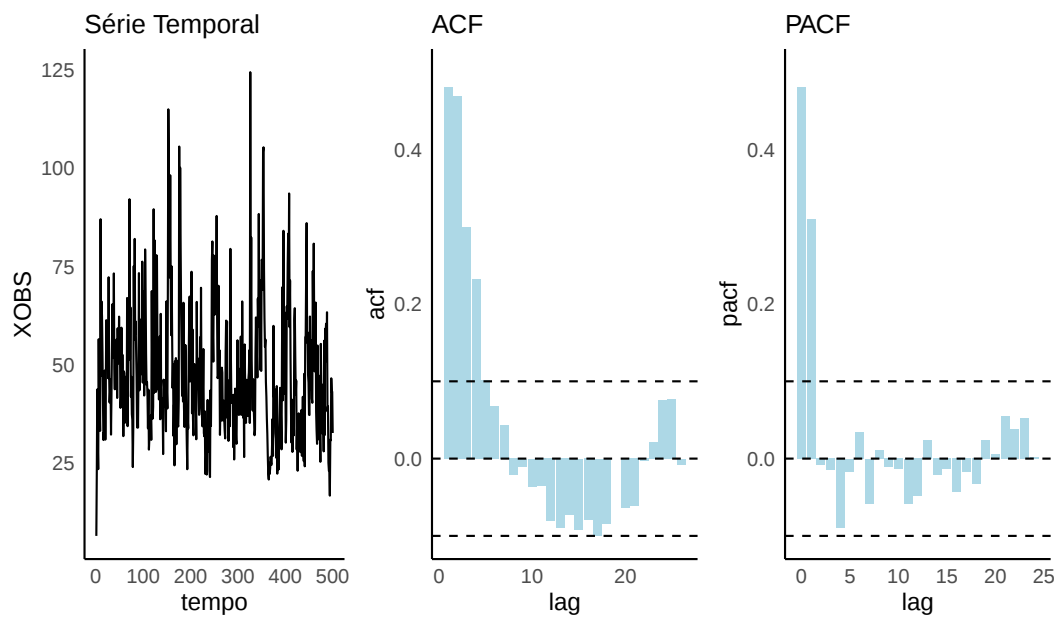
A Figura 7 exibe curvas da densidade do erro. Aqui, fixou-se o λ e variou-se o a . Note que, à medida que aumenta o valor do parâmetro autorregressivo, a distribuição tende a ter caudas mais pesadas, apresentando pouca variabilidade.

Figura 5 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 para os valores $\alpha_1 = \alpha_2 = a = 0.2, \lambda = 12$



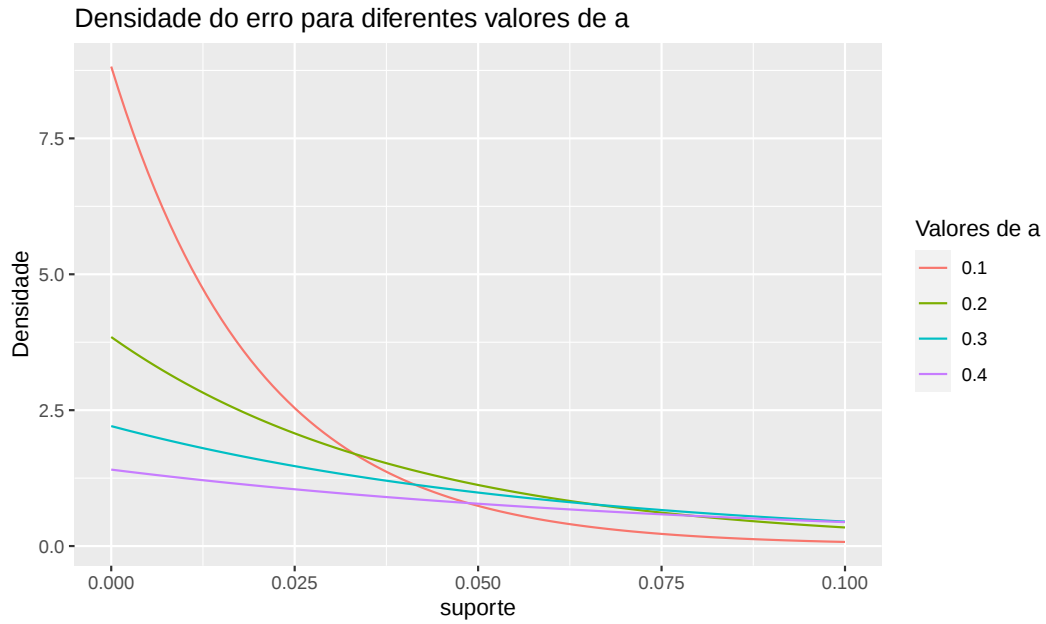
Fonte: Elaboração Própria (2025)

Figura 6 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 para os valores $\alpha_1 = \alpha_2 = a = 0.3, \lambda = 12$



Fonte: Elaboração Própria (2025)

Figura 7 – Gráfico da densidade do erro para amostra de tamanho 500, com $\lambda = 12$, variando o parâmetro a .



Fonte: Elaboração Própria (2025)

3.1.5 Previsão

Processos autorregressivos são amplamente utilizados na abordagem de previsão de séries temporais. Através de um modelo bem ajustado de uma ST, é possível entender a estrutura de dependência e fazer previsão.

Desse modo, por meio dos dados anteriores de uma série (passado), consegue-se deduzir os valores futuros por meio de previsões. Assim, se uma dada observação de uma série X for observada num instante t (X_t), consegue-se obter a previsão para o instante $t + h$. Logo, a observação para o período h passos a frente pode ser denotada por X_{t+h} . Considere uma ST AR(2) do tipo

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t.$$

De acordo com a teoria de Box et al. (2015), a previsão um passo a frente de X_2 , tendo como referência X_1 , é dada por:

$$\hat{X}_{2|1} = X_1 \left(\frac{\gamma(1)}{\gamma(0)} \right).$$

Assim, o preditor para o processo h movimentos a frente é dado por

$$\hat{X}_{t+h|t} = \alpha_1 X_{t+h-1} + \alpha_2 X_{t+h-2}.$$

Na previsão de séries temporais, é importante distinguir componentes determinísticos e estocásticos, além de caracterizar a estrutura de dependência temporal (HAMILTON, 2020). Diante disso, os processos Markovianos surgem como ferramentas fundamentais. Formalmente, um processo estocástico X_t é dito Markoviano de ordem 1 se satisfaz a propriedade:

$$\mathbb{P}(X_t \leq x \mid X_{t-1}, X_{t-2}, \dots) = \mathbb{P}(X_t \leq x \mid X_{t-1}),$$

isso significa que X_t depende do estado imediatamente anterior X_{t-1} , e não de valores mais distantes no passado ($X_{t-2}, X_{t-3}, \dots$). De maneira similar, para um processo markoviano de segunda ordem X_t depende apenas dos dois estados imediatamente anteriores (X_{t-1} e X_{t-2}), e não de todo o histórico passado, com :

$$\mathbb{P}(X_t \leq x \mid X_{t-1}, X_{t-2}, X_{t-3}, \dots) = \mathbb{P}(X_t \leq x \mid X_{t-1}, X_{t-2}).$$

Em comparação com modelos AR(1), os processos AR(2) se destacam-se por sua capacidade de capturar dinâmicas mais complexas, visto que incorporam não apenas o valor do período imediatamente anterior, mas também o de dois períodos passados. Essa estrutura permite uma representação mais rica das dependências temporais, sendo capaz de modelar não apenas relações de curto prazo, restritas a modelos AR(1), como também padrões que se estendem por intervalos mais longos. Essa característica é particularmente relevante em séries temporais onde a influência de *lags* mais distantes desempenha um papel significativo na determinação dos valores futuros.

Uma vantagem relevante dos modelos AR(2) consiste em sua capacidade de representar comportamentos oscilatórios, como ciclos e flutuações periódicas, algo que um AR(1) não é capaz de representar de forma adequada. Enquanto o AR(1) se restringe a comportamentos de decaimento exponencial ou tendências monotônicas, o AR(2), em virtude de sua estrutura de raízes características, é capaz de produzir dinâmicas oscilatórias amortecidas. Essa característica torna-se particularmente relevante em áreas como meteorologia e hidrologia, nas quais os fenômenos naturais costumam apresentar padrões cíclicos decorrentes da influência de múltiplos *lags*.

Além disso, os modelos AR(2) se mostram mais adequados quando a ST apresenta dependências mais prolongadas, em que o valor atual é afetado não apenas pelo *lag* 1, mas também pelo *lag* 2. Nessas situações, o AR(1) tende a subestimar a influência de observações mais afastadas, o que pode gerar previsões menos precisas. Ao incorporar explicitamente esses efeitos, o AR(2) oferece um ajuste mais aderente à verdadeira estrutura dos dados, elevando a qualidade das inferências e das previsões.

Outro aspecto importante é a redução de erros de previsão associada aos modelos AR(2). Ao incorporar mais informação passada, esses modelos tendem a apresentar menores resíduos em comparação com os AR(1), resultando em estimativas mais precisas. Essa vantagem é essencial em contextos onde pequenos ganhos de acurácia podem ter implicações operacionais significativas, como na previsão de eventos climáticos extremos ou na gestão de recursos hídricos, áreas em que ganhos incrementais na precisão das previsões impactam diretamente a eficácia das estratégias de prevenção e controle.

Considerando esses aspectos, a escolha entre modelos AR(1) e AR(2) deve basear-se na análise da ACF e nas características específicas da ST. Quando há indícios claros de relevância no *lag* 2 ou a presença de padrões oscilatórios, o modelo AR(2) se destaca como a opção mais adequada. Sua habilidade em representar dinâmicas mais complexas e em diminuir a incerteza nas previsões, torna o modelo AR(2) mais adequado para situações em que a precisão das previsões e o ajuste aos dados observados são fatores decisivos.

3.1.6 Estimação do Processo LAR₁(2)

Com o intuito de estimar os parâmetros do modelo LAR₁(2), utilizou-se os métodos Yule-Walker e MQC. Para o primeiro método, construiu-se a forma fechada para estimar os parâmetros autorregressivos do modelo. E, por meio do segundo método, estimou-se o parâmetro λ . Finalmente, realiza-se uma simulação de Monte Carlo para avaliar o desempenho dos estimadores.

3.1.6.1 Método Yule-Walker

Esse método é utilizado em ST estacionárias para estimar os parâmetros AR em termos das autocorrelações amostrais. Assim, reescrevendo a Equação (3.14) tem-se

$$\rho_X(h) = \alpha_1 \rho_X(h-1) + \alpha_2 \rho_X(h-2) + \frac{\mathbb{I}_0(h) \gamma_\epsilon(0)}{\sigma^2 \gamma_X(0)},$$

em que o termo $\mathbb{I}_0(h)$ representa a função indicadora definida como:

$$\mathbb{I}_0(h) = \begin{cases} 1, & \text{se } h = 0, \\ 0, & \text{se } h \neq 0. \end{cases}$$

Vamos encontrar os valores das estimativas dos parâmetros em função das autocorrelações amostrais. Avaliando a expressão acima para $h = 0$, considere ainda $\hat{\alpha}_1$ e $\hat{\alpha}_2$ são as esti-

mativas para os parâmetros α_1 e α_2 , respectivamente. As autocorrelações amostrais entre as observações separadas por um intervalo de 1 e 2 períodos são determinadas por $\tilde{\rho}(1)$ e $\tilde{\rho}(2)$, respectivamente. Dessa forma, segue que:

$$\begin{cases} \tilde{\rho}_X(1) &= \hat{\alpha}_1 \tilde{\rho}_X(0) + \hat{\alpha}_2 \tilde{\rho}_X(1), \\ \tilde{\rho}_X(2) &= \hat{\alpha}_1 \tilde{\rho}_X(1) + \hat{\alpha}_2 \tilde{\rho}_X(0). \end{cases}$$

Logo, da primeira equação tem-se:

$$\hat{\alpha}_1 = \tilde{\rho}_X(1)[1 - \alpha_2].$$

Resolvendo o sistema:

$$\tilde{\rho}_X(2) = \tilde{\rho}_X(1)(1 - \alpha_2)\tilde{\rho}_X(1) + \alpha_2 \iff \tilde{\rho}_X(2) = \tilde{\rho}_X^2(1) - \hat{\alpha}_2 \tilde{\rho}_X^2(1) + \alpha_2.$$

Isolando o $\hat{\alpha}_2$, obtém-se: $\hat{\alpha}_2 = \frac{\tilde{\rho}_X(2) - \tilde{\rho}_X^2(1)}{1 - \tilde{\rho}_X^2(1)}$.

De $\hat{\alpha}_1 = \tilde{\rho}(1)[1 - \alpha_2]$, substituindo a expressão encontrada para $\hat{\alpha}_2$:

$$\hat{\alpha}_1 = \frac{\tilde{\rho}_X(1) - \tilde{\rho}_X(1)\tilde{\rho}_X(2)}{1 - \tilde{\rho}_X^2(1)}.$$

Assim, obtêm-se as formas fechadas dos parâmetros AR em função das autocorrelações amostrais. Além disso, como a distribuição da v.a ϵ_t é uma mistura, consegue-se também encontrar um estimador para o parâmetro λ . É sabido que:

$$\begin{cases} \gamma_X(0) = \frac{\gamma_\epsilon(0)}{1 - \hat{\alpha}_1 \tilde{\rho}(1) - \hat{\alpha}_2 \tilde{\rho}(2)}, \\ \sigma^2[1 - \hat{\alpha}_1 \tilde{\rho}(1) - \hat{\alpha}_2 \tilde{\rho}(2)] = \gamma_\epsilon(0), \end{cases}$$

em que $\gamma_\epsilon(0)$ a variância da mistura que é composta por distribuições Gama e exponenciais.

Para derivar as equações de Yule-Walker, pode-se expressar o processo AR(2) como uma média móvel infinita (MA(∞)) dos choques passados ϵ_t :

$$X_t = \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}.$$

Os coeficientes ψ_i são definidos recursivamente por:

$$\psi_0 = 1, \quad \psi_1 = \alpha_1, \quad \psi_i = \alpha_1 \psi_{i-1} + \alpha_2 \psi_{i-2} \text{ para } i \geq 2.$$

Esses coeficientes capturam o efeito persistente de choques passados ϵ_{t-i} sobre X_t .

Do mesmo modo, multiplicando ambos os lados da equação por X_{t-h} , para $h \geq 0$, e tomando o valor esperado:

$$\begin{aligned}\mathbb{E}(X_t, X_{t-h}) &= a[\mathbb{E}(X_{t-1}, X_{t-h}) + \mathbb{E}(X_{t-2}, X_{t-h})] + \mathbb{E}(\epsilon_t, X_{t-h}) \\ \Leftrightarrow \gamma_X(h) &= a[\gamma_X(h-1) + \gamma_X(h-2)] + \mathbb{E}\left(\epsilon_t \sum_{i=0}^{\infty} \psi_i \epsilon_{t-h-i}\right) \\ \Leftrightarrow \gamma_X(h) &= a[\gamma_X(h-1) + \gamma_X(h-2)] \\ \Leftrightarrow \rho_X(h) &= a[\rho_X(h-1) + \rho_X(h-2)].\end{aligned}$$

Desse modo, tem-se o estimador para a :

$$\hat{a} = \begin{cases} \frac{\tilde{\rho}(1)}{[1+\tilde{\rho}(1)]}, & h = 1 \\ \frac{\tilde{\rho}(2)}{[\tilde{\rho}(1)+1]}, & h = 2 \\ \frac{\tilde{\rho}(3)}{\tilde{\rho}(2)+\tilde{\rho}(1)}, & h = 3. \end{cases} \quad (3.15)$$

Os termos $\tilde{\rho}$ e h representam autocorrelação amostral e a defasagem, respectivamente. Assim, de acordo com a Expressão (3.15), consegue-se aplicar essas estimativas para realizar a previsão do Caso 1 do modelo AR(2).

Para N suficientemente grande, os estimadores Yule-Waker convergem em distribuição para a distribuição normal, veja Brockwell e Davis (1991). Especificamente,

$$\sqrt{N}(\hat{a} - a) \xrightarrow{D} N(\mathbf{0}, \sigma_\epsilon^2 \cdot \Gamma_2^{-1}) \quad \text{e} \quad \hat{\sigma}_\epsilon^2 \xrightarrow{P} \sigma_\epsilon^2. \quad (3.16)$$

Na Expressão (3.16), Γ_2 representa a matriz de autocovariância da série até o lag 2. Sendo:

$$\Gamma_2 = \begin{pmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{pmatrix},$$

em que γ_0 é a variância da série e γ_1 representa a autocovariância no lag 1. Consequentemente:

$$\sigma_\epsilon^2 \Gamma_2^{-1} = \frac{\sigma_\epsilon^2}{\gamma(0)^2 - \gamma(1)^2} \begin{pmatrix} \gamma(0) & -\gamma(1) \\ -\gamma(1) & \gamma(0) \end{pmatrix}.$$

Para encontrar o estimador de λ , considere:

$$\gamma_X(h) = a[\gamma_X(h-1) + \gamma_X(h-2)] + \mathbb{I}_0(h)\sigma_\epsilon^2.$$

Dividindo ambos os lados por $\text{Var}(X_t)$ e avaliando em $h = 0$:

$$\begin{aligned}\rho_X(h) &= a[\rho_X(h-1) + \rho_X(h-2)] + \frac{\sigma_\epsilon^2}{\sigma^2(X_t)} \\ \rho_X(0) &= a[\rho_X(-1) + \rho_X(-2)] + \frac{\sigma^2\epsilon}{\sigma^2(X_t)} \\ \frac{\sigma^2\epsilon}{\sigma^2(X_t)} &= 1 - a[\rho_X(1) + \rho_X(2)],\end{aligned}$$

em que $\sigma^2\epsilon$, denota a variância da mistura das distribuições encontradas para o erro e $\sigma^2(X_t)$ representa a variância da distribuição Lindley.

Dessa forma, o estimador para o parâmetro λ pode ser:

$$1 - \hat{a}[\tilde{\rho}_X(1) + \tilde{\rho}_X(2)] = \frac{\sigma^2\epsilon}{\sigma^2(X_t)},$$

em que a variância da mistura da distribuição vem diretamente da 3.1.1, sendo denotada por:

$$\sigma^2\epsilon = \frac{T_1^2}{\hat{\lambda}^2} + \frac{2T_2^2}{\hat{\lambda}^2} + \frac{T_3^2\hat{a}}{(1 + \hat{a})^2}.$$

Em vista da intratabilidade das expressões, não é possível obter uma forma fechada para o parâmetro λ . Porém, pode-se definir a expressão:

$$1 - \hat{a}[\tilde{\rho}_X(1) + \tilde{\rho}_X(2)] = \phi(\lambda), \quad (3.17)$$

em que $\phi(\lambda)$ descreve a razão entre variância de X_t e variância de ϵ_t , com:

$$\phi(\lambda) = \frac{\text{Var}(\epsilon_t)}{\text{Var}(X_t)}.$$

Sendo assim, a Equação (3.17) representa uma função de λ que depende do estimador do parâmetro autorregressivo $\hat{a}$ com forma fechada em (3.15).

3.1.6.2 Mínimos Quadrados Condicionais

Por meio de uma ST observada, consegue-se estimar os parâmetros do modelo $\text{LAR}_1(2)$ através do método dos MQC, para mais detalhes, consulte Klimko e Nelson (1978). Este artifício de estimação é fundamentado na minimização da soma dos quadrados dos desvios entre os valores observados e os valores esperados condicionais das observações. Sendo assim, tem-se:

$$Q_n(\alpha_1, \alpha_2, \lambda) = \sum_{t=3}^n [X_t - \mathbb{E}(X_t | X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2})^2].$$

Para o caso geral, temos:

$$Q_n(\alpha_1, \alpha_2, \lambda) = \sum_{t=3}^n \left(X_t - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - (1 - \alpha_1 - \alpha_2) \frac{(2 + \lambda)}{\lambda + \lambda^2} \right)^2.$$

Já em particular, considerando o *Caso 1*, em que os termos autorregressivos são iguais obtém-se:

$$Q_n(a, \lambda) = \sum_{t=3}^n \left(X_t - a(X_{t-1} - X_{t-2}) - (1 - 2a) \frac{(2 + \lambda)}{\lambda + \lambda^2} \right)^2.$$

A partir da equação acima, pode-se encontrar os estimadores por meio da solução do sistema:

$$\left(\frac{\partial Q_n}{\partial a}, \frac{\partial Q_n}{\partial \lambda} \right) = (0, 0).$$

Como a forma fechada para o parâmetro autorregressivo foi encontrada pelo método Yule-Walker, utilizou-se, para o *Caso 1*, o método dos MQC para estimar o parâmetro λ .

Na próxima Seção, são apresentados os resultados das simulações. As estimativas para os parâmetros autorregressivos pelo método MQC foi suprimida, uma vez que apresentaram resultados bem inferiores ao método Yule-Walker.

3.1.7 Resultados das Simulações

Considere que se deseja estimar o termo AR a e o parâmetro λ . Para tal adotam-se amostras de 1000 réplicas de Monte Carlo, quantidade considerada suficiente para garantir a convergência dos estimadores e a estabilidade dos resultados. O método Yule-Walker é usado em duas situações para os parâmetros autorregressivos: $h = 1$ e $h = 2$ e o método MQC para estimar o parâmetro λ .

Foram considerados 4 diferentes cenários para os valores dos parâmetros. Em cada cenário, fixou-se o valor de $\lambda = 12$, dentro da região de estabilidade, variando os valores de a . Os resultados dispostos na Tabela 1 representam os valores das estimativas dos parâmetros a e λ e seus respectivos EQM.

Para os dois primeiros cenários, o método Yule-Walker em $h = 2$ apresentou estimativas mais próximas para os verdadeiros valores dos parâmetros. Para $h = 1$ os dois últimos cenários favoreceram as estimativas neste caso. No entanto, pode-se afirmar que os dois casos são confiáveis para estimar os parâmetros autorregressivos, visto que, à medida que aumenta-se o número de observações, as estimativas tendem a convergir para os valores verdadeiros. Além disso, os valores do erro quadrático médio diminuem, nos dois casos, ao aumentar o número de observações.

Nas estimativas para o parâmetro λ , nota-se que seus resultados são diretamente influenciados pelo valor do termo autorregressivo a . À medida que aumenta-se o valor do parâmetro autorregressivo, as estimativas para o λ tendem a convergir mais rapidamente para o valor verdadeiro, além de ocorrer a diminuição do EQM.

Tabela 1 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados

Cenários	Yule-Walker ($h = 1$)			Yule-Walker ($h = 2$)		Mínimos Quadrados	
	N	$\hat{a}$	EQM($\hat{a}$)	$\hat{a}$	EQM($\hat{a}$)	$\hat{\lambda}$	EQM($\hat{\lambda}$)
(0.1, 12)	50	0.0033	0.0534	0.0427	0.0368	11.5552	0.1978
	100	0.0770	0.0091	0.0853	0.0080	11.5553	0.1977
	150	0.0860	0.0057	0.0888	0.0054	11.5554	0.1975
(0.2, 12)	50	0.1082	0.0411	0.1194	0.0311	11.6236	0.1416
	100	0.1769	0.0066	0.1786	0.0063	11.6243	0.1411
	150	0.1845	0.0042	0.1850	0.0039	11.6246	0.1409
(0.30, 12)	50	0.2378	0.0240	0.2097	0.0236	11.7115	0.0832
	100	0.2846	0.0039	0.2747	0.0043	11.7130	0.0823
	150	0.2883	0.0025	0.2818	0.0028	11.7134	0.0820
(0.4, 12)	50	0.3978	0.0035	0.3296	0.0095	11.8285	0.0294
	100	0.4096	0.0010	0.3793	0.0017	11.8312	0.0284
	150	0.4080	0.0007	0.3842	0.0011	11.8320	0.0282

Fonte: Elaboração Própria (2025)

3.1.8 Aplicação

As moedas fiduciárias tradicionais, como o dólar, o euro ou o iene japonês, são acompanhadas por stablecoins (moedas estáveis) mantidas numa conta bancária específica. Verificou-se que o modelo de série temporal dos preços do Bitcoin (BTC) não é estacionário. Uma abordagem alternativa para cumprir a estacionariedade é usar as volatilidades dos preços.

Assim, após uma crise financeira ou uma nova regulamentação governamental, pode haver mudanças significativas nos mercados que não são necessariamente capturadas pelos retornos. Nesses casos, as mudanças na variância indicam a chegada de informações e, portanto, qualquer medida de dispersão pode refletir a relação entre o fluxo de informações e a volatilidade nas séries temporais financeiras. Com isso, usar volatilidades em vez de retornos de ativos pode levar a um modelo mais previsível e, portanto, pode ser modelado muito bem, como em

Alizadeh, Brandt e Diebold (2002) e Chiang e Wang (2011) que usaram a volatilidade no mercado de ações e ajustaram um modelo de autorregressão condicional (CARR) a elas.

Um tipo de volatilidade baseada no intervalo durante um período fixo (por exemplo, um dia ou um mês) é definida como:

$$X_t = [\log H_t - \log L_t],$$

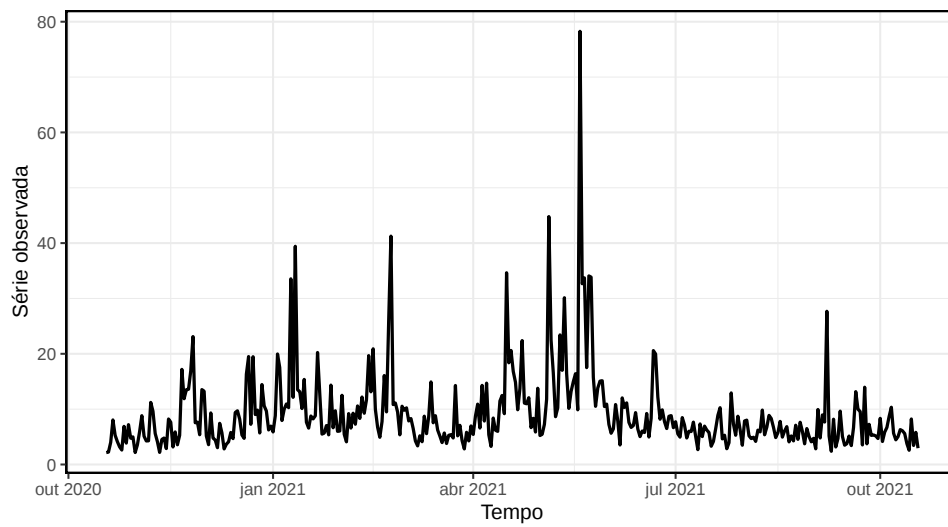
onde H_t e L_t representam o preço mais alto e mais baixo, respectivamente, do Bitcoin (BTC) no intervalo $[t - 1, t]$, ou seja, um dia ou um mês.

Dessa forma, para avaliar o desempenho do modelo $LAR_1(2)$, aplicou-se um conjunto de dados reais que representa a cotação da bolsa de valores de um mercado no período entre outubro de 2020 até outubro de 2021. A base de dados pode ser encontrada em <https://finance.yahoo.com/quote/BTC-USD/history>.

Essa base de dados é interessante, pois permite analisar a volatilidade do mercado como o número de ações compradas e vendidas em um dia de negociação na bolsa de valores. Uma grande variação entre preços de ações altos e baixos pode indicar um mercado instável. Para garantir que o conjunto de dados é apropriado, foram aplicados os testes Dickey-Fuller e Philip-Perron. Esses testes são utilizados a fim de verificar se uma série é estacionária. Dessa forma, considerou-se como hipótese nula, a série não ser estacionária. Com um nível de significância de 0.05, rejeita-se a hipótese nula, em ambos os testes, com p-valor em torno de 0.01, mostrando assim, que a série é estacionária, logo adequada para a modelagem.

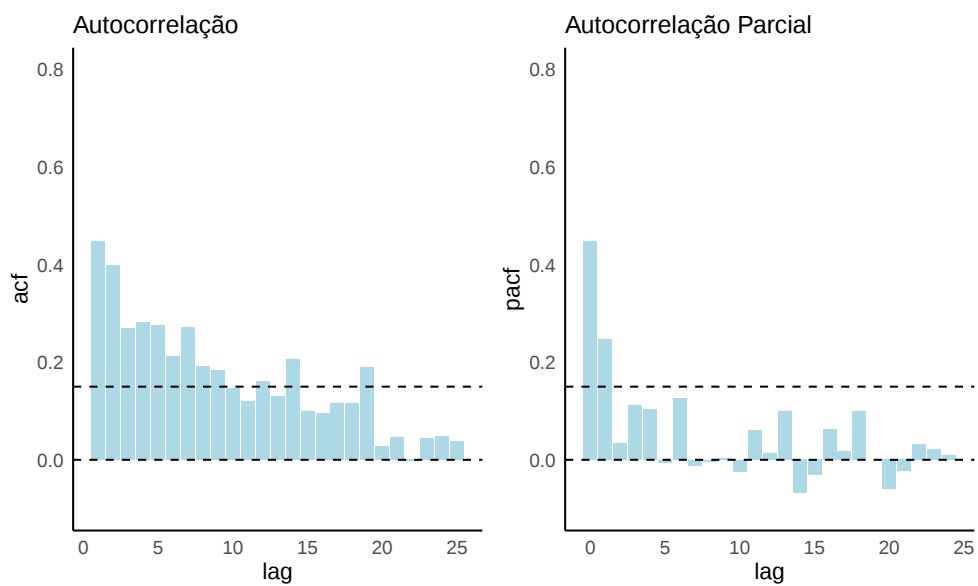
Além disso, o comportamento dos gráficos de autocorrelação como um decaimento exponencial e autocorrelação significativa nos dois primeiros *lags* evidenciam que se trata de um processo $AR(2)$. A Figura 8 representa a ST para o conjunto de dados. Em seguida tem-se os gráficos de autocorrelação e autocorrelação parcial. O par, ACF e PACF indica a presença de valores estacionários e de memória curta, enquanto a série observada mostra um aumento exponencial no meio da série observada.

Figura 8 – Fechamento da bolsa de valores



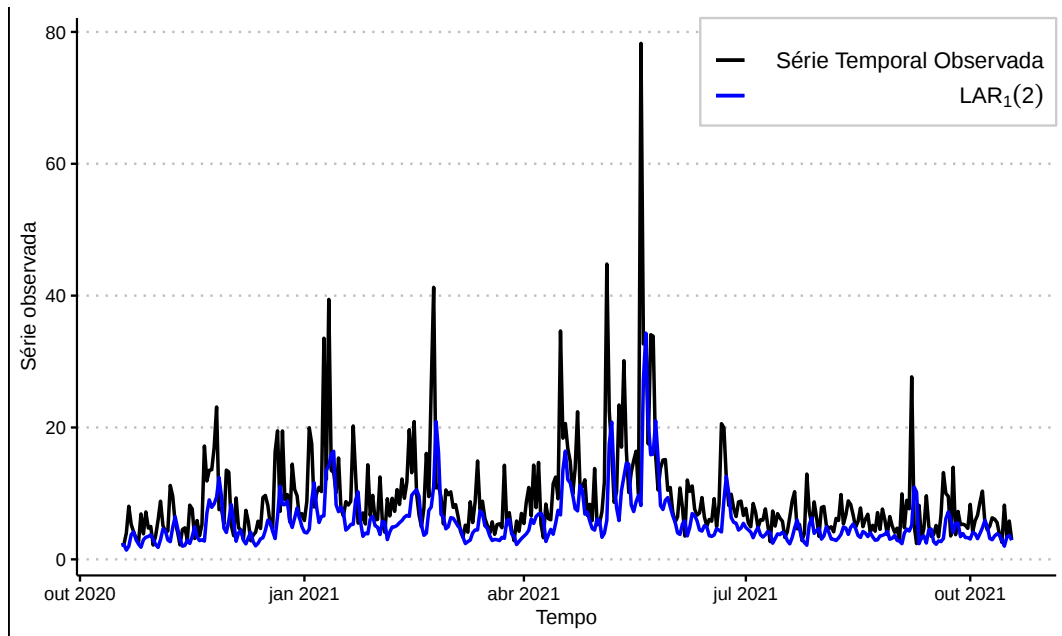
Fonte: Elaboração Própria (2025)

Figura 9 – Autocorrelação e autocorrelação parcial para a série



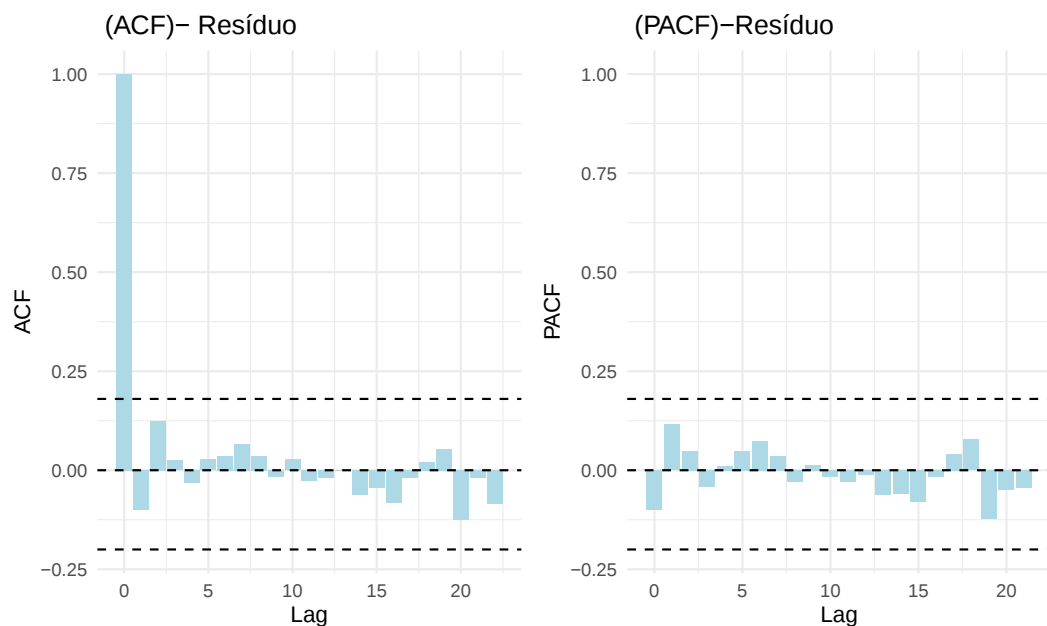
Fonte: Elaboração Própria (2025)

Dessa maneira, utilizando a base de dados e já conhecendo a forma fechada para o estimador autorregressivo, faz-se a predição do modelo $LAR_1(2)$, por meio do valor esperado condicional já desenvolvido. A Figura 10 mostra a série que representa o banco de dados e as séries do novo modelo obtidas por previsão.

Figura 10 – Gráfico para as séries observada e $LAR_1(2)$ 

Fonte: Elaboração Própria (2025)

Já o resíduo é obtido como a diferença entre as série observada e a série predita dividido pela variância do erro, para o processo $LAR_1(2)$. A Figura 11 mostra o gráfico da ACF e PACF do resíduo, indicando que não há correlação significativa entre eles para diferentes defasagens. Além disso, pelo Box Test, não há evidências para se rejeitar a hipótese nula, sugerindo então que o resíduo tem características de um ruído branco.

Figura 11 – Análise dos resíduos para o $LAR_1(2)$ 

Fonte: Elaboração Própria (2025)

3.1.9 Considerações Parciais

Com a motivação de existir uma lacuna na literatura sobre trabalhos que utilizam processos AR(2) com distribuição marginal, neste Capítulo está sendo desenvolvido um novo modelo AR(2) com distribuição marginal Lindley. Atribuindo uma restrição para os parâmetros AR de uma ST de segunda ordem, de modo que eles sejam iguais, denominou-se essa condição de *Caso 1* sendo o modelo chamado de $LAR_1(2)$. Desse modo, através de técnicas matemáticas como a transformada de Laplace e uso do método de frações parciais obteve-se uma distribuição para a inovação como sendo uma mistura de distribuições $\text{Gama}(2, \lambda)$, $\exp(\lambda)$ e $\exp(\frac{\lambda+1}{a})$. Propriedades Matemáticas como densidade espectral, variância e esperança condicional, ACF e PACF e equação de previsão foram desenvolvidas para o novo modelo proposto.

Para estimar os parâmetros do modelo, encontrou-se a forma fechada dos termos autorregressivos em função das autocorrelações amostrais, pelo método de estimação Yule-Walker. Por meio do método dos MQC, que é determinado pelo valor esperado condicional, foi feita a estimativa do parâmetro adicional (λ). A avaliação do desempenho desse estimador foi realizada utilizando simulação de Monte Carlo, mostrando que o método de estimação Yule-Walker produziu estimativas precisas e consistentes para estimar os parâmetros autorregressivos do modelo. Além disso, através do método dos MQC, obteve-se boas estimativas para o parâmetro adicional λ . Embora esse método também possa estimar os parâmetros autorregressivos, a preferência recaiu sobre o primeiro devido aos seus melhores resultados.

Por último, aplicou-se o novo modelo a um conjunto de dados com 366 observações sobre a cotação da bolsa de valores no período de Outubro de 2020 até Outubro de 2021. Para isso, utilizou-se os valores das cotações máximas e mínimas diárias. Através da equação de previsão, o $LAR_1(2)$ teve ajuste satisfatório ao banco de dados utilizado, garantindo assim a estacionariedade. Além disso, o resíduo não possui correlações significativas, indicando se tratar de ruído branco. Dessa maneira, a contribuição para o modelo aplicado dentro do mercado de ações, facilita identificar padrões, fazendo previsão de preços e tomada de decisões sobre investimento de compra e venda, analisando as tendências.

Na próxima Seção, esse mesmo modelo AR(2) com distribuição marginal Lindley é estendido para o caso em que os parâmetros autorregressivos são diferentes.

3.2 CASO 2: PROCESSO AR(2) COM DISTRIBUIÇÃO MARGINAL LINDLEY

Na Seção anterior, utilizando uma restrição nos parâmetros autorregressivos, foi desenvolvido um novo modelo autorregressivo de segunda ordem não condicional utilizando a distribuição Lindley como marginal. Após a construção de toda a estrutura, além de desenvolver várias propriedades estruturais, amplia-se, aqui, a análise de modo que os parâmetros autorregressivos sejam diferentes. Antes considerou-se:

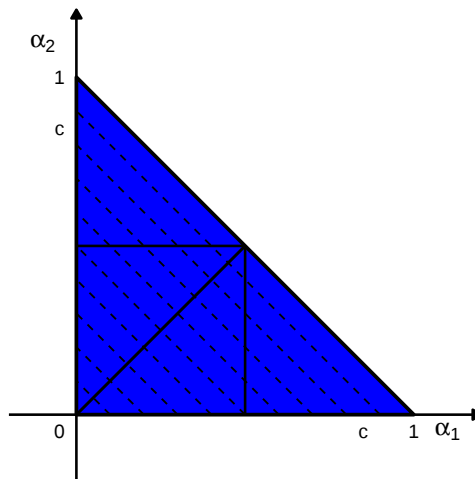
$$\alpha_1 = \alpha_2 = a, \quad 0 < a < 1.$$

E agora, será a seguinte relação:

$$\alpha_2 = k = c - \alpha_1, \quad 0 < \alpha_1 < c < 1. \quad (3.18)$$

A Figura 12 representa a região de estabilidade, considerando a relação descrita na Equação (3.18), que caracteriza a condição que os parâmetros autorregressivos sejam diferentes. Nesse trabalho considerou-se $c = 0.9$. A restrição para $\alpha_1 + \alpha_2 = 0.9$ define uma linha reta no plano cartesiano, de modo que, para o sistema ser estável, essa linha precisa estar dentro do triângulo de estabilidade. Ao fazer a limitação de modo que $\alpha_1 + \alpha_2 < 1$, evita-se a não estacionariedade da série, além de obter um decaimento controlado da autocorrelação. Assim, essa escolha é compatível com a teoria de estabilidade, pois para qualquer par (α_1, α_2) , seguindo essa restrição, ao longo da linha, o sistema será estável, pois todas as raízes da equação característica estarão fora do círculo unitário.

Figura 12 – Triângulo de estabilidade para um processo AR(2), Caso 2



Fonte: Elaboração Própria (2025)

Utilizando os mesmos artifícios para a construção do modelo $LAR_1(2)$, aplicou-se a transformada inversa de Laplace assim como foi desenvolvido para encontrar a Equação (3.5), obtendo:

$$\phi_\epsilon(s) = \frac{(\lambda + s + 1)(\lambda + \alpha_1 s)^2(\lambda + ks)^2}{(\lambda + s)^2[(1 - k)(\lambda + \alpha_1 s + 1)(\lambda + ks)^2 + k(\lambda + ks + 1)(\lambda + \alpha_1 s)^2]}, \quad (3.19)$$

com:

$$\alpha_2 = k = c - \alpha_1, \quad 0 < c < \alpha_1 < 1.$$

Nota-se que pela Equação (3.19), para esse caso, tem-se também uma evidência de que a distribuição de inovação ϵ_t é uma mistura envolvendo a distribuição Gama. Tal fato será expresso na proposição a seguir.

Proposição 3.2.1 *Seja X_t um processo autorregressivo de média nula e entrada ϵ_t com $\alpha_2 = c - \alpha_1$. Então, a distribuição de ϵ_t é uma mistura de distribuições Gama e exponenciais com parâmetros diferentes. Isto é, a fda de ϵ_t , $F_{\epsilon_t}(\cdot)$ é tal que:*

$$F_{\epsilon_t}(l) = T_1^* F_{\epsilon_t,1}(l) + T_2^* F_{\epsilon_t,2}(l) + T_{31}^* F_{\epsilon_t,3}(l) + T_{32}^* F_{\epsilon_t,4}(l) + T_{33}^* F_{\epsilon_t,5}(l),$$

sendo $F_{\epsilon_t}(\cdot)$ a fda de ϵ_t , $\epsilon_{t,1} \sim \exp(\lambda)$, $\epsilon_{t,2} \sim \Gamma(2, \lambda)$, $\epsilon_{t,3} \sim \exp(\phi_1)$, $\epsilon_{t,4} \sim \exp(\phi_2)$ e $\epsilon_{t,5} \sim \exp(\phi_3)$, em que $\phi_i = -r_i$ representa as raízes do polinômio de grau 3 que aparece ao aplicar a inversa de Laplace em:

$$(1 - k)(\lambda + \alpha_1 s + 1)(\lambda + ks)^2 + k(\lambda + ks + 1)(\lambda + \alpha_1 s)^2, \quad (3.20)$$

com r_1, r_2, r_3 raízes do polinômio, na variável s , dada por:

$$\begin{aligned} z(s) = & (\alpha_1^2 k^2 + \alpha_1 k^2) s^3 + (2\alpha_1 k^2 \lambda + 2\alpha_1 k \lambda + \alpha_1^2 k \lambda + k^2 \lambda + \alpha_1^2 k + k^2) s^2 \\ & + (2\alpha_1 k \lambda^2 + 2\alpha_1 k \lambda + 2k \lambda + 2k \lambda^2 + k^2 \lambda^2 + \alpha_1 \lambda^2) s + \lambda^3 + \lambda^2 + k \lambda^3 + k \lambda^2. \end{aligned} \quad (3.21)$$

A partir de então, para o Caso 2, que é o caso mais abrangente, o modelo será chamado de $LAR_2(2)$. Além disso, para qualquer par de parâmetros e considerando $\alpha_1 + \alpha_2 = 0.9$, o sistema $AR(2)$ será estável. Assim, pode-se obter uma infinidade de combinação dos parâmetros como, por exemplo:

$$\begin{cases} \alpha_1 = 0.8, & \alpha_2 = 0.1, \\ \alpha_1 = 0.35, & \alpha_2 = 0.55, \\ \alpha_1 = 0.75, & \alpha_2 = 0.15. \end{cases}$$

É importante destacar que, para esse estudo, iremos considerar que as raízes do polinômio de grau 3 sejam reais, evitando, assim, os casos complexos. Além disso, as 3 raízes são negativas, inclusive para os casos em que existem raízes complexas. Dessa maneira, ao resolver a Equação (3.20) por meio do software matemático Matlab, as raízes r_1 , r_2 e r_3 são sempre menores que 0 e possuem a forma geral com a seguinte estrutura:

$$r_1 = \left\{ \frac{\delta_2}{\delta_1} + \delta_1 - \frac{\delta_7}{3ak(a-k+1)} \right\},$$

em que:

$$\begin{aligned} \delta_1 &= \sqrt{((\delta_5 + \delta_4 - \delta_3)^2 - \delta_5 - \delta_4 + \delta_3)^{\frac{1}{3}}}, & \delta_2 &= \frac{\delta_7^2}{9a^2k^2(a-k+a)^2} - \frac{\lambda\delta_6}{3ak^2(a-k+1)}, \\ \delta_3 &= \frac{\lambda\delta_6\delta_7}{6a^2k^3(a-k+1)^2}, & \delta_4 &= \frac{\lambda^2(\lambda+1)}{2ak^2(a-k+1)}, & \delta_5 &= \delta_7^3 27a^3k^3(a-k+1)^3, \\ \delta_6 &= 2k + 2ak + a\lambda + 2k\lambda - k^2\lambda - 2k^2 + ak\lambda & \text{e} & \delta_7 = k + 2a\lambda + k\lambda + a^2\lambda - k^2\lambda + a^2 - k^2. \end{aligned}$$

$$r_2 = \left\{ -\frac{\zeta_2}{2\zeta_1} - \frac{\zeta_5}{3a^2k^2 - 3ak^3 + 3ak^2} - \frac{\zeta_1}{2} + \frac{\sqrt{3} \left(\frac{\zeta_2}{\zeta_1} - \zeta_1 \right) i}{2} \right\},$$

tais que:

$$\zeta_1 = \left(\sqrt{\left(\left(\zeta_4 + \frac{\zeta_5^3}{27(a^2k^2 - ak^3 + ak^2)^3} - \zeta_3 \right)^2 - \zeta_2^3 \right) - \frac{\zeta_5^3}{27(a^2k^2 - ak^3 + ak^2)^3} - \zeta_4 + \zeta_3} \right)^{\frac{1}{3}},$$

$$\begin{aligned} \zeta_2 &= \frac{\zeta_5^2}{9(a^2k^2 - ak^3 + ak^2)^2} - \frac{\zeta_6}{3a^2k^2 - 3ak^3 + 3ak^2}, & \zeta_3 &= \frac{\zeta_5\zeta_6}{6(a^2k^2 - ak^3 + ak^2)^2}, \\ \zeta_4 &= \frac{\lambda^3 + \lambda^2}{2a^2k^2 - 2ak^3 + 2ak^2}, & \zeta_5 &= a^2k + k^2\lambda - k^3\lambda + k^2 - k^3 + 2ak\lambda + a^2k\lambda \end{aligned}$$

e

$$\zeta_6 = 2k\lambda + a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2.$$

Por último,

$$r_3 = \left\{ -\frac{\eta_2}{2\eta_1} - \frac{\eta_5}{3a^2k^2 - 3ak^3 + 3ak^2} - \frac{\eta_1}{2} - \frac{\sqrt{3} \left(\frac{\eta_2}{\eta_1} - \eta_1 \right) i}{2} \right\},$$

onde:

$$\eta_1 = \left(\sqrt{\left(\left(\eta_4 + \frac{\eta_5^3}{27(a^2k^2 - ak^3 + ak^2)^3} - \eta_3 \right)^2 - \eta_2^3 \right) - \frac{\eta_5^3}{27(a^2k^2 - ak^3 + ak^2)^3} - \eta_4 + \eta_3} \right)^{\frac{1}{3}},$$

$$\begin{aligned}\eta_2 &= \frac{\eta_5^2}{9(a^2k^2 - ak^3 + ak^2)^2} - \frac{\eta_6}{3a^2k^2 - 3ak^3 + 3ak^2}, \quad \eta_3 = \frac{\eta_5\eta_6}{6(a^2k^2 - ak^3 + ak^2)^2}, \\ \eta_4 &= \frac{\lambda^3 + \lambda^2}{2a^2k^2 - 2ak^3 + 2ak^2}, \quad \eta_5 = a^2k + k^2\lambda - k^3\lambda + k^2 - k^3 + 2ak\lambda + a^2k\lambda\end{aligned}$$

e

$$\eta_6 = 2k\lambda + a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2.$$

Por uma convenção estabelecida, considere $\alpha_1 = a$ e $\alpha_2 = k$ nas expressões que representam as raízes do polinômio de terceiro grau. O mesmo vale na identificação dos pesos da Proposição 3.2.1 determinados por:

teste

$$T_1^* = \frac{\lambda^4(\lambda + 1) + \frac{\lambda^6(a-1)^2(k-1)^2(\lambda+1)}{\lambda^2(k-1)^3(\lambda-a\lambda+1)-k\lambda^2(a-1)^2(\lambda-k\lambda+1)} - \frac{\lambda^2(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\lambda+r)^2}}{\lambda^4(\lambda + 1)}$$

$$T_2^* = -\frac{\lambda^2(a-1)^2(k-1)^2}{\lambda^2(k-1)^3(\lambda-a\lambda+1)-k\lambda^2(a-1)^2(\lambda-k\lambda+1)},$$

$$T_3^* = \frac{(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\lambda+r)^2},$$

$$T_{31}^* = -\frac{(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\phi_1) \left[(\lambda+r)^2 \left(2k\lambda + a\lambda^2 + \phi_1^2(3a^2k^2 - 3ak^3 + 3ak^2) + 2k\lambda^2 - 2k^2\lambda \right) + \phi_1 \left(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 + 4ak\lambda + 2a^2k\lambda \right) - k^2\lambda^2 + 2ak\lambda + ak\lambda^2 \right]},$$

$$T_{32}^* = -\frac{(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\phi_2) \left[(\lambda+r)^2 \left(2k\lambda + a\lambda^2 + \phi_2^2(3a^2k^2 - 3ak^3 + 3ak^2) + 2k\lambda^2 - 2k^2\lambda \right) + \phi_2 \left(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 + 4ak\lambda + 2a^2k\lambda \right) - k^2\lambda^2 + 2ak\lambda + ak\lambda^2 \right]} \quad \text{e}$$

$$T_{33}^* = -\frac{(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\phi_3) \left[(\lambda+r)^2 \left(2k\lambda + a\lambda^2 + \phi_3^2(3a^2k^2 - 3ak^3 + 3ak^2) + 2k\lambda^2 - 2k^2\lambda \right) + \phi_3 \left(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 + 4ak\lambda + 2a^2k\lambda \right) - k^2\lambda^2 + 2ak\lambda + ak\lambda^2 \right]},$$

de modo que

$$T_{31}^* = -\frac{T_3^*}{k_1\phi_1}, \quad T_{32}^* = -\frac{T_3^*}{k_2\phi_2}, \quad T_{33}^* = -\frac{T_3^*}{k_3\phi_3},$$

com os coeficientes k_1 , k_2 e k_3 em função dos parâmetros autorregressivos e do termo λ , da forma:

$$\left\{ \begin{array}{l} k_1 = (3a^2k^2 - 3ak^3 + 3ak^2)\phi_1^2 + 2k\lambda + \phi_1(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 \\ \quad + 4ak\lambda + 2a^2k\lambda) + (a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2) \\ k_2 = (3a^2k^2 - 3ak^3 + 3ak^2)\phi_2^2 + 2k\lambda + \phi_2(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 \\ \quad + 4ak\lambda + 2a^2k\lambda) + (a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2) \\ k_3 = (3a^2k^2 - 3ak^3 + 3ak^2)\phi_3^2 + 2k\lambda + \phi_3(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 \\ \quad + 4ak\lambda + 2a^2k\lambda) + (a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2) . \end{array} \right.$$

A demonstração da Proposição 3.2.1 pode ser encontrada no Apêndice C.

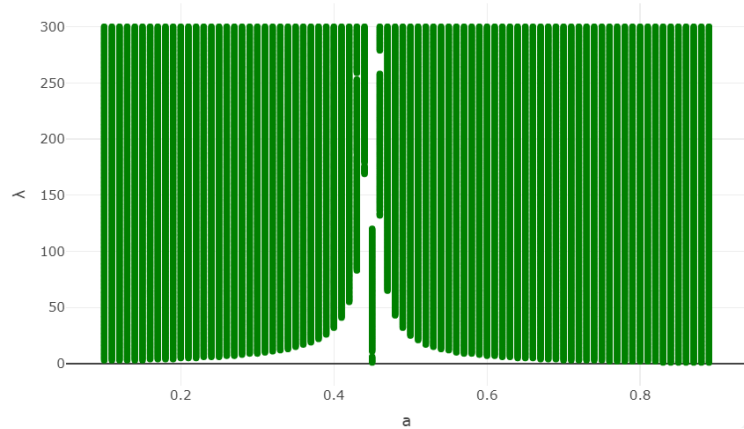
Um ponto de destaque refere-se ao caso em que os parâmetros autorregressivos são iguais. Quando se tem $\alpha_1 = \alpha_2 = 0.45$, o polinômio descrito na Equação (3.21) apresenta duas raízes reais e iguais. Com o auxílio do software Matalab, consegue-se obter a expressão geral que representa essas raízes, sendo elas em função de λ do tipo:

$$R_1 = -\frac{20}{9}(\lambda + 1) \quad \text{e} \quad R_2 = R_3 = -\frac{20\lambda}{9}.$$

A partir desse ponto, é necessário investigar a região de estabilidade para cada configuração de parâmetros. A Figura 13 apresenta a região do espaço (a, λ) em que todas as raízes do polinômio de grau 3 são reais e negativas. Por questões de tratabilidade, considerou-se o limite superior para o parâmetro λ de modo que:

$$0 < \lambda \leq 300.$$

Figura 13 – Região em que as raízes do polinômio são reais



Fonte: Elaboração Própria (2025)

A Figura 14 mostra o espaço paramétrico da região satisfeita referente aos parâmetros a , λ e k para cada uma das raízes r_1 , r_2 e r_3 . Como $k = 0.9 - a$, a análise reduz-se a duas dimensões, de modo que a variação do parâmetro a determina automaticamente a de k . Essa figura foi construída em R permitindo o usuário determinar qual a região dos parâmetros $(\alpha_1, \alpha_2, \lambda)$ tem-se raízes reais negativas e também satisfaz as condições da v.a ϵ_t ser uma mistura. Tal estrutura do código segue os passos:

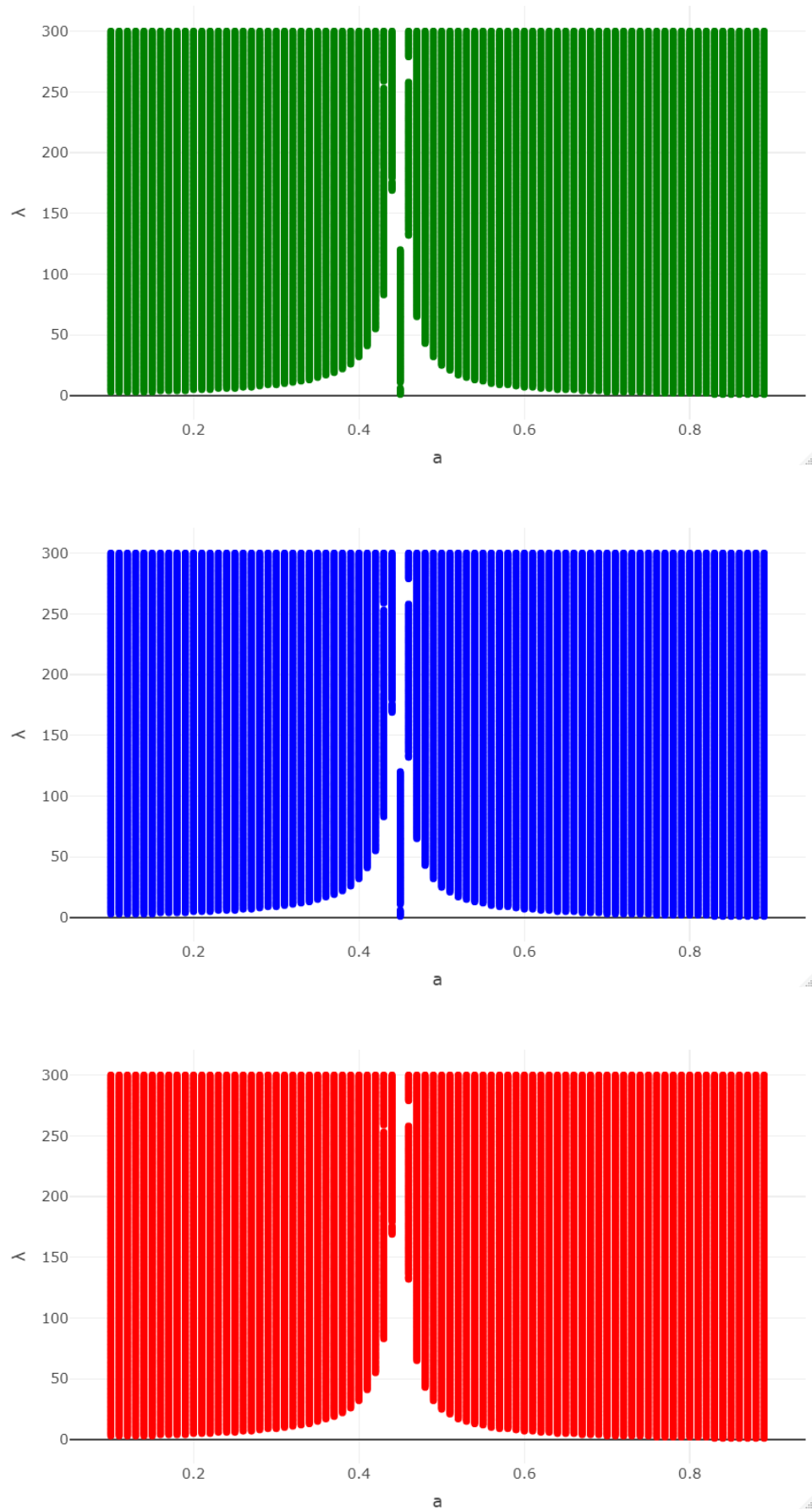
1. Carrega as bibliotecas `dplyr` e `plotly` para manipulação dos dados e visualização dos gráficos, respectivamente.
2. Define-se uma função para verificar se todas as raízes são reais.
3. O próximo passo consiste em definir a função soma, que corresponde à soma dos pesos.

Além disso, considera-se:

- $\alpha_1 = a$ variando de 0.1 até 0.89 com incrementos de 0.01,
 - $\alpha_2 = k = 0.9 - a$,
 - λ variando de 1 até 300 com incremento de 1.
4. Em seguida, o código retorna a soma dos pesos, que deve ser próxima de 1 com uma tolerância de 0.00001. Para cada combinação (a, λ) :
 - Se as raízes forem reais e a soma for aproximadamente 1, o par é armazenado.
 - Caso contrário, é descartado.

5. Utiliza-se dois comandos de repetição aninhados, onde um laço está contido dentro do outro. Isso permite testar todas as combinações possíveis de a e λ , com o loop interno para λ e o externo para a .
6. As funções `rowwise()` e `mutate()` garantem que o cálculo seja executado linha por linha e calcula a soma dos pesos, nessa ordem.
7. Utilizando a função `expand.grid()`, cria-se uma tabela combinando todos os valores de a e λ .
8. O código também imprime os resultados para um valor específico de a . Dessa forma, ao escolher um a específico, gera-se todos os valores de λ que satisfazem a condição da v.a ser uma mistura.
9. Por fim, tem-se o gráfico interativo, ao posicionar o cursor sobre qualquer ponto da região válida, são exibidos os valores de a , λ , soma, além das raízes do polinômio. Isso permite que o usuário visualize os intervalos de λ verticalmente para cada a .

Figura 14 – Espaço paramétrico para os parâmetros a e λ , considerando as raízes r_1, r_2 e r_3 do polinômio.



Fonte: Elaboração Própria (2025)

Na Figura 14, cada região representa a combinação dos parâmetros a , k e λ que satisfazem a condição da v.a ϵ_t ser uma mistura. No entanto, como a Equação (3.20) possui 3 raízes, para a construção desses modelos, considerou-se as regiões para cada conjunto de raízes (r_1, r_2, r_3) , respectivamente. Além disso, caso as raízes sejam complexas, a parte imaginária é desprezível da ordem 10^{-5} . Assim, considerando a região representada pela raiz r_1 , para $\alpha_1 = 0.1$ e $\alpha_2 = 0.8$ a partir de $\lambda = 170$ todas as raízes do polinômio de grau 3 são reais com soma dos pesos igual a 1, dentro da margem mínima de erro. No entanto, ao escolher uma outra raiz do polinômio, como a região destacada de vermelho que representa a raiz r_3 , para $\alpha_1 = 0.1$ e $\alpha_2 = 0.8$ a partir de $\lambda = 3$ tem-se as condições satisfeitas pra a v.a ser uma mistura. Na próxima Seção são exploradas as propriedades para o novo modelo considerando o *Caso 2*.

3.2.1 Propriedades Estruturais

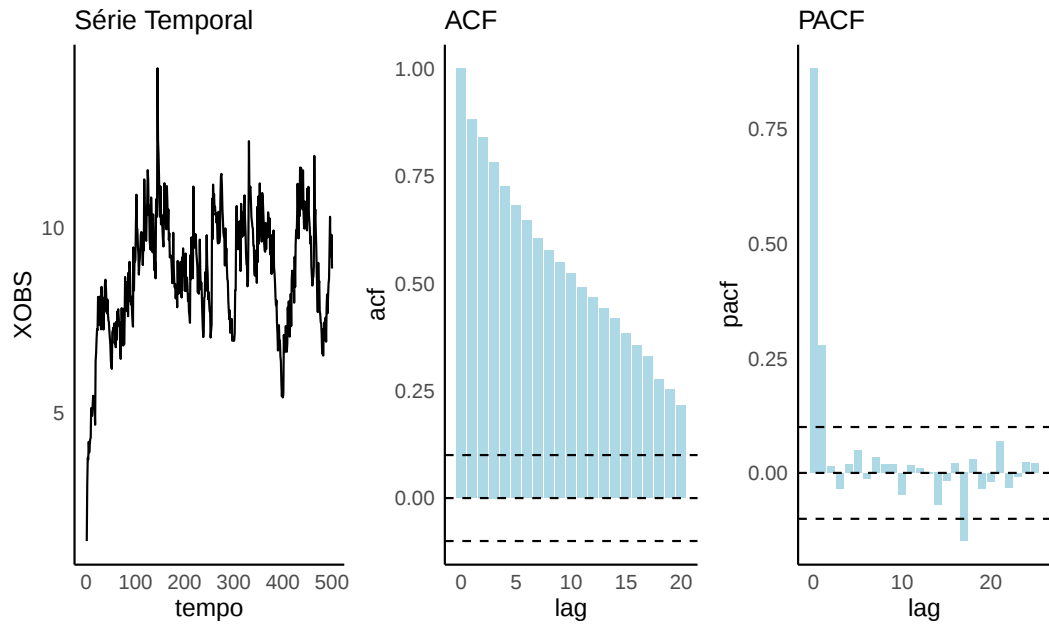
Para a construção dos novos modelos AR(2), foram desenvolvidas propriedades como esperança e variância condicional, densidade do erro e predição dos modelos. Expressões originalmente desenvolvidas para o *Caso 1*, as quais são estendidas para o caso geral, *Caso 2*.

Desse modo, por meio do gerador para o erro, construiu-se as Figuras 15 e 16 que representam os gráficos da série, autocorrelação e autocorrelação parcial para duas configurações de parâmetros distintos. Na Figura 17, tem-se a densidade do erro. Tais figuras correspondem ao *Caso 2* do modelo $LAR_2(2)$.

Além disso, é possível verificar que o comportamento dos gráficos da ACF e PACF indicam que se trata de um processo AR(2). Outro aspecto importante refere-se ao desempenho observado no gráfico da PACF. Isso pode ser explicado pela influência dos parâmetros autorregressivos (α_1, α_2) na estrutura do processo AR(2). Ao observar a Figura 15, com $\alpha_1 = 0.6$ e $\alpha_2 = 0.3$, o processo é dominado pelo termo de *lag* 1, X_{t-1} , resultando um PACF mais evidente que o segundo *lag*.

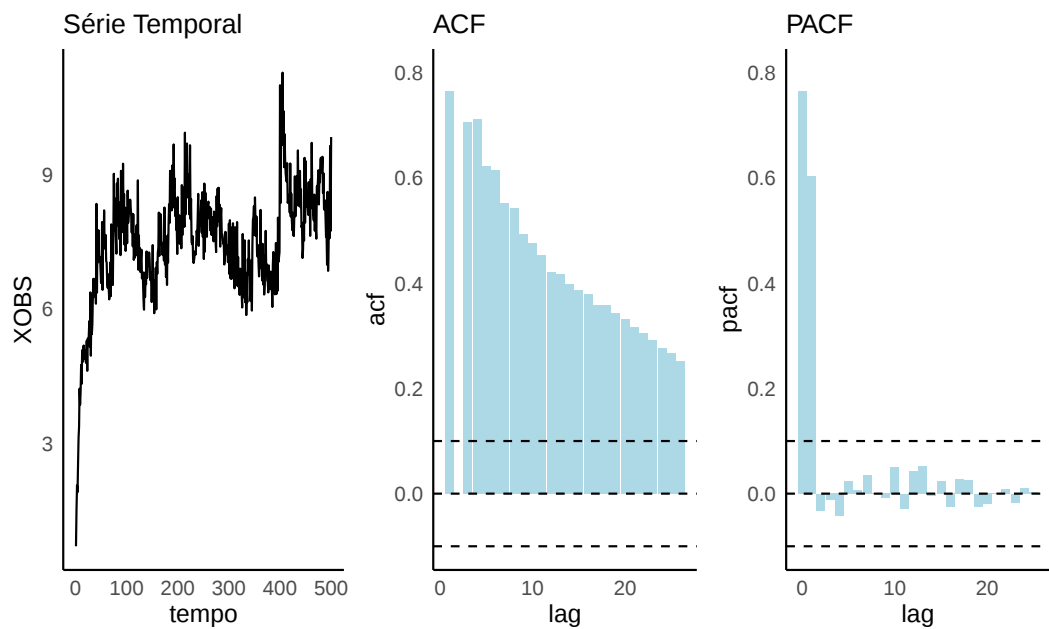
À medida que aumenta-se o parâmetro α_2 e, consequentemente diminui o α_1 , a influência do termo de *lag* 2 se torna mais significativa. Desse modo, o segundo *lag* na PACF passa a ser mais visível e a estrutura AR(2) fica mais evidente. Quando o parâmetro α_2 se torna dominante, conforme a Figura 16, a PACF reflete claramente a dependência do processo em relação aos dois *lags* anteriores. Portanto, a mudança na visibilidade dos *lags* na PACF é diretamente relacionada à variação dos parâmetros autorregressivos.

Figura 15 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.6, \alpha_2 = 0.3, \lambda = 100$



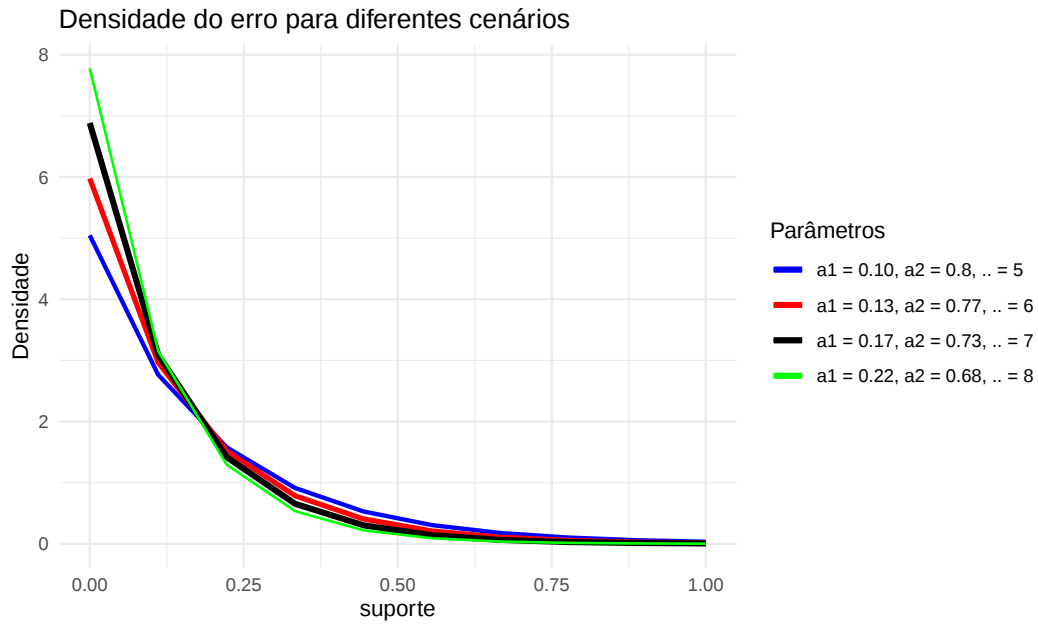
Fonte: Elaboração Própria (2025)

Figura 16 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.22, \alpha_2 = 0.68, \lambda = 100$



Fonte: Elaboração Própria (2025)

Figura 17 – Gráficos da densidade do erro para amostra de tamanho 500, e diferentes valores para os parâmetros.



Fonte: Elaboração Própria (2025)

A Figura 17 possui 4 configurações de parâmetros distintas. Nota-se que, em configurações com valores elevados de α_2 , e consequentemente reduzidos em α_1 , resultam valores mais concentrados, tendo pouca variabilidade. No entanto, à medida que aumentam os valores de α_1 e λ , diminuindo α_2 tem-se um crescimento da variabilidade do erro.

Assim, ao analisar as propriedades estruturais inerentes a um $AR(2)$ é essencial garantir solidez metodológica e confiabilidade nos resultados inferenciais. Adicionalmente, tais propriedades garantem que o modelo não apenas reproduz adequadamente a dinâmica linear subjacente à ST, mas também cumpra requisitos teóricos fundamentais como estacionariedade e especificação correta da estrutura de dependência, os quais são indispensáveis para a realização de inferências estatisticamente consistentes, previsões com menor erro de estimação e generalização dos resultados para diferentes horizontes temporais (BOX et al., 2015). A seguir, é apresentado dois métodos para estimar os parâmetros do modelo $LAR_2(2)$.

3.2.2 Método Yule-Waker

É possível obter uma forma fechada para estimar os parâmetros AR em função das autocorrelações amostrais, de modo análogo ao desenvolvido para o Caso 1. Considere um $AR(2)$:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t.$$

Da mesma maneira, multiplicando ambos os lados da equação por X_{t-h} , para $h \geq 0$ e tomando o valor esperado, tem-se que

$$\begin{aligned}\mathbb{E}(X_t X_{t-h}) &= \alpha_1 [\mathbb{E}(X_{t-1} X_{t-h})] + \alpha_2 [\mathbb{E}(X_{t-2} X_{t-h})] + \mathbb{E}(\epsilon_t X_{t-h}) \\ \iff \gamma_X(h) &= \alpha_1 [\gamma_X(h-1)] + \alpha_2 [\gamma_X(h-2)] + \mathbb{E}(\epsilon_t \sum_{i=0}^{\infty} \psi_i \epsilon_{t-h-i}) \\ \iff \gamma_X(h) &= \alpha_1 [\gamma_X(h-1)] + \alpha_2 [\gamma_X(h-2)] \\ \iff \rho_X(h) &= \alpha_1 [\rho_X(h-1)] + \alpha_2 [\rho_X(h-2)].\end{aligned}$$

Assim, as formas fechadas das estimativas dos parâmetros autorregressivos em função das estimativas das autocorrelações amostrais são da forma:

$$\hat{\alpha}_1 = \frac{\hat{\rho}(3) - \hat{\rho}(1)\hat{\rho}(2)}{\hat{\rho}(2) - \hat{\rho}^2(1)} \quad \text{e} \quad \hat{\alpha}_2 = \hat{\rho}(2) - \hat{\alpha}_1 \hat{\rho}(1).$$

Em geral, um resultado assintótico importante, derivado do *Caso 1* é :

$$\sqrt{N} \begin{pmatrix} \hat{\alpha}_1 - \alpha_1 & \hat{\alpha}_2 - \alpha_2 \end{pmatrix} \xrightarrow[N \rightarrow \infty]{D} N \left(\mathbf{0}, \sigma^2 \varepsilon \cdot \Gamma_2^{-1} \right) \quad \text{e} \quad \hat{\sigma}^2 \varepsilon \xrightarrow[N \rightarrow \infty]{P} \sigma_\varepsilon^2.$$

Da mesma forma, como desenvolvido para o *Caso 1*, pode-se encontrar um estimador para o parâmetro λ , sendo $\sigma^2 \epsilon$ a variância da mistura, dada por:

$$\sigma^2 \epsilon = \frac{T_1^{*2}}{\hat{\lambda}^2} + \frac{2T_2^{*2}}{\hat{\lambda}^2} + \frac{T_{31}^{*2}}{\hat{\phi}_1^2} + \frac{T_{32}^{*2}}{\hat{\phi}_2^2} + \frac{T_{33}^{*2}}{\hat{\phi}_3^2}.$$

3.2.3 Mínimos Quadrados Condicionais

Sabe-se que a equação de predição é determinada pelo valor esperado condicional do modelo. Conforme desenvolvido para o *Caso 1*, que é um caso particular, aqui tem-se:

$$g_t(\alpha_1, \alpha_2, \lambda) = \mathbb{E}(X_t \mid X_{t-1}, X_{t-2}),$$

que equivale a:

$$\begin{aligned}Q_n(\alpha_1, \alpha_2, \lambda) &= \sum_{t=3}^N [X_t - g_t(\alpha_1, \alpha_2, \lambda)]^2, \\ Q_n(\alpha_1, \alpha_2, \lambda) &= \sum_{t=3}^N \left\{ X_t - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - (1 - \alpha_1 - \alpha_2) \frac{2 + \lambda}{\lambda + \lambda^2} \right\}^2.\end{aligned}$$

Assim, na próxima Seção, são apresentados os resultados das simulações utilizando esses dois métodos.

3.2.4 Resultados das Simulações

A Tabela 2 mostra, para o *Caso 2*, as simulações dos parâmetros com 1000 réplicas de Monte Carlo, pelos métodos Yule-Walker e MQC, em que fixou-se o parâmetro λ variando os parâmetros autorregressivos. Diferente do *Caso 1*, foram incluídas as estimativas também dos parâmetros autorregressivos pelo método dos mínimos quadrados. Nota-se que, nos 5 cenários, as estimativas dos parâmetros autorregressivos pelo método Yule-Walker apresentaram resultados mais precisos que os obtidos via MQC. Além disso, verifica-se a influência do termo autorregressivo (α_2) nos resultados das estimativas do parâmetro λ . Conforme aumenta-se o termo α_2 as estimativas para o λ tendem a convergir melhor para o valor verdadeiro, além de apresentarem uma diminuição significativa para o EQM a cada observação. Os cenários foram feitos obedecendo a ordem dos parâmetros: α_1, α_2 e λ .

Tabela 2 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados Condicionais

Cenários	N	Yule-Walker				Mínimos Quadrados Condicionais					
		$\hat{\alpha}_1$	EQM($\hat{\alpha}_1$)	$\hat{\alpha}_2$	EQM($\hat{\alpha}_2$)	$\hat{\alpha}_1$	EQM($\hat{\alpha}_1$)	$\hat{\alpha}_2$	EQM($\hat{\alpha}_2$)	$\hat{\lambda}$	EQM($\hat{\lambda}$)
(0.35, 0.55, 100)	50	0.3568	0.0314	0.4977	0.0285	0.5109	0.0292	0.5000	0.0058	100.4915	0.2416
	100	0.3605	0.0280	0.5195	0.0239	0.5006	0.0227	0.5028	0.0022	100.4794	0.2299
	150	0.3537	0.0222	0.5302	0.0181	0.5002	0.0225	0.5012	0.0023	100.4675	0.2189
(0.30, 0.60, 100)	50	0.3312	0.0318	0.5231	0.0319	0.5038	0.0431	0.5072	0.0102	100.4883	0.2384
	100	0.3233	0.0232	0.5595	0.0206	0.5009	0.0403	0.5025	0.0094	100.4719	0.2228
	150	0.3065	0.0159	0.5771	0.0134	0.5003	0.0401	0.5011	0.0097	100.4542	0.2074
(0.28, 0.62, 100)	50	0.3261	0.0336	0.5284	0.0348	0.5018	0.0502	0.5093	0.0133	100.4872	0.2374
	100	0.3019	0.0217	0.5795	0.0196	0.5009	0.0488	0.5024	0.0138	100.4661	0.2172
	150	0.2995	0.0162	0.5876	0.0141	0.5003	0.0485	0.5011	0.0141	100.4462	0.1992
(0.24, 0.66, 100)	50	0.3059	0.0314	0.5481	0.0355	0.5025	0.0693	0.5085	0.0233	100.4857	0.2359
	100	0.2774	0.0198	0.6034	0.0187	0.5010	0.0681	0.5024	0.0248	100.4661	0.2172
	150	0.2575	0.0104	0.6293	0.0093	0.5003	0.0677	0.5010	0.0252	100.4432	0.1991
(0.18, 0.72, 100)	50	0.2751	0.0310	0.5817	0.0380	0.5029	0.1044	0.5080	0.0347	100.4849	0.2351
	100	0.2278	0.0144	0.6565	0.0146	0.5009	0.1030	0.5022	0.0474	100.4643	0.2156
	150	0.2066	0.0071	0.6815	0.0072	0.5003	0.1026	0.5009	0.0479	100.4430	0.1966

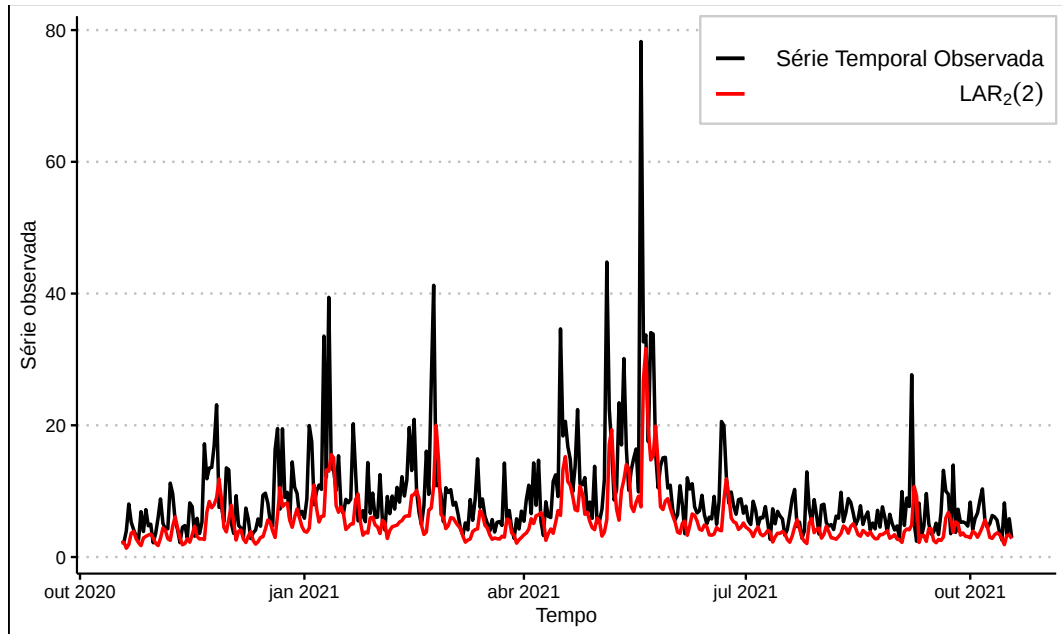
Fonte: Elaboração Própria (2025)

3.2.5 Aplicação

Com o objetivo de avaliar o desempenho do modelo $LAR_2(2)$, utilizou-se o mesmo conjunto de dados referente às cotações da bolsa de valores previamente aplicado ao modelo $LAR_1(2)$. A Figura 18 apresenta a série observada juntamente com as previsões do $LAR_2(2)$, considerando

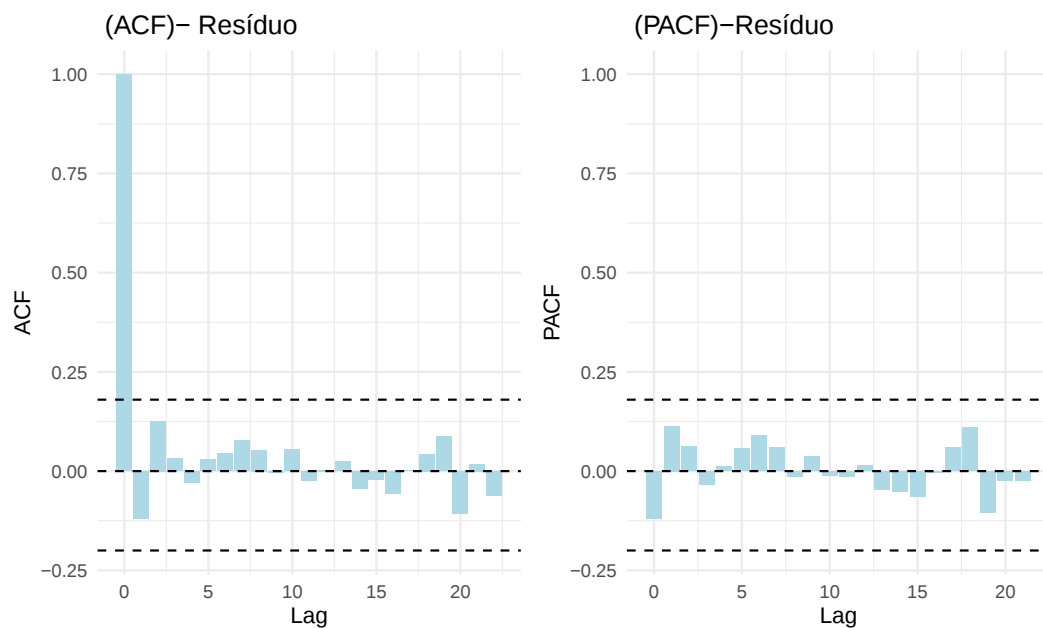
$h = 1$, horizonte de previsão. Observa-se, mais uma vez, que o modelo proporciona um bom ajuste aos dados reais. A Figura 19, por sua vez, exibe os gráficos da ACF e da PACF dos resíduos, nesta ordem.

Figura 18 – Gráfico para as séries observadas e $LAR_2(2)$



Fonte: Elaboração Própria (2025)

Figura 19 – Autocorrelação e autocorrelação parcial do resíduo

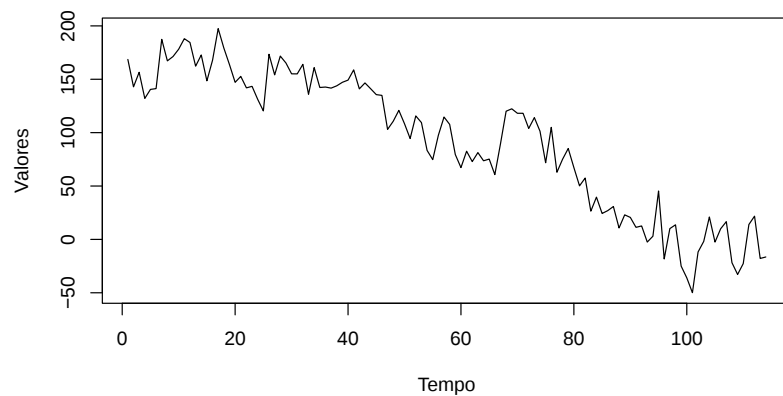


Fonte: Elaboração Própria (2025)

A seguir, é apresentada uma segunda aplicação para o novo modelo proposto. A base de

dados utilizada refere-se a dados climáticos diários da capital da Índia, Nova Delhi, podendo ser encontrada em <https://www.kaggle.com/datasets/sumanthvrao/daily-climate-time-series-data>. A coleta foi realizada no ano de 2017 englobando as variáveis temperatura e humidade. Essa base de dados é importante pois auxilia na previsão de eventos climáticos como furacões, inundações e secas. Para garantir que a base de dados é ideal, utilizou-se o teste Dickey-Fuller. Com um nível de significância de 5%, rejeita-se a hipótese nula, com p-valor próximo a 0.04. A seguir tem-se a ST formada com esse conjunto de dados contendo 114 valores. A Figura 21 representa a série observada juntamente com o modelo $LAR_2(2)$.

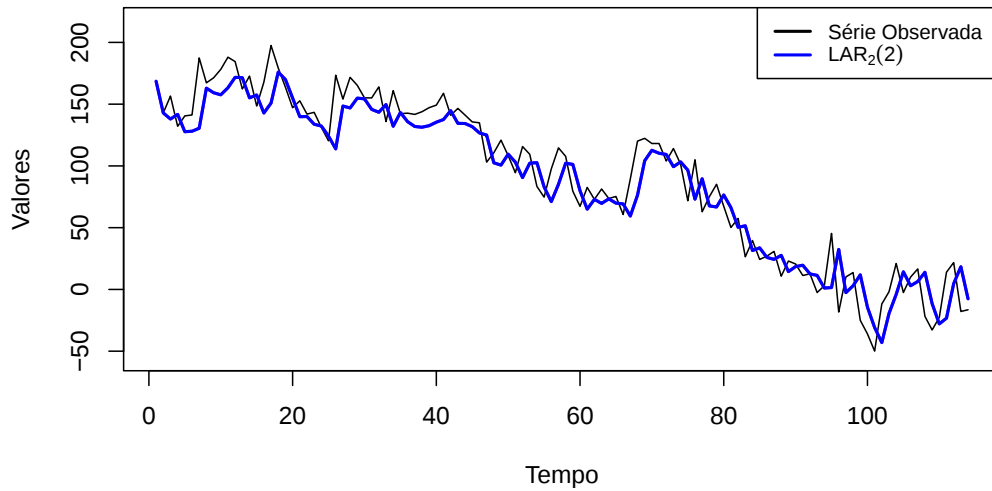
Figura 20 – Gráfico para a série



Fonte: Elaboração Própria (2025)

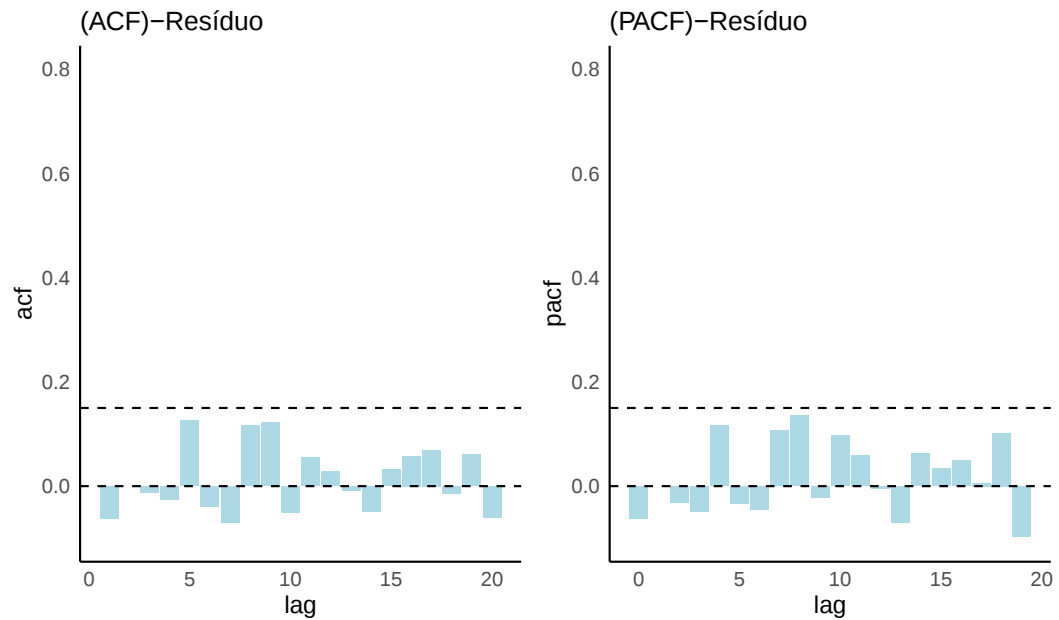
Novamente, observa-se que o modelo proposto se ajustou bem a esse conjunto de dados. Além disso, ao analisar a Figura 22, nota-se que os resíduos não possuem dependência. Logo, não apresentam autocorrelações significativas, indicando que tem-se fortes indícios de ser um ruído branco. Por fim, ao aplicar o Box-Test (Ljung-Box) obteve-se um p -valor maior que 5%. Pela estrutura do teste, tem-se evidências significativas para não rejeitar a hipótese nula, de que os resíduos são independentes e não há correlação entre eles.

Figura 21 – Gráfico para as séries observadas e $LAR_2(2)$



Fonte: Elaboração Própria (2025)

Figura 22 – Autocorrelação e autocorrelação parcial do resíduo



Fonte: Elaboração Própria (2025)

3.2.6 Considerações Parciais

Devido às dificuldades na literatura de se propor modelos $AR(2)$ não condicionais com distribuição marginal, foi proposto um modelo $AR(2)$ utilizando a Lindley como distribuição

marginal.

Dessa forma, esse Capítulo foi desenvolvido em duas partes denominadas de *Caso 1* e *Caso 2*, analisando a região de estabilidade para um modelo $AR(2)$. No *Caso 1*, considerou-se os parâmetros autorregressivos α_1 e α_2 tais que $\alpha_1 = \alpha_2 = a$. A partir daí, foi desenvolvida toda uma estrutura para as propriedades, considerando essa restrição para o modelo. Já para o *Caso 2*, admitiu-se que $\alpha_2 = c - \alpha_1$, refazendo todo o processo a fim de encontrar a distribuição de inovação, de tal modo como foi desenvolvido no *Caso 1* para o modelo $LAR_2(2)$.

Assim, a distribuição da inovação para o *Caso 2* do novo modelo $LAR_2(2)$ é também uma mistura de distribuições Gama de parâmetros $(2, \lambda)$, exponencial com parâmetro λ e 3 exponenciais com parâmetros ϕ_1 , ϕ_2 e ϕ_3 que correspondem o oposto das raízes negativas de um polinômio de grau 3 descrito na Equação (3.21). Esse polinômio é função dos parâmetros α_1 , α_2 e λ , diferenciando dos resultados para o *Caso 1*. Em seguida, obteve-se uma expressão geral para cada uma dessas raízes por meio do software Matemático Matlab, o qual comprova que todas são negativas. Além disso, para o caso em que $\alpha_1 = \alpha_2 = 0.45$ também encontrou-se uma expressão geral para as raízes do polinômio. Nela, nota-se que, das 3 raízes reais negativas, existem 2 raízes reais e iguais.

A partir desse ponto, foram exploradas as propriedades como esperança e variância condicionais, além de construir um gerador para o erro aleatório. Propriedades estas generalizadas provenientes das desenvolvidas para o *Caso 1*, que é o caso particular.

Nessa perspectiva, para estimar os parâmetros encontrou-se a forma fechada dos termos autorregressivos em função das autocorrelações amostrais, pelo método de estimação Yule-Walker. Por meio do método dos MQC, que se baseia no valor esperado condicional, foi feita a estimativa do parâmetro adicional (λ) e dos termos autorregressivos.

Ao comparar os dois métodos de estimação, o método Yule-Walker obteve estimativas mais próximas dos valores verdadeiros dos termos autorregressivos em relação aos MQC. A capacidade preditiva do novo modelo também foi avaliada aplicando-o a dois conjuntos de dados reais: um relacionado à cotação da bolsa de valores e o outro conjunto sobre dados meteorológicos coletados na cidade de Delhi. Em ambos os casos, o modelo $LAR_2(2)$ apresentou um bom ajuste. Além disso, ao analisar os resíduos do modelo, nos dois conjuntos de dados reais, não foi detectada a presença de autocorrelações significativas. Isso indica que eles possuem características de serem um ruído branco.

A partir da construção desse modelo, será estendida a proposta de estudo utilizando a Gama-Lindley como a distribuição marginal, visto no próximo capítulo.

4 A CONSTRUÇÃO DE UM NOVO MODELO GAMA-LINDLEY DE SEGUNDA ORDEM

Nesse Capítulo, é apresentado um novo modelo AR(2) com distribuição marginal Gama-Lindley. Antes disso, é feita uma discussão sobre a Gama-Lindley e suas propriedades. Na Seção 4.1, é realizada uma revisão sobre a distribuição Gama-Lindley e algumas de suas extensões. Na Seção 4.2, é apresentado um novo modelo de ST de ordem 2 não condicional com distribuição marginal Gama-Lindley. Na Subseção 4.2.1, são apresentadas as propriedades do novo modelo. A aplicação a um conjunto de dados reais é feita na Subseção 4.2.5. Na Subseção 4.2.6, são feitas as considerações parciais do Capítulo. Na Seção 4.3, o modelo é proposto considerando que os parâmetros autorregressivos são diferentes. Os resultados das simulações são apresentados na Subseção 4.3.2. Por fim, a aplicação a conjunto de dados e as considerações são efetuadas nas Subseções 4.3.3 e 4.3.4, nessa ordem.

4.1 DISTRIBUIÇÃO GAMA-LINDLEY

Algumas distribuições são derivadas de outras já conhecidas. Uma das suas vantagens é permitir representar uma maior variabilidade dos dados à medida que se trabalha com mais parâmetros. Dentro desse contexto, tem-se a distribuição Gama-Lindley. Formulada por Zeghdoudi e Nedjar (2015), a distribuição Gama-Lindley pode ser considerada como uma extensão da distribuição Lindley. Essa distribuição foi construída como mistura de uma Lindley (λ) de peso $\frac{1}{\beta}$ com uma Gama ($2, \lambda$) de peso $\left(\frac{\beta-1}{\beta}\right)$, com $0 < \beta < 1$. Note que a combinação dos pesos representa uma soma ponderada das distribuições individuais simples. Essa distribuição forneceu um bom ajuste ao ser aplicada a conjuntos de dados relacionado a tempos de falha.

Entretanto, existem na literatura outras distribuições Gama-Lindley, como a estudada por Lima (2015), que apresentou uma distribuição Gama-Lindley derivada da família de distribuições gama-G, para mais detalhes, consulte Zografos e Balakrishnan (2009). Essa família de distribuições gera novas distribuições que dependem das funções Gama e Gama incompleta. Este modelo, ao ser aplicado a dados de tempo de falhas, apresentou bom ajuste em comparação a duas outras distribuições: Lindley e exponencial geométrico complementar.

Dessa forma, por simplicidade, optou-se, por utilizar distribuição Gama-Lindley proposta por Zeghdoudi e Nedjar (2015) como a distribuição da variável de saída do novo processo, de tal forma como foi desenvolvido para o processo Gama-Lindley AR(1), veja Mello et al.

(2021). O benefício de se usar tal distribuição é o fato da Gama-Lindley ajustar bem as características da distribuição Gama e Lindley obtendo um bom ajuste a dados complexos. Além disso, apresenta ferramentas analíticas completas como fdp sendo útil para uma inferência precisa além de possuir facilidade para implementação computacional. Em acréscimo, a Gama-Lindley consegue ultrapassar as restrições das distribuições individuais, por possuir uma maior flexibilidade de formas e caudas pesadas.

A distribuição Gama-Lindley possui a seguinte fdp:

$$f(t; \lambda, \beta) = \frac{\lambda^2}{\beta(1 + \lambda)} [(\beta\lambda + \beta - \lambda)t + 1] \exp(-\lambda t), \quad t > 0, \quad \text{com } \lambda > 0, \quad \beta > 0,$$

sendo λ o parâmetro de escala e β o parâmetro de forma. Consequentemente, a fda é dada por

$$F(t; \lambda, \beta) = 1 - \left(\frac{[(\beta\lambda + \beta - \lambda)(\lambda t + 1) + \lambda]}{\beta(1 + \lambda)} \right) \exp(-\lambda t), \quad t > 0.$$

Para fins de notação, escrevemos que X tem distribuição Gama-Lindley com parâmetros λ e β como $X \sim \text{GAL}(\lambda, \beta)$.

O valor esperado e a variância de X são, respectivamente:

$$\begin{aligned} \mathbb{E}(X) &= \frac{2\beta(1 + \lambda) - \lambda}{\lambda\beta(1 + \lambda)} \\ \text{e} \quad \text{Var}(X) &= \frac{-(-2\beta\lambda + \lambda)^2 + 2\beta^2(1 + 3\lambda) - 2\beta(\beta\lambda - 3\beta\lambda^2 + 2\lambda^2)}{\lambda^2\beta^2(1 + \lambda)^2}. \end{aligned} \quad (4.1)$$

Utilizaremos o artifício envolvendo a transformada de Laplace afim de determinar as propriedades do novo modelo. Considerando a densidade de $X \sim \text{GAL}(\lambda, \beta)$, a sua transformada de Laplace é da forma:

$$\phi_X(s) = \frac{\lambda^2(\beta\lambda + \beta + s)}{\beta(1 + \lambda)(s + \lambda)^2}.$$

Com base nas definições acima, é proposto um novo processo autorregressivo não condicional de segunda ordem com distribuição marginal Gama-Lindley.

4.2 A CONSTRUÇÃO DE UM NOVO MODELO AUTORREGRESSIVO DE SEGUNDA ORDEM

Admitindo a estacionariedade de uma ST, a transformada de Laplace para o erro aleatório (LAWRANCE; LEWIS, 1980), é dada por:

$$\phi_\epsilon(s) = \frac{\phi_X(s)}{[(1 - \alpha_2)\phi_X(\alpha_1 s) + \alpha_2\phi_X(\alpha_2 s)]}, \quad (4.2)$$

em que

$$\phi_X(\alpha_1 s) = \frac{\lambda^2(\beta\lambda + \beta + \alpha_1 s)}{\beta(1 + \lambda)(\alpha_1 s + \lambda)^2} \quad \text{e} \quad \phi_X(\alpha_2 s) = \frac{\lambda^2(\beta\lambda + \beta + \alpha_2 s)}{\beta(1 + \lambda)(\alpha_2 s + \lambda)^2}.$$

Substituindo na Equação (4.2) tem-se:

$$\phi_\epsilon(s) = \left[\frac{(\beta\lambda + \beta + s)(\alpha_1 s + \lambda)^2(\alpha_2 s + \lambda)^2}{(s + \lambda)^2[(1 - \alpha_2)(\beta\lambda + \beta + \alpha_1 s)(\alpha_2 s + \lambda)^2 + \alpha_2(\beta\lambda + \beta + \alpha_2 s)(\alpha_1 s + \lambda)^2]} \right]. \quad (4.3)$$

Da mesma forma como foi desenvolvido o modelo AR(2) com distribuição marginal Lindley, devido à tratabilidade e à complexidade das expressões envolvidas, a construção do modelo será realizada em duas etapas. Primeiramente, consideremos a restrição em que os parâmetros autorregressivos sejam iguais, na qual chamamos de Caso 1:

$$\alpha_1 = \alpha_2 = a, \quad a \in (0, 1).$$

Reescrevendo a Equação (4.2) tem-se:

$$\phi_\epsilon(s) = \left[\frac{(\beta\lambda + \beta + s)(as + \lambda)^2}{(s + \lambda)^2[(\beta\lambda + \beta + as)]} \right]. \quad (4.4)$$

Observe que a partir da Equação (4.4), a seguinte discussão mostra que a distribuição de inovação é uma mistura envolvendo a distribuição Gama. Essa ocorrência será expressa na proposição a seguir.

Proposição 4.2.1 *Seja $X_t \sim GLAR(\lambda, \beta)$ e $\alpha_1 = \alpha_2 = a$. Então, a distribuição de ϵ_t é uma mistura de distribuições Gama e exponenciais com parâmetros diferentes.*

$$X_t = a(X_{t-1} + X_{t-2}) + \epsilon_t,$$

$$\text{tal que } F_{\epsilon_t}(l) = T_1^* F_{\epsilon_{t,1}}(l) + T_2^* F_{\epsilon_{t,2}}(l) + T_3^* F_{\epsilon_{t,3}}(l)$$

em que $F_{\epsilon_t}(\cdot)$ é a fda de ϵ_t , com v.a componentes

$$\epsilon_{t,1} \sim \exp(\lambda), \quad \epsilon_{t,2} \sim \Gamma(2, \lambda) \quad \text{e} \quad \epsilon_{t,3} \sim \exp\left(\frac{\beta\lambda + \beta}{a}\right).$$

Os pesos T_1^* , T_2^* e T_3^* são dados por:

$$T_1^* = - \frac{\frac{\beta\lambda^2(a-1)^2(\lambda+1)(\beta-\lambda+\beta\lambda)}{\beta-a\lambda+\beta\lambda} - \beta\lambda^2(\lambda+1) + \frac{a\beta\lambda^2(a-1)(\lambda+1)(\beta-\lambda+\beta\lambda)^2}{(\beta-a\lambda+\beta\lambda)^2}}{\lambda^2(\beta + \beta\lambda)},$$

$$T_2^* = \frac{(a-1)^2(\beta - \lambda + \beta\lambda)}{\beta - a\lambda + \beta\lambda} \quad \text{e} \quad T_3^* = \frac{a(\beta - \lambda + \beta\lambda)^2 \left(\beta - \frac{\beta+\beta\lambda}{a} + \beta\lambda \right)}{\left(\lambda - \frac{\beta+\beta\lambda}{a} \right)^2 (\beta + \beta\lambda)}.$$

De forma similar ao $\text{LAR}_1(2)$, considerando os parâmetros autorregressivos iguais, esse modelo será chamado de:

$$X_t \sim \text{GLAR}_1(2).$$

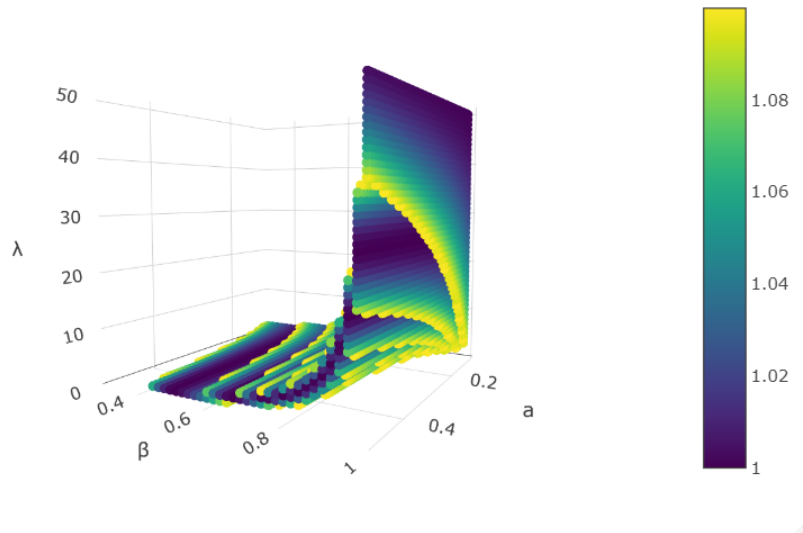
A Proposição a seguir trata sobre o espaço paramétrico dos parâmetros β e λ .

Proposição 4.2.2 *Considerando a soma dos pesos T_1^* , T_2^* e T_3^* e de acordo com a região de estabilidade para o parâmetro autorregressivo, λ e β obedecem a seguinte relação:*

$$\left\{ (\lambda, \beta) : 0 < \beta < 1 \quad \text{e} \quad \lambda = \frac{-\beta}{\beta - 1} \right\}.$$

A demonstração dessa relação entre os parâmetros pode ser conferida no Apêndice E. A Figura 23 representa a região de estabilidade dos parâmetros β e λ variando o termo autorregressivo a . Como existe uma relação entre os parâmetros λ e β , dado o parâmetro a tem-se a região de λ .

Figura 23 – Região satisfeita para os parâmetros β e λ .



Fonte: Elaboração Própria (2025)

Dessa forma, para determinar a distribuição do erro, foi aplicada a transformada inversa de Laplace em $\phi_\epsilon(s)$, utilizando o método de frações parciais. Assim, obteve-se uma mistura que descreve a distribuição da v.a ϵ_t . A demonstração da Proposição 4.2.1 pode ser conferida no Apêndice D. Assim como o modelo $\text{LAR}_1(2)$, o parâmetro autorregressivo influencia diretamente nas condições de estacionariedade para o novo processo $\text{AR}(2)$. De modo similar, a

região de estacionariedade corresponde a área do primeiro quadrante (α_1, α_2) , com α_1 pertencente ao eixo das abscissas e α_2 pertencente ao eixo das ordenadas, em que todas as raízes estão fora no círculo unitário (ver Figura 2). De modo semelhante, para o modelo $\text{GLAR}_1(2)$ o sistema será estável se $\alpha_1 = \alpha_2 = a < 0.49$. Na Seção seguinte é apresentada as propriedades estruturais do novo modelo proposto.

4.2.1 Propriedades Estruturais

A seguir, é apresentada as propriedades como esperança e variância condicional, autocorrelação, autocorrelação parcial e densidade do erro para o modelo $\text{GLAR}(2)$.

4.2.1.1 Esperança e Variância Condicional

Essas propriedades são semelhantes as que foram desenvolvidas para os modelos $\text{LAR}(2)$. O que as diferem são as distribuições do erro aleatório ϵ_t , apesar que ambos os modelos correspondem a misturas de distribuições gamas e exponencias. Assim, da Equação (4.5) e considerando os parâmetros autorregressivos iguais, a esperança condicional é dada por:

$$\mathbb{E}(X_{t+k}|X_{t-1}, X_{t-2}) = a(X_{t-1} + X_{t-2}) + \mathbb{E}(\epsilon_t), \quad (4.5)$$

com $\mathbb{E}(\epsilon_t)$ o valor esperado da mistura, que é uma combinação ponderada de pesos referentes às suas distribuições. Como o processo é estacionário, tem-se:

$$\mathbb{E}(\epsilon_t) = (1 - 2a)\mathbb{E}(X_t),$$

com $\mathbb{E}(X_t)$ descrito na Equação (4.1). Assim, substituindo tem-se:

$$\mathbb{E}(\epsilon_t) = (1 - 2a) \left(\frac{2\beta(1 + \lambda) - \lambda}{\lambda\beta(1 + \lambda)} \right).$$

Do mesmo modo, tem-se a variância condicional é dada por:

$$\text{Var}(X_t) = \alpha_1 \gamma_X(1) + \alpha_2 \gamma_X(2) + \text{Var}(\epsilon_t)$$

com,

$$\text{Var}(\epsilon_t) = \frac{(1 + \alpha_2)[1 - (\alpha_2 + \alpha_1)][1 - (\alpha_2 - \alpha_1)]\text{Var}(X_t)}{1 - \alpha_2},$$

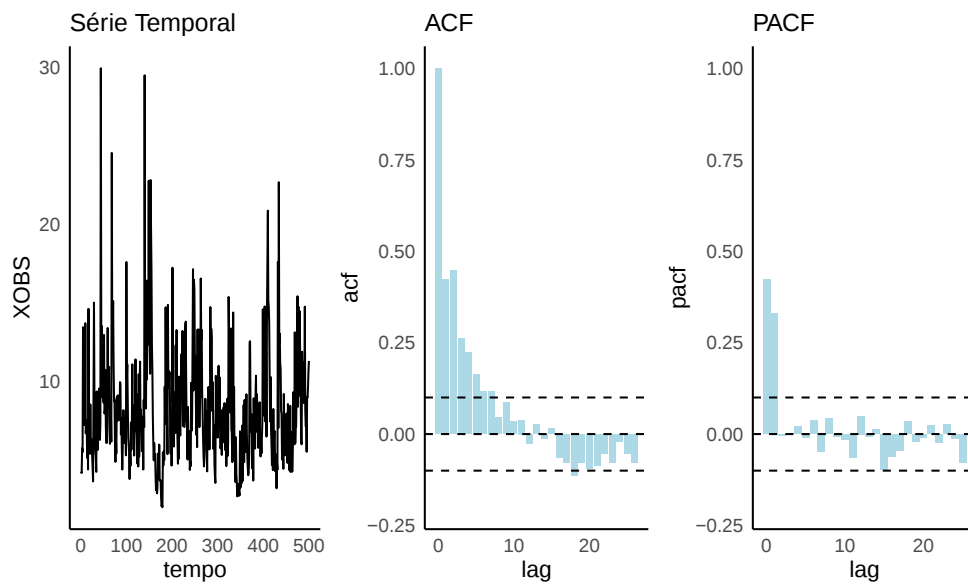
sendo $\text{Var}(X_t)$, a variância da distribuição Gama-Lindley. A variância do erro ϵ_t é dada por:

$$\sigma_\epsilon^2 = \frac{T_1^{*2}}{\hat{\lambda}^2} + \frac{2T_2^{*2}}{\hat{\lambda}^2} + \frac{T_3^{*2}\hat{a}^2}{\hat{\beta}^2(\hat{\lambda} + 1)^2}.$$

4.2.1.2 Autocorrelação, Autocorrelação Parcial e Densidade do erro

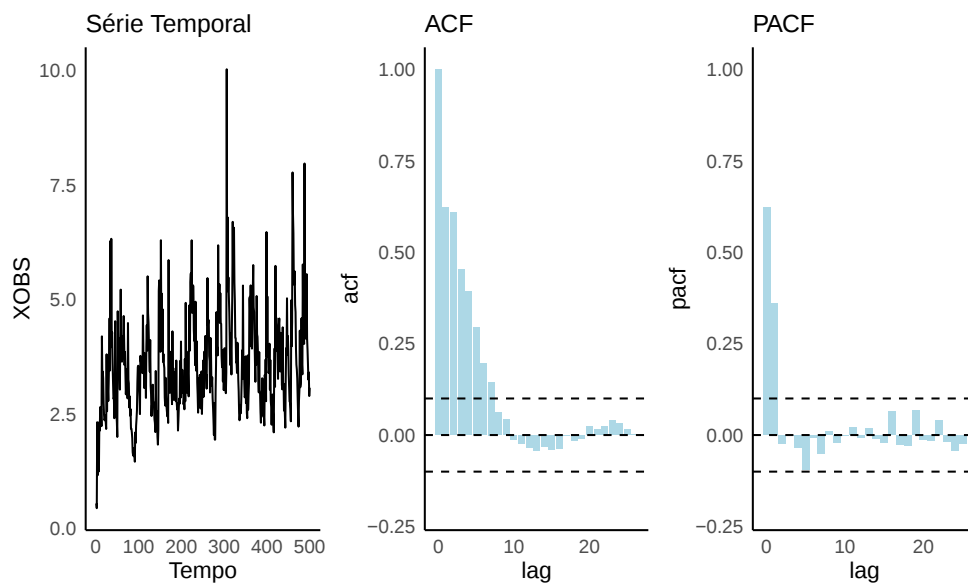
Foi implementada, no software R, uma função geradora de números pseudo-aleatórios para construir a ST correspondente ao modelo $GLAR_1(2)$. As Figuras 24 e 25 apresentam o gráfico da série, autocorrelação e autocorrelação parcial, nessa ordem.

Figura 24 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 com os valores $a = 0.3, \lambda = 1, \beta = 0.5$.



Fonte: Elaboração Própria (2025)

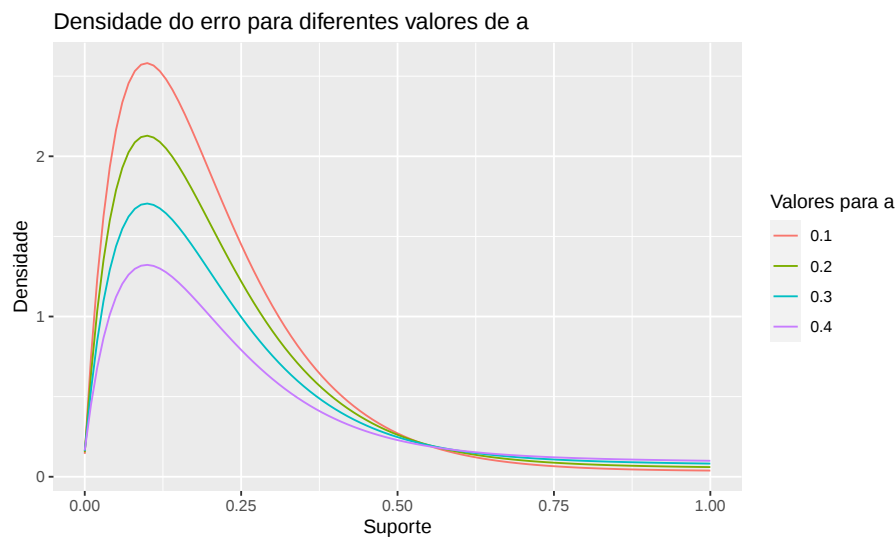
Figura 25 – Gráficos da série, função de autocorrelação e autocorrelação parcial para amostras de tamanho 500 com os valores $a = 0.4, \lambda = 1.5, \beta = 0.6$.



Fonte: Elaboração Própria (2025)

Nota-se claramente, que os dois gráficos da autocorrelação parcial apresentam significância estatística nos dois primeiros *lags*. Os gráficos da autocorrelação possuem um decaimento exponencial. Certamente, por essas características, tem-se evidências de que se trata de um processo AR(2). Na Figura 26 é apresentado o gráfico da densidade do erro. Fixou-se os parâmetros $\beta = 0.5$ e $\lambda = 0.1$, variando o parâmetro autorregressivo.

Figura 26 – Gráfico da densidade do erro para amostra de tamanho 500, com $\lambda = 0.1$, $\beta = 0.5$, variando o parâmetro a .



Fonte: Elaboração Própria (2025)

É fácil perceber que, à medida que aumentamos os valores do parâmetro a , a distribuição tende a ter caudas mais pesadas.

4.2.2 Previsão

Para prever valores futuros da série estacionária X_t , considerando os parâmetros do modelo conhecidos, a equação de predição é dada pela Esperança Condicional. Dessa forma, tem-se:

$$x_n^{n+1} = \mathbb{E}(X_{n+1} | X_{t-1}, X_{t-2}),$$

com a esperança condicional para o processo, determinada pela Equação:

$$\mathbb{E}(X_{t+k} | X_{t-1}, X_{t-2}) = a(X_{t-1} + X_{t-2}) + \mathbb{E}(\epsilon_t).$$

Ainda podemos reescrever a expressão acima substituindo $\mathbb{E}(\epsilon_t)$

$$\mathbb{E}(X_{t+k} | X_{t-1}, X_{t-2}) = a(X_{t-1} + X_{t-2}) + (1 - 2a)\mathbb{E}(X_t).$$

Para o modelo $\text{GLAR}_1(2)$ temos:

$$\mathbb{E}(X_{t+k}|X_{t-1}, X_{t-2}) = a(X_{t-1} + X_{t-2}) + (1 - 2a) \left(\frac{2\beta(1 + \lambda) - \lambda}{\lambda\beta(1 + \lambda)} \right).$$

Dessa forma, consegue-se estimar os parâmetros do novo modelo utilizando o método dos MQC, visto na próxima seção.

4.2.3 Estimação

Para estimar os parâmetros do modelo $\text{GLAR}_1(2)$, dois métodos de estimação são utilizados: O método Yule-Wallker, para estimar os parâmetros autorregressivos e o método dos MQC, para estimar os parâmetros que correspondem a variável de saída do modelo.

4.2.3.1 Mínimos Quadrados Condicionais

O método de estimação por Mínimos Quadrados consiste em minimizar a soma dos quadrados dos desvios ao valor esperado condicional, veja Klimko e Nelson (1978). Assim,

$$Q_n(a, \beta, \lambda) = \sum_{t=3}^n \left[X_t - \mathbb{E}(X_t|X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}) \right]^2.$$

Desse modo, para estimar os parâmetros a , λ e β do novo modelo $\text{GLAR}_1(2)$ é necessário minimizar a função:

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

4.2.3.2 Método Yule-Walker

Assim como no modelo $\text{LAR}_1(2)$ esse método apresenta boa eficiência ao estimar os parâmetros autorregressivos por meio das autocorrelações amostrais. Dessa forma, usaremos a Expressão (3.15), na qual encontrou-se a forma fechada para estimar o parâmetro autorregressivo a .

4.2.4 Resultado das Simulações

Foram feitas simulações de Monte Carlo com o objetivo de determinar as estimativas para os parâmetros a , λ e β . Para estimar o termo autorregressivo a , utilizou-se o método de

Yule-Walker. Os parâmetros λ e β foram estimados por meio do método MQC, com 1000 réplicas de Monte Carlo, quantidade suficiente para garantir a convergência dos estimadores e a estabilidade dos resultados para cada caso.

Mesmo obtendo uma forma fechada para estimar o parâmetro a , os resultados das estimativas foram bem inferiores as encontradas pelo Método de Yule-Walker. Por isso, eles foram suprimidos via MQC.

Assim, os resultados obtidos na Tabela 3 apresentam as simulações referentes ao parâmetro autorregressivo a pelo método de Yule-Walker e os parâmetros β e λ , utilizando o método dos MQC. Foram considerados 4 cenários distintos com 50, 100 e 150 observações. Em cada configuração, variou-se os parâmetros autorregressivos, considerando a região de estabilidade para o $GLAR_1(2)$.

Nota-se, que em todos os cenários, as estimativas para os parâmetros autorregressivos foram satisfatórias, convergindo mais rapidamente à medida que se aumenta o parâmetro a . Da mesma forma pode-se observar que o método dos MQC é uma alternativa satisfatória para estimar os parâmetros λ e β .

Assim, analisando a região de estabilidade para o modelo $GLAR_1(2)$, à proporção que aumenta o valor do parâmetro autorregressivo, nos dois métodos, as estimativas convergem mais rapidamente para os verdadeiros valores dos parâmetros. Dessa forma, o último cenário testado obteve os melhores resultados para estimar tanto o parâmetro autorregressivo quanto os parâmetros β e λ . Os cenários foram feitos de acordo com a ordem (a, λ, β) a seguir.

Tabela 3 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados

Cenários	Yule-Walker ($h = 1$)			Mínimos Quadrados			
	N	$\hat{a}$	EQM($\hat{a}$)	$\hat{\lambda}$	$\hat{\beta}$	EQM($\hat{\lambda}$)	EQM($\hat{\beta}$)
(0.1, 0.1, 0.5)	50	0.0039	0.0263	0.1635	0.4490	0.0056	0.0092
	100	0.0744	0.0091	0.1583	0.4547	0.0039	0.0051
	150	0.0850	0.0077	0.1580	0.4575	0.0037	0.0045
(0.2, 0.1, 0.5)	50	0.0961	0.0345	0.1520	0.4495	0.0038	0.0070
	100	0.1734	0.0233	0.1455	0.4553	0.0027	0.0072
	150	0.1878	0.0219	0.1439	0.4535	0.0027	0.0067
(0.3, 0.1, 0.5)	50	0.2077	0.0587	0.1413	0.4617	0.0064	0.0027
	100	0.2767	0.0472	0.1336	0.4543	0.0019	0.0010
	150	0.2864	0.0462	0.1339	0.4568	0.0018	0.0088
(0.4, 0.1, 0.5)	50	0.3563	0.0852	0.1322	0.4692	0.0019	0.0031
	100	0.3893	0.0807	0.1277	0.4677	0.0012	0.0037
	150	0.3940	0.0805	0.1275	0.4681	0.0010	0.0030

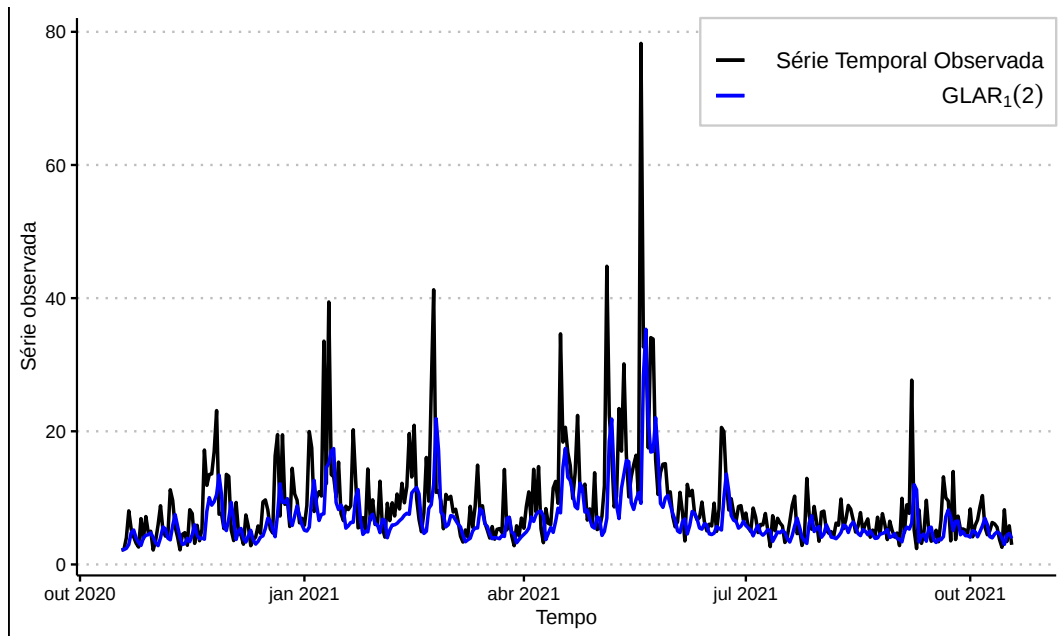
Fonte: Elaboração Própria (2025)

4.2.5 Aplicação

Com o intuito de observar a capacidade de ajuste do modelo $GLAR_1(2)$, aplicou-se um conjunto de dados reais na área do mercado financeiro por meio das moedas bitcoins. O mesmo banco de dados utilizado para avaliar o desempenho do modelos $LAR(2)$.

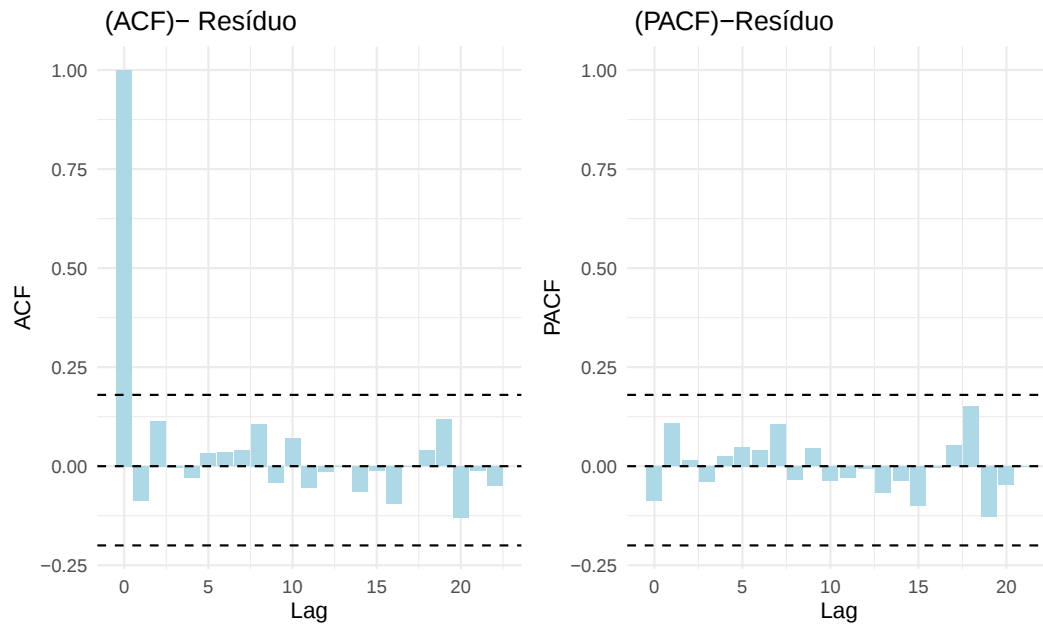
Em vista disso, após construir a forma fechada para estimar o parâmetro a e realizar as estimativas para os parâmetros λ e β , por meio da esperança condicional, fez-se a predição do modelo $GLAR_1(2)$. A Figura 27 representa as séries geradas e seu ajuste ao conjunto de dados. Nota-se que o novo modelo se ajusta bem ao conjunto de dados observados.

Figura 27 – Gráfico para as séries observadas e $GLAR_1(2)$



Fonte: Elaboração Própria (2025)

A Figura 28 apresenta os gráficos da autocorrelação e autocorrelação parcial do resíduo, indicando que não há correlação significativa entre eles para diferentes defasagens. Desse modo, foi utilizado o Box Test, que é um teste estatístico válido para verificar a presença de autocorrelação em uma ST. Assim, conclui-se que não existem evidências para rejeitar a hipótese nula, ou seja, não há autocorrelação até uma dada defasagem. Portanto, tem-se vestígios de que o resíduo tenha características de um ruído branco, já que o ruído branco não apresentam correlações significativas entre si.

Figura 28 – Análise dos resíduos para o $GLAR_1(2)$ 

Fonte: Elaboração Própria (2025)

4.2.6 Considerações Parciais

Neste Capítulo, um novo modelo autorregressivo de ordem 2 não condicional com distribuição marginal Gama-Lindley está sendo desenvolvido. Devido à complexidade matemática e tratabilidade, considerou-se que os parâmetros autorregressivos sejam iguais. Dessa forma, utilizando técnicas matemáticas como a transformada de Laplace, foi proposto o novo modelo, nessas condições, chamado de $GLAR_1(2)$.

Desse modo, encontrou-se a distribuição para o erro aleatório como sendo uma mistura de distribuições $Gama(2, \lambda)$, $\exp(\lambda)$ e $\exp\left(\frac{\beta\lambda+\beta}{a}\right)$. Assim, como foi desenvolvido para o modelo $LAR_1(2)$, o estudo se aprofundou em desenvolver propriedades matemáticas, como esperança e variância condicional, ACF, PACF, equação de previsão e um gerador para o erro aleatório.

Com a finalidade de estimar os parâmetros do modelo $GLAR_1(2)$ dois métodos de estimação foram utilizados: método Yule-Walker e os mínimos quadrados. O primeiro foi responsável para estimar os parâmetros autorregressivos a , em que foi deduzido a sua forma fechada, anteriormente. Pelo método dos mínimos quadrados, estimou os parâmetros provenientes da distribuição da inovação: β e λ . A avaliação do desempenho desses estimadores foi feita através de simulação de Monte Carlo. Com isso, constata-se que os dois métodos utilizados obtiveram bom desempenho ao estimar os parâmetros.

Por fim, aplicou-se o modelo proposto a um conjunto de dados reais no mercado de ações. O mesmo banco de dados utilizado para os modelos LAR(2), o que comprovou seu bom ajuste. Portanto, de acordo com essa base de dados, o modelo GLAR₁(2) pode ser eficiente ao fazer análises de riscos financeiros e tomadas de decisões, prevendo movimentos futuros de preços de ações e modelando as correlações de diferentes ativos financeiros. Na próxima seção, será apresentado esse novo modelo de forma mais ampla, no qual os parâmetros autorregressivos são diferentes.

4.3 CASO 2- PROCESSO AR(2) COM DISTRIBUIÇÃO MARGINAL GAMA-LINDLEY

De maneira semelhante ao que foi feito para o LAR₂(2), utilizou-se a transformada inversa de Laplace tendo a transformada para o erro aleatório de acordo com a Equação (4.3).

$$\phi_{\epsilon}(s) = \left[\frac{(\beta\lambda + \beta + s)(\alpha_1 s + \lambda)^2(\alpha_2 s + \lambda)^2}{(s + \lambda)^2[(1 - \alpha_2)(\beta\lambda + \beta + \alpha_1 s)(\alpha_2 s + \lambda)^2 + \alpha_2(\beta\lambda + \beta + \alpha_2 s)(\alpha_1 s + \lambda)^2]} \right].$$

Da mesma forma, tem-se indícios de que v.a ϵ_t possui a distribuição Gama em sua composição. A proposição a seguir demonstra isso.

Proposição 4.3.1 *Seja X_t um processo estacionário de média nula e entrada ϵ_t , com $\alpha_2 = c - \alpha_1$. Então, a distribuição de ϵ_t é uma mistura de distribuições gama e exponenciais com parâmetros diferentes. Ou seja:*

$$F_{\epsilon_t}(l) = T_1^* F_{\epsilon_t,1}(l) + T_2^* F_{\epsilon_t,2}(l) + T_{31}^* F_{\epsilon_t,3}(l) + T_{32}^* F_{\epsilon_t,4}(l) + T_{33}^* F_{\epsilon_t,5}(l),$$

em que $F_{\epsilon_t}(\cdot)$ é a fda de ϵ_t , com:

$$\epsilon_{t,1} \sim \exp(\lambda), \quad \epsilon_{t,2} \sim \Gamma(2, \lambda), \quad \epsilon_{t,3} \sim \exp(\varphi_1), \quad \epsilon_{t,4} \sim \exp(\varphi_2), \quad \epsilon_{t,5} \sim \exp(\varphi_3).$$

As raízes do polinômio de grau 3 são representadas por $r_i = -\varphi_i$, e aparecem ao aplicar a inversa de Laplace em:

$$(1 - k)(\beta\lambda + as + \beta)(\lambda + ks)^2 + k(\beta\lambda + ks + \beta)(\lambda + as)^2. \quad (4.6)$$

Novamente, considerou-se, por simplicidade, na Expressão (4.6) que $\alpha_1 = a$ e $\alpha_2 = c - \alpha_1 = k$. O polinômio, na sua forma expandida, em função dos parâmetros autorregressivos

(a, k) e dos parâmetros β e λ é da forma:

$$\begin{aligned} z(s) = & (a^2k^2 - ak^3 + ak^2)s^3 + (\beta k^2 - \beta k^3 + 2ak\lambda + a^2\beta k + \beta k^2\lambda - \beta k^3\lambda + a^2\beta k\lambda)s^2 \\ & + (a\lambda^2 + k^2\lambda^2 - 2\beta k^2\lambda^2 + 2\beta k\lambda - ak\lambda^2 + 2\beta k\lambda^2 - 2\beta k^2\lambda + 2a\beta k\lambda^2 + 2a\beta k\lambda)s \\ & + \beta\lambda^3 + \beta\lambda^2. \end{aligned}$$

Os pesos são dados por:

$$\begin{aligned} T_1^* = & - \left(\frac{(\lambda - k\lambda)^2 (k\lambda^2(\beta + \beta\lambda) - \lambda^2(k-1)(\beta + \beta\lambda)) (\lambda - a\lambda)^2 (\beta - \lambda + \beta\lambda)}{k(\lambda - a\lambda)^2 (\beta + \beta\lambda - k\lambda) - (\lambda - k\lambda)^2 (k-1)(\beta - a\lambda + \beta\lambda)} \right. \\ & \left. - \lambda^4 (\beta + \beta\lambda) + \frac{\lambda^2 (\lambda + kr)^2 (\lambda + ar)^2 (\beta + r + \beta\lambda)}{(\lambda + r)^2} \right) \\ & \div \left(\lambda^2 (k\lambda^2(\beta + \beta\lambda) - \lambda^2(k-1)(\beta + \beta\lambda)) \right), \end{aligned}$$

$$T_2^* = \frac{(\lambda - k\lambda)^2 (\lambda - a\lambda)^2 (\beta - \lambda + \beta\lambda)}{\lambda^2 (k(\lambda - a\lambda)^2 (\beta + \beta\lambda - k\lambda) - (\lambda - k\lambda)^2 (k-1)(\beta - a\lambda + \beta\lambda))},$$

$$T_3^* = \frac{(\lambda + kr)^2 (\lambda + ar)^2 (\beta + r + \beta\lambda)}{(\lambda + r)^2}$$

e

$$T_{31}^* = -\frac{T_3^*}{k_1\varphi_1} = -\frac{(\lambda + kr)^2 (\lambda + ar)^2 (\beta + r + \beta\lambda)}{\varphi_1(\lambda + r)^2 w^*},$$

com:

$$\begin{aligned} w^* = & \varphi_1 \left(2\beta k^2 - 2\beta k^3 + 4ak\lambda + 2a^2\beta k + 2\beta k^2\lambda - 2\beta k^3\lambda + 2a^2\beta k\lambda \right) + a\lambda^2 \\ & + \varphi_1^2 \left(3a^2k^2 - 3ak^3 + 3ak^2 \right) + k^2\lambda^2 - 2\beta k^2\lambda^2 + 2\beta k\lambda - ak\lambda^2 + 2\beta k\lambda^2 - 2\beta k^2\lambda, \end{aligned}$$

$$T_{32}^* = -\frac{T_3^*}{k_2\varphi_2} = -\frac{(\lambda + kr)^2 (\lambda + ar)^2 (\beta + r + \beta\lambda)}{\varphi_2(\lambda + r)^2 l^*},$$

com:

$$\begin{aligned} l^* = & \varphi_2 \left(2\beta k^2 - 2\beta k^3 + 4ak\lambda + 2a^2\beta k + 2\beta k^2\lambda - 2\beta k^3\lambda + 2a^2\beta k\lambda \right) + a\lambda^2 \\ & + \varphi_2^2 \left(3a^2k^2 - 3ak^3 + 3ak^2 \right) + k^2\lambda^2 - 2\beta k^2\lambda^2 + 2\beta k\lambda - ak\lambda^2 + 2\beta k\lambda^2 - 2\beta k^2\lambda. \end{aligned}$$

$$T_{33}^* = -\frac{T_3^*}{k_3\varphi_3} = -\frac{(\lambda + kr)^2(\lambda + ar)^2(\beta + r + \beta\lambda)}{\varphi_3(\lambda + r)^2m^*},$$

com:

$$m^* = \varphi_3 \left(2\beta k^2 - 2\beta k^3 + 4ak\lambda + 2a^2\beta k + 2\beta k^2\lambda - 2\beta k^3\lambda + 2a^2\beta k\lambda \right) + a\lambda^2 \\ + \varphi_3^2 \left(3a^2k^2 - 3ak^3 + 3ak^2 \right) + k^2\lambda^2 - 2\beta k^2\lambda^2 + 2\beta k\lambda - ak\lambda^2 + 2\beta k\lambda^2 - 2\beta k^2\lambda.$$

Assim como no modelo $LAR_2(2)$, os parâmetros das distribuições exponenciais são raízes negativas do polinômio de grau 3. Desse modo, com o auxílio do software Matlab essas raízes possuem a seguinte estrutura:

$$\left\{ R_1 = \Phi_1 - \frac{\Phi_7}{3a^2k^2 - 3ak^3 + 3ak^2} - \frac{\Phi_2}{\Phi_1} \right\},$$

$$\left\{ R_2 = \frac{\Phi_2}{2\Phi_1} - \frac{\Phi_7}{3a^2k^2 - 3ak^3 + 3ak^2} - \frac{\Phi_1}{2} - \frac{\sqrt{3} \left(\frac{\Phi_2}{\Phi_1} + \phi_1 \right) i}{2} \right\}$$

e

$$\left\{ R_3 = \frac{\Phi_2}{2\Phi_1} - \frac{\Phi_7}{3a^2k^2 - 3ak^3 + 3ak^2} - \frac{\Phi_1}{2} + \frac{\sqrt{3} \left(\frac{\Phi_2}{\Phi_1} + \Phi_1 \right) i}{2} \right\},$$

em que:

$$\Phi_1 = \left(\sqrt{(\Phi_4 + \Phi_5 - \Phi_3)^2 + \Phi_2^3} - \Phi_5 - \Phi_4 + \Phi_3 \right)^{\frac{1}{3}},$$

$$\Phi_2 = \frac{\Phi_6}{3a^2k^2 - 3ak^3 + 3ak^2} - \frac{\Phi_7^2}{9(a^2k^2 - ak^3 + ak^2)^2},$$

$$\Phi_3 = \frac{\Phi_7\Phi_6}{6(a^2k^2 - ak^3 + ak^2)^2}, \quad \Phi_4 = \frac{\beta\lambda^3 + \beta\lambda^2}{2a^2k^2 - 2ak^3 + 2ak^2}, \quad \Phi_5 = \frac{\Phi_7^3}{27(a^2k^2 - ak^3 + ak^2)^3},$$

$$\phi_6 = \lambda(a\lambda + 2\beta k - 2\beta k^2 + k^2\lambda + 2a\beta k - ak\lambda + 2\beta k\lambda - 2\beta k^2\lambda + 2a\beta k\lambda)$$

e

$$\phi_7 = \beta k^2 - \beta k^3 + 2ak\lambda + a^2\beta k + \beta k^2\lambda - \beta k^3\lambda + a^2\beta k\lambda.$$

A demonstração da Proposição 4.3.1 pode ser conferida no Apêndice F.

Considerando o caso em que os parâmetros autorregressivos são iguais: $\alpha_1 = \alpha_2 = 0.45$, de acordo com o Caso 2, utilizando o software Matemático Matlab, consegue obter uma expressão geral que representa as raízes do polinômio de terceiro grau de acordo com a Equação (4.3). Novamente observa-se que, além de todas as raízes serem negativas, existem duas raízes reais e iguais da forma:

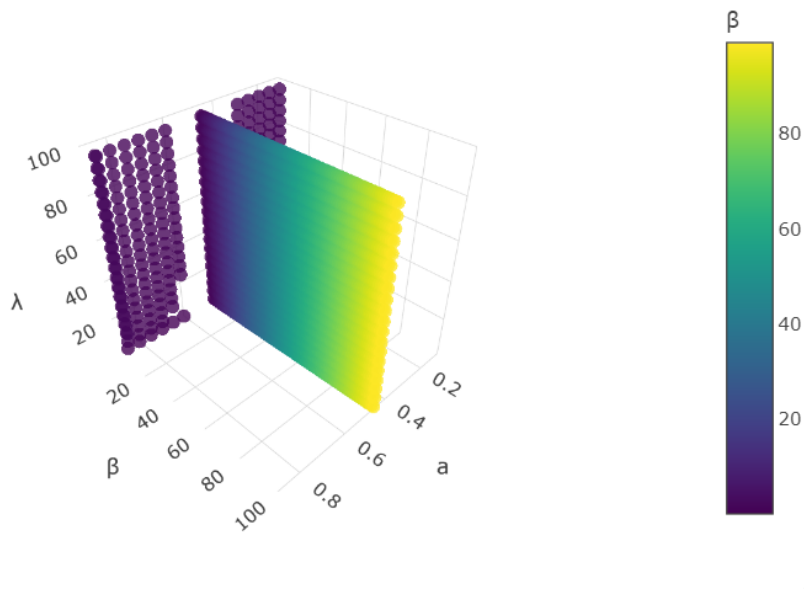
$$R_1 = \frac{-20\beta}{9}(\lambda + 1), \quad R_2 = R_3 = -\frac{20\lambda}{9}.$$

Note que a expressão que representa as raízes reais e iguais não depende do parâmetro β e possuem a mesma estrutura das raízes para os parâmetros autorregressivos iguais determinadas para o modelo $LAR_2(2)$. A primeira raiz, chamada de R_1 depende dos parâmetros β e λ .

A seguir, tem-se o espaço paramétrico que compõe a região dos parâmetros a , β e λ , em que a área em destaque representa a tripla ordenada que satisfaz as condições da v.a ϵ_t ser uma mistura. Essa região foi construída obedecendo cada uma das raízes reais do polinômio de terceiro grau da Expressão (4.6), sendo a parte imaginária desprezível da ordem 10^{-4} . Isso, porque para cada raiz escolhida possui uma região de estabilidade diferente a ser considerada. O parâmetro α_2 foi suprimido, considerando $\alpha_2 = k = c - \alpha_1$. A Figura 29 representa a região a , β e λ em que as raízes do polinômio de grau 3 são reais e negativas. Para a construção do gráfico, considerou-se uma restrição para os parâmetros β e λ de modo que ele pertence ao intervalo:

$$0.1 \leq \beta, \lambda < 100.$$

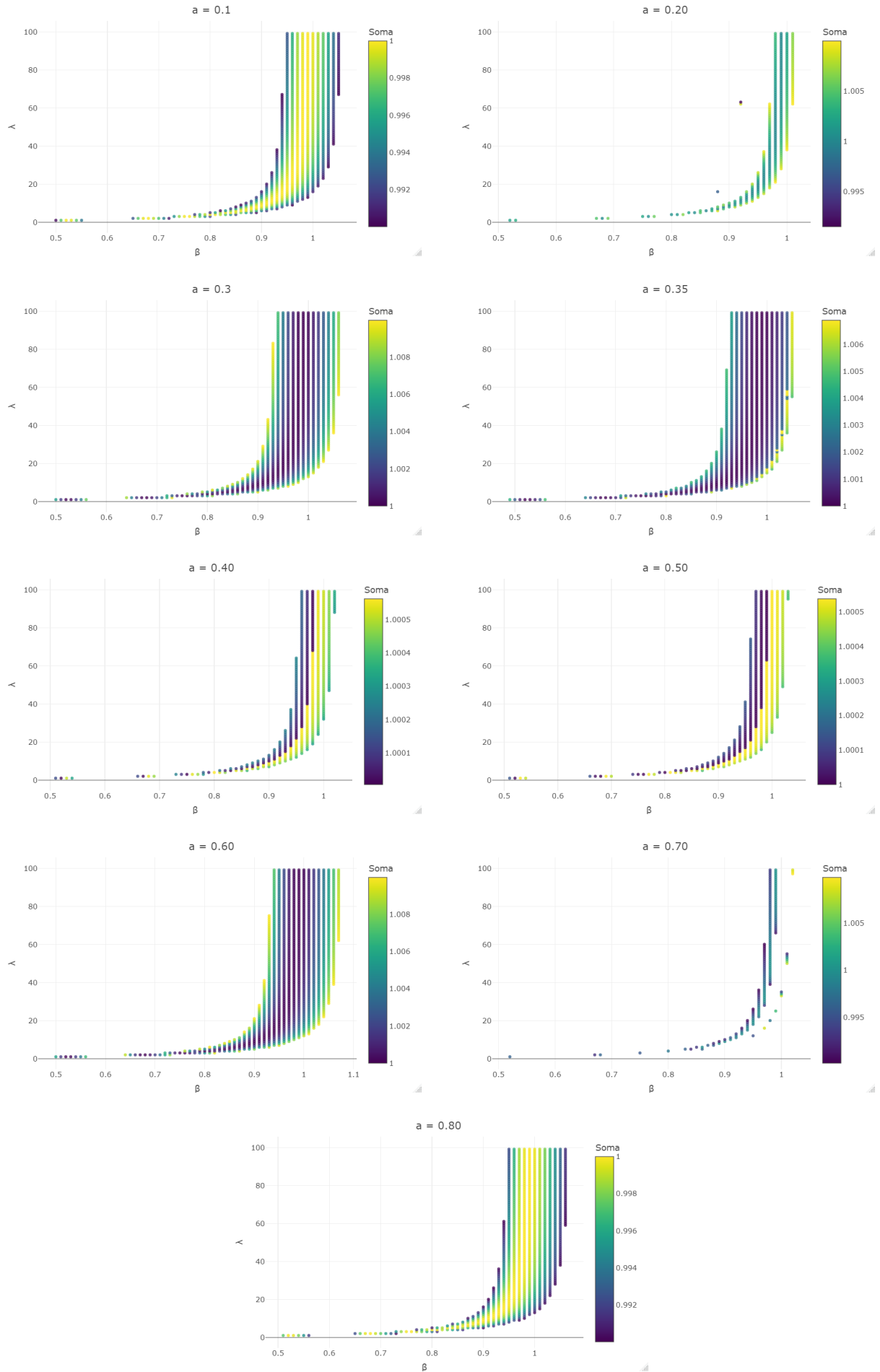
Figura 29 – Região em que as raízes do polinômio de grau 3 são reais



Fonte: Elaboração Própria (2025)

Na Figura 30 fixou-se valores para os parâmetros autorregressivos em 9 cenários diferentes, e a partir daí plotou o comportamento dos parâmetros β e λ . Na maioria dos casos, β está compreendido entre 0.1 e 0.99. Apenas considerando o último cenário, para $\alpha_1 = a = 0.80$ e $\alpha_2 = 0.10$ os valores de β variam entre 0.1 e 1.25. Os gráficos foram construídos no R, considerando $0.1 \leq \beta \leq 20$ e $0.1 \leq \lambda \leq 100$.

Figura 30 – Espaço paramétrico para os parâmetros β e λ .



Fonte: Elaboração Própria (2025)

Da mesma forma como construído para o modelo $LAR_2(2)$, o usuário pode determinar o espaço parâmetro para os parâmetros β e λ , com o parâmetro $\alpha_1 = a$ fixado.

1. Utilizando a função soma determinada pela soma dos pesos em função dos parâmetros a , k , β e λ .
2. Calcula-se as raízes reais do polinômio de grau 3, considerando a parte imaginária desprezível da ordem 10^{-5} .
3. O código contém 2 loops for aninhados com um a fixo que iteram sobre diferentes valores de β e λ :
 - O parâmetro β varia de 0.1 até 20 com incremento de 0.01
 - O parâmetro λ varia de 0.1 até 100 com incremento de 1

Para cada valor de β , o algoritmo testa diferentes combinações de λ . Por meio de um comando de condição, se a soma for próxima de 1, imprime-se os valores de β e λ , o resultado da soma e as 3 raízes reais.

4. O código gera um gráfico iterativo em que:

- Eixo y representa λ
- Eixo x representa os valores de β

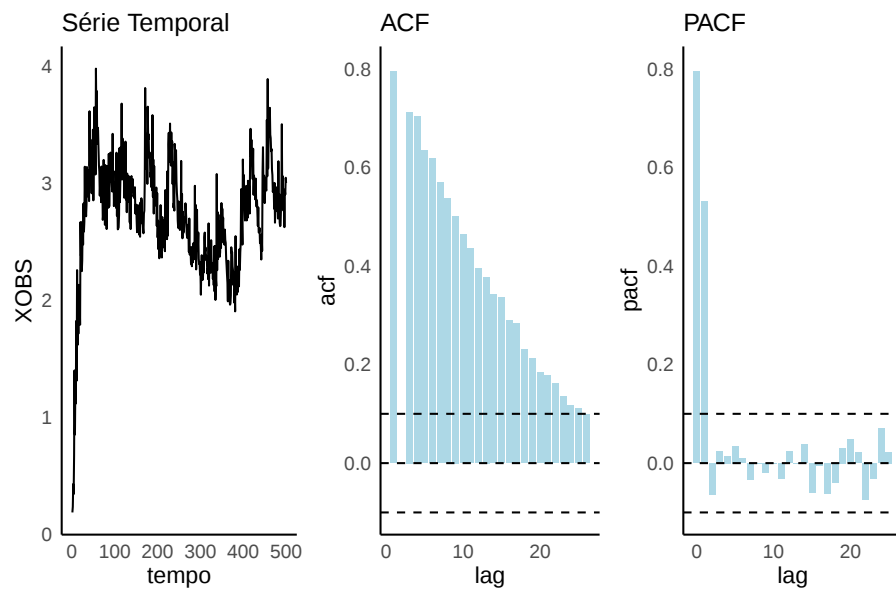
Por meio da função `Tooltip`, quando o usuário posiciona o cursor sobre o mouse, mostra-se os valores de λ , β , a soma e os valores das 3 raízes reais.

Note que, nas configurações da Figura 30 aparecem “buracos” no espaço paramétrico que podem estar relacionados às condições essenciais do modelo. Isso, porque, para v.a ϵ_t ser uma mistura, a soma dos pesos precisa ser igual a 1. Pequenas variações de β e λ podem fazer a soma sair da faixa aceitável, gerando descontinuidade da região, em que nenhuma combinação de parâmetros atende ao critério. Outro ponto a ser considerado, refere-se ao polinômio característico descrito na Equação (4.6) que estabelece que todas as raízes sejam reais e negativas. No entanto, para certos valores de β e λ as raízes podem ser complexas ou positivas, o que indica instabilidade do sistema. Assim, essas condições explicam a formação de buracos no espaço paramétrico.

4.3.1 Propriedades Estruturais

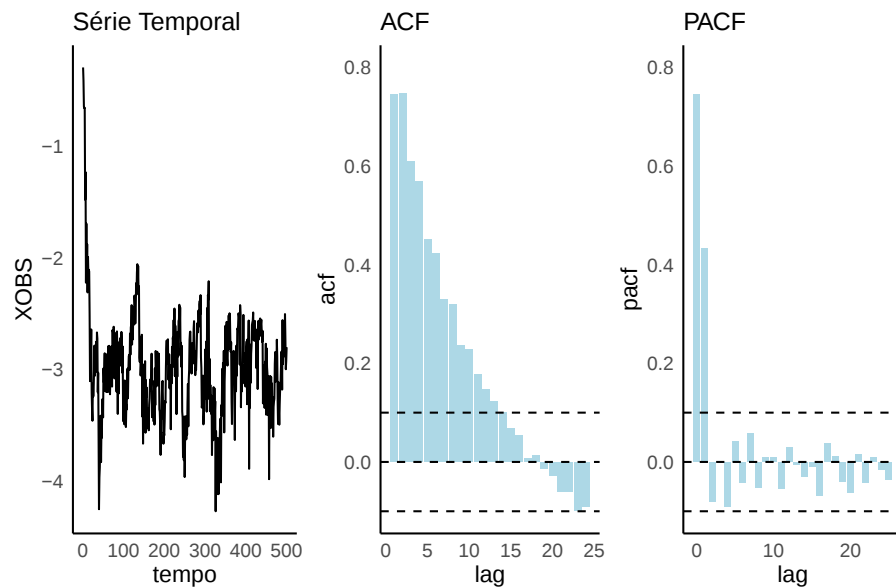
Semelhante aos modelos anteriores, após a construção do gerador para o erro aleatório, são apresentados abaixo os gráficos da ST, da ACF, PACF e da densidade do erro, considerando configurações de parâmetros que respeitem a região de estabilidade do modelo.

Figura 31 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.25$, $\alpha_2 = 0.65$, $\beta = 0.85$, $\lambda = 5$.



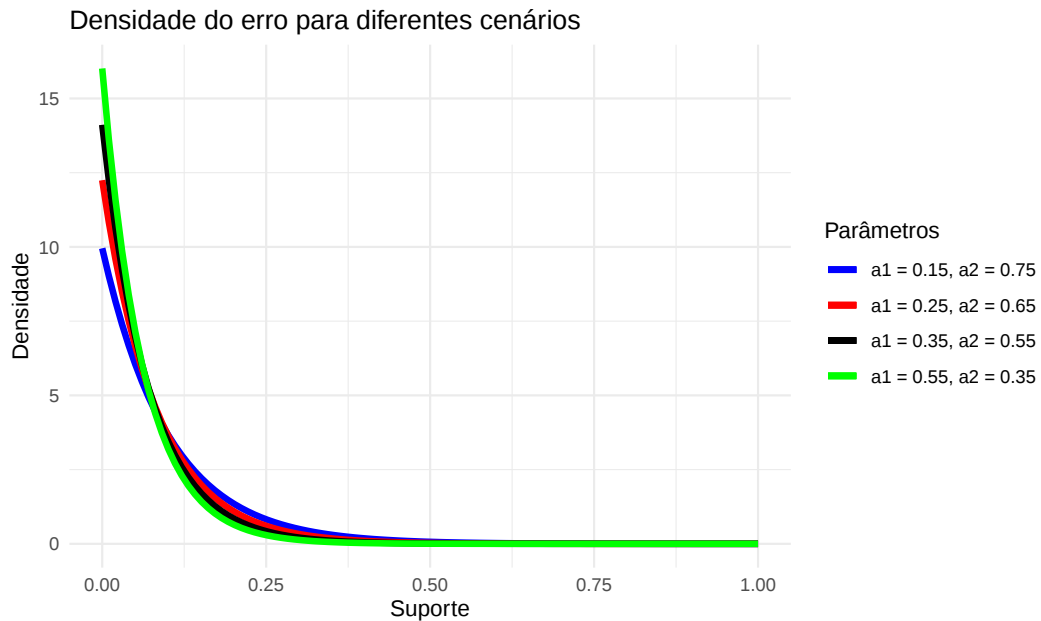
Fonte: Elaboração Própria (2025)

Figura 32 – Gráficos da série, função de autocorrelação e autocorrelação parcial de amostras de tamanho 500 para os valores $\alpha_1 = 0.32$, $\alpha_2 = 0.58$, $\beta = 0.93$, $\lambda = 18$.



Fonte: Elaboração Própria (2025)

Figura 33 – Gráficos da densidade do erro para amostra de tamanho 500, com $\lambda = 10$, $\beta = 0.92$ e diferentes valores para os autorregressivos.



Fonte: Elaboração Própria (2025)

Nota-se que, nas Figuras 31 e 32, os gráficos das ACF possuem um decaimento exponencial, já os gráficos das PACF apresentam um pico significativo apenas nos 2 primeiros *lags*, características de processos AR(2). Já na Figura 33, nota-se que as curvas seguem uma tendência decrescente, indicando que os erros estão mais concentrados, próximos de zero. Ao comparar os diferentes cenários, nota-se que à medida que $\alpha_1 = a_1$ aumenta e $\alpha_2 = a_2$, do cenário azul até o verde, a densidade do erro se torna mais concentrada, evidenciando uma menor variabilidade dos erros. Além disso, verifica-se que os parâmetros β e λ , quando aumentados, também contribuem para redução da dispersão dos erros.

4.3.2 Simulação

Com a finalidade de estimar os parâmetros do *Caso 2* para o modelo GLAR₂(2), utilizou-se também o método Yule-Walker, em que foi desenvolvido uma expressão geral para encontrar as estimativas dos parâmetros autorregressivos e MQC, utilizado para estimar todos os parâmetros.

4.3.2.1 Mínimos Quadrados Condicionais

Na construção dos modelos feitos para o Caso 1 foram desenvolvidas várias propriedades estruturais, como o valor esperado condicional para o caso em geral. Dessa forma, para o modelo AR(2) com distribuição marginal Gama-Lindley, tem-se:

$$Q_n(\alpha_1, \alpha_2, \beta, \lambda) = \sum_{t=3}^n \left(X_t - \alpha_1(X_{t-1}) - \alpha_2(X_{t-2}) - (1 - \alpha_1 - \alpha_2) \frac{2\beta(1 + \lambda) - \lambda}{\lambda\beta(1 + \lambda)} \right)^2.$$

4.3.2.2 Resultados das Simulações

Dessa forma, tendo o valor esperado condicional e as formas fechadas para estimar os termos autorregressivos, foram feitas as estimativas para os parâmetros. Nesse processo, fixou-se os valores para os parâmetros β e λ , variando os parâmetros autorregressivos. As simulações ocorreram em 5 cenários por meio de 1000 réplicas de Monte Carlo. A Tabela 4 indica os resultados das estimativas obtidos pelos método Yule-Walker e MQC.

Novamente os resultados da estimação dos parâmetros autorregressivos pelo método Yule-Walker foram superiores aos registrados pelo métodos dos MQC. No entanto, observa-se a influência dos parâmetros autorregressivos nos resultados das estimativas dos parâmetros λ e β . Nota-se que à medida que o valor do termo autorregressivo (α_1) aumenta, as estimativas para o β tendem a convergir para os valores verdadeiros, pelo método dos MQC. Entretanto ocorre o inverso para as estimativas do parâmetro λ . Mesmo assim, em todos os cenários obteve-se bons resultados na estimação dos parâmetros. Os cenários foram construídos seguindo a ordem dos parâmetros $\alpha_1, \alpha_2, \beta$ e λ , dentro da região de estabilidade.

Tabela 4 – Estimativas para os parâmetros do modelo pelos métodos de Yule-Walker e Mínimos Quadrados

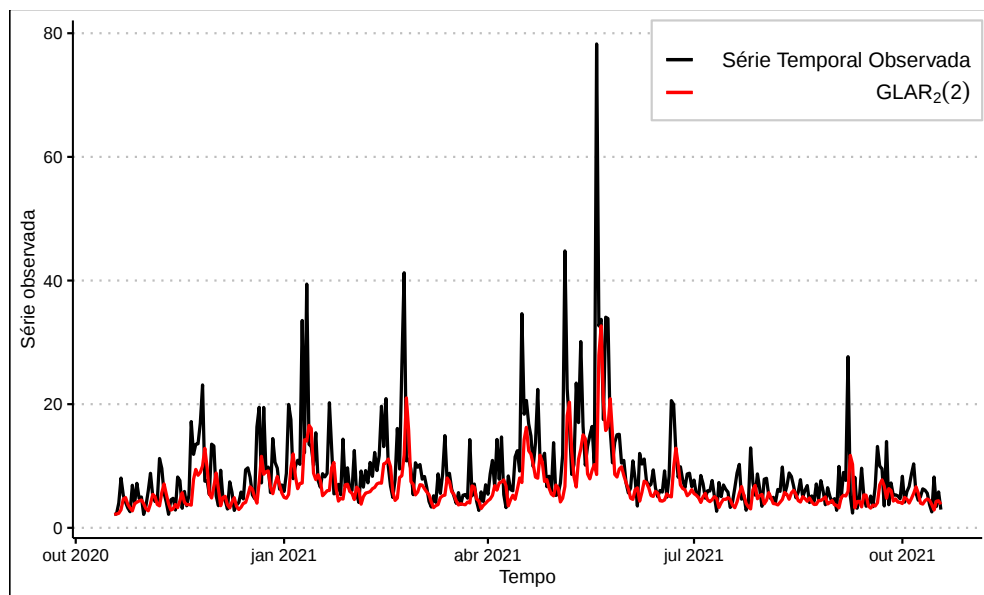
Cenários	N	Yule-Walker				Mínimos Quadrados							
		$\hat{\alpha}_1$	$EQM(\hat{\alpha}_1)$	$\hat{\alpha}_2$	$EQM(\hat{\alpha}_2)$	$\hat{\alpha}_1$	$EQM(\hat{\alpha}_1)$	$\hat{\alpha}_2$	$EQM(\hat{\alpha}_2)$	$\hat{\beta}$	$EQM(\hat{\beta})$	$\hat{\lambda}$	$EQM(\hat{\lambda})$
(0.11, 0.79, 18, 0.95)	50	0.2093	0.0202	0.6044	0.0464	0.5037	0.1550	0.5064	0.0804	0.7997	0.0273	18.1119	0.2082
	100	0.1685	0.0072	0.6916	0.0151	0.5000	0.1527	0.5015	0.0831	0.9008	0.0100	18.0091	0.0160
	150	0.1614	0.0055	0.7170	0.0091	0.5001	0.1521	0.5003	0.0839	0.9244	0.0035	17.9979	0.0040
(0.21, 0.69, 18, 0.95)	50	0.2780	0.0289	0.5701	0.0350	0.5044	0.0867	0.5059	0.0335	0.8370	0.0208	18.0601	0.1081
	100	0.2453	0.0134	0.6330	0.0141	0.5011	0.0847	0.5022	0.0352	0.8882	0.0124	18.0075	0.0081
	150	0.2331	0.0091	0.6488	0.0090	0.5004	0.0843	0.5010	0.0357	0.8920	0.0118	18.0067	0.0043
(0.31, 0.59, 18, 0.95)	50	0.3395	0.0346	0.5160	0.0329	0.5062	0.0387	0.5054	0.0073	0.8237	0.0229	18.1425	0.2723
	100	0.3342	0.0247	0.5459	0.0216	0.5023	0.0370	0.5013	0.0079	0.8563	0.0176	18.0403	0.0681
	150	0.3191	0.0167	0.5648	0.0140	0.5014	0.0366	0.5003	0.0080	0.8746	0.0145	18.0330	0.0523
(0.41, 0.49, 18, 0.95)	50	0.3744	0.0362	0.4800	0.0286	0.5233	0.0135	0.4881	0.0006	0.7959	0.0277	18.8164	1.6250
	100	0.3909	0.0299	0.4898	0.0242	0.5117	0.0121	0.4861	0.0005	0.8718	0.0148	18.2684	0.5334
	150	0.4032	0.0250	0.4810	0.0202	0.5165	0.0118	0.4853	0.0004	0.9087	0.0084	18.1422	0.2842
(0.51, 0.49, 18, 0.95)	50	0.4147	0.0456	0.4393	0.0321	0.5412	0.0014	0.4697	0.0068	0.7807	0.0304	19.4600	2.9168
	100	0.4371	0.0383	0.4385	0.0284	0.5376	0.0012	0.4659	0.0062	0.9187	0.0074	18.2430	0.4884
	150	0.4507	0.0345	0.4291	0.0257	0.5372	0.0011	0.4645	0.0060	0.9339	0.0048	18.1038	0.2122

Fonte: Elaboração Própria (2025)

4.3.3 Aplicação

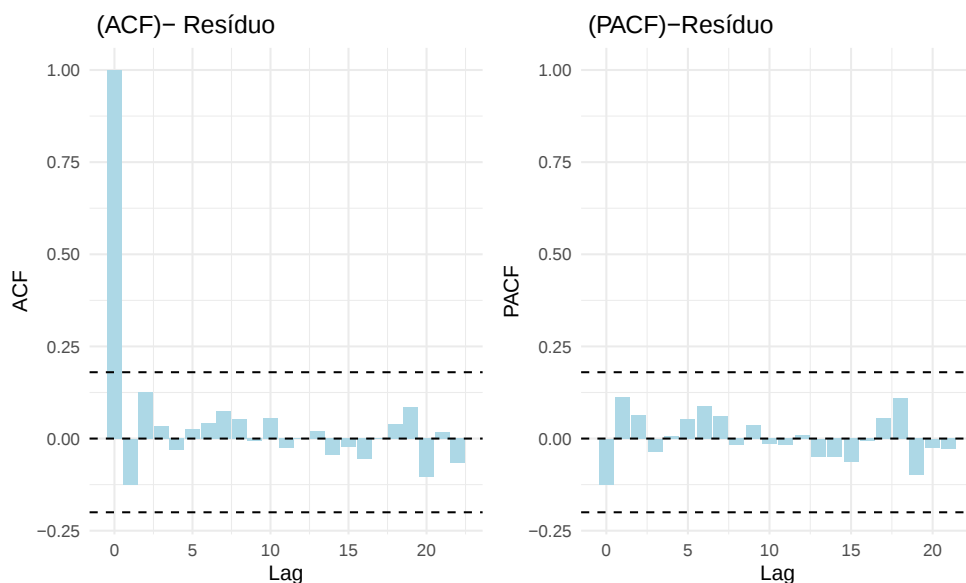
Foram utilizadas as mesmas bases de dados, que avaliaram a capacidade preditiva dos modelos anteriores. Dessa maneira, a Figura 34 representa o valor da série ajustada por predição a série observada, para o primeiro conjunto de dados sobre a bolsa de valores. A Figura 35 apresenta a ACF e PACF para os resíduos, indicando que não existe dependência.

Figura 34 – Gráfico para as séries observadas e $GLAR_2(2)$



Fonte: Elaboração Própria (2025)

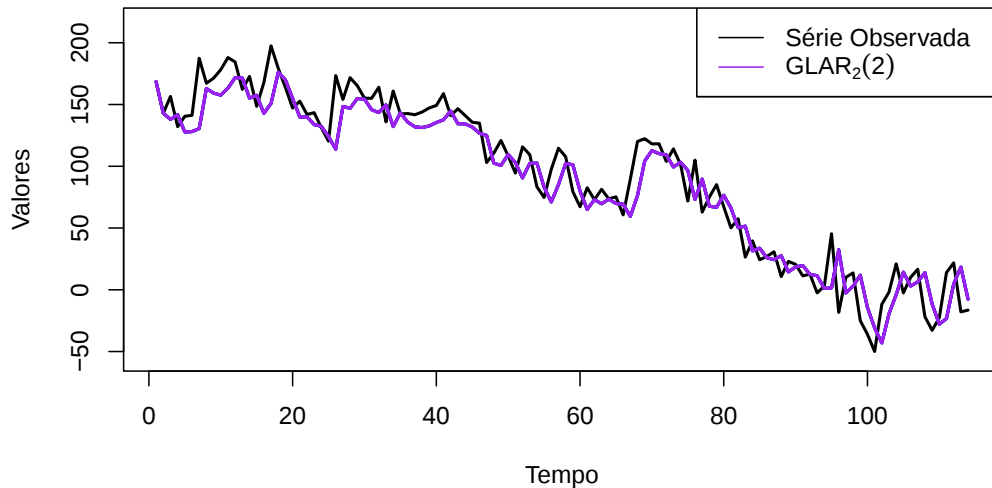
Figura 35 – Análise dos resíduos para o GALAR(2)



Fonte: Elaboração Própria (2025)

Da mesma forma, utilizou-se o conjunto de dados em hidrologia, de tal forma como foi usado para avaliar a capacidade preditiva do modelo $LAR_2(2)$. A Figura 36 representa o gráfico da série observada juntamente com o modelo $GLAR_2(2)$.

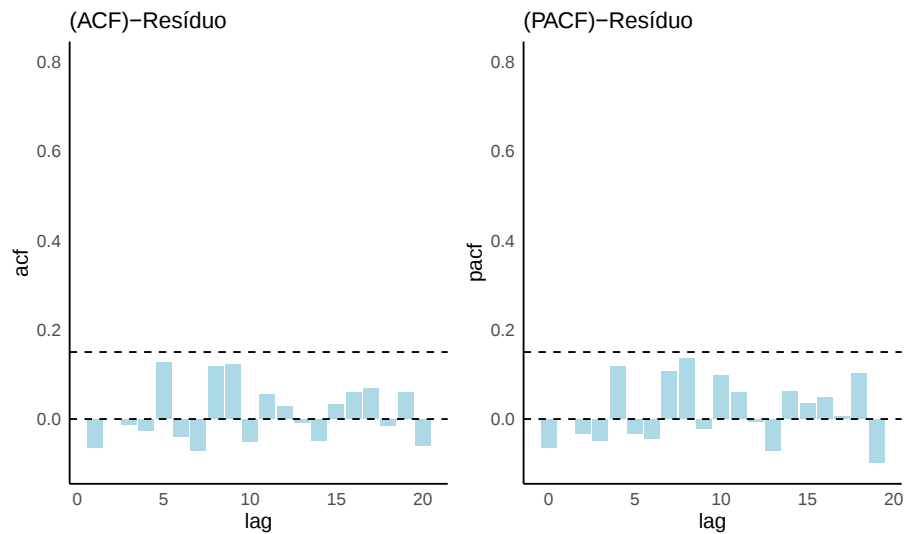
Figura 36 – Gráfico para as séries observadas e $GLAR_2(2)$



Fonte: Elaboração Própria (2025)

Nota-se que o modelo apresentou um bom ajuste a esse conjunto de dados reais conforme verifica-se a proximidade entre as curvas da série observada e a série do novo modelo obtida por previsão. Além do mais, a Figura 37 representa os gráficos da ACF e PACF.

Dessa forma, observa-se a ausência de autocorrelações significativas, comportando-se como um ruído branco. Para a análise dos modelos, trata-se de uma boa evidência, comprovando que a sua estrutura capturou adequadamente os padrões dos dados. O teste de Ljung-Box foi aplicado, e os resultados indicam que não se pode rejeitar a hipótese nula de independência dos resíduos. Portanto, conclui-se que o modelo demonstra eficiência satisfatória para representar a série em análise.

Figura 37 – Análise dos resíduos do GLAR₂(2)

Fonte: Elaboração Própria (2025)

4.3.4 Considerações Parciais

Neste Capítulo, foi apresentado um modelo autorregressivo não condicional de ordem 2 com distribuição marginal Gama-Lindley. Com isso, após análise para o *Caso 1*, considerou-se o *Caso 2*, com parâmetros autorregressivos diferentes. Sendo assim, a distribuição de inovação para esse modelo, assim como nos modelos LAR(2) é também uma mistura de distribuições Gama de parâmetros $(2, \lambda)$, exponencial com parâmetro λ e três exponenciais com parâmetros φ_1 , φ_2 e φ_3 que correspondem o oposto das raízes negativas de um polinômio de grau 3, conforme a Equação (4.3). Esse polinômio é função dos termos autorregressivos (α_1, α_2) e dos parâmetros adicionais (β, λ) .

Após essa etapa, foi construído o gerador para o erro aleatório, além de apresentar as propriedades estruturais como esperança condicional, variância condicional e densidade do erro. Para estimar os parâmetros autorregressivos do modelo utilizou-se também os métodos de estimação Yule-Walker. Pelo método dos MQC estimou-se além dos termos autorregressivos os parâmetros λ e β . Novamente o método Yule-Walker apresentou melhores resultados das estimativas dos parâmetros autorregressivos.

Além disso, percebe-se também a influência do termo autorregressivo para estimar os parâmetros β e λ . À medida que aumenta o termo α_1 as estimativas para o β tendem a convergir mais rapidamente para o valor verdadeiro. No entanto, ocorre o efeito contrário em relação à estimativa do parâmetro λ . Mesmo assim, o método dos MQC mostrou ser eficiente

para estimar os termos adicionais β e λ . Para analisar a capacidade preditiva do $\text{GLAR}_2(2)$, aplicou-se dois conjuntos de dados reais. O primeiro referente à cotação da bolsa de valores e o segundo sobre um conjunto de dados ligados à meteorologia. Em ambos os casos, o modelo $\text{GLAR}_2(2)$ obteve um bom ajuste. Dessa maneira, o modelo proposto é capaz de generalizar padrões presente a esses conjuntos de dados.

5 COMPARAÇÃO DOS MODELOS

Neste Capítulo, realiza-se uma comparação entre os dois modelos autorregressivos propostos, LAR(2) e GLAR(2). Para avaliar a eficiência desses modelos com base em um conjunto de dados reais, foram utilizados os seguintes critérios de seleção: o Critério de Informação de Akaike (AIC), o Critério de Informação Bayesiano (BIC) e a Raiz do Erro Quadrático Médio da Previsão (RMSE). Com o valor esperado condicional e a variância condicional desenvolvidos na Seção 4.2.1 para os novos modelos, é possível realizar a comparação utilizando a função de log-verossimilhança gaussiana. Por fim, os modelos LAR(2) e GLAR(2) serão comparados com outros modelos desenvolvidos na literatura, como o EAR(2), GAR(2) e INGAR(2). Na Seção 5.1 é feita uma revisão sobre os modelos propostos e a distribuição de inovação para cada um deles. Além disso, faz-se uma descrição sobre os critérios utilizados para comparar os modelos, além de uma breve análise sobre os modelos propostos na literatura. Na Seção 5.2 tem-se os resultados da comparação dos modelos.

5.1 ANÁLISE DOS MODELOS PROPOSTOS

Neste trabalho foram desenvolvidos dois novos modelos autorregressivos não condicionais, com distribuição marginal Lindley e Gama-Lindley. Entretanto, por simplicidade, para a construção desses novos modelos AR(2) atribuiu-se uma restrição aos parâmetros autorregressivos de modo que eles fossem iguais.

Desse modo, para o *Caso 1* a distribuição de inovação correspondente aos dois modelos é uma mistura envolvendo 2 distribuições exponenciais e 1 Gama. Para os dois modelos, existe uma semelhança entre os parâmetros da distribuição exponencial e gama, que são λ e $(2, \lambda)$, respectivamente.

Entretanto, o que os difere são os parâmetros da exponencial referente a cada modelo. Para o modelo com distribuição marginal Lindley, a exponencial possui parâmetro $\left(\frac{\lambda+1}{a}\right)$. O parâmetro da distribuição exponencial usando o Gama-Lindley é $\left(\frac{\beta\lambda+\beta}{a}\right)$. Note que, assim como ocorre com as funções de densidade de probabilidade, tomando $\beta = 1$, a distribuição de inovação da Gama-Lindley recai para a Lindley. Dessa forma, para existirem dois modelos diferentes é necessário que $\beta \neq 1$.

Além disso, para o modelo LAR₁(2) construiu-se o espaço paramétrico da região que

compõe o parâmetro λ . Desse modo, a região foi determinada utilizando a definição formal de limite e também expansão de séries de Fourier, gerando assim duas novas proposições. Para o modelo $GLAR_1(2)$ existe uma relação entre os parâmetros β e λ de acordo com a Proposição 4.2.2.

Da mesma maneira, considerando o *Caso 2*, que corresponde ao caso geral, a distribuição de inovação da v.a ϵ_t é também uma mistura envolvendo distribuições exponenciais e Gama. De forma similar, ocorre a mesma semelhança entre dois parâmetros das distribuições: exponencial (λ) e Gama ($2, \lambda$).

Assim, de modo análogo, a diferença no caso geral para os modelos $LAR(2)$ e $GLAR(2)$ refere-se aos parâmetros da distribuição exponencial de cada modelo, que correspondem as raízes reais de um polinômio de grau 3. Para o primeiro modelo, o polinômio está em função dos termos autorregressivos e do termo adicional λ . No segundo modelo, o polinômio é função dos parâmetros autorregressivos α_1, α_2 e dois termos adicionais β e λ . O ponto semelhante entre eles está nos parâmetros da distribuição exponencial que correspondem ao oposto das raízes reais negativas de cada um desses polinômios.

Em geral, a distribuição de inovação para os dois modelos propostos apresentam formatos semelhantes no que diz respeito à quantidade de distribuições Gama e exponenciais. No *Caso 1* para os dois modelos, a mistura é formada por 2 exponenciais e uma Gama. Para o *Caso 2*, tem-se a mistura composta por 4 exponenciais e uma Gama. Além disso, nota-se a influência dos parâmetros autorregressivos nos resultados das estimativas dos parâmetros. No *Caso 1*, as estimativas dos parâmetros tanto do $LAR_1(2)$ quanto o $GLAR_1(2)$ convergem para os valores verdadeiros à medida que aumenta-se o parâmetro autorregressivo. No caso geral, os parâmetros autorregressivos também influenciam nos resultados das estimativas dos parâmetros λ e β .

Para avaliar o desempenho dos modelos propostos utilizou-se os critérios de seleção AIC, BIC e RMSE que serão expostos a seguir.

5.1.1 Critérios de Seleção

1) AKAIKE (AIC) Trata-se de um critério bastante utilizado na seleção de modelos, que foi desenvolvido por Akaike (1973). Sua estrutura obedece a seguinte expressão:

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

em que $\hat{C}$ representa o vetor de parâmetros que maximiza a função de log-verossimilhança. O termo posterior adicionado, funciona como uma penalização do modelo e está diretamente relacionado ao número de parâmetros p .

2) BAYESIANO (BIC) Também muito utilizado para comparação de modelos, esse critério foi proposto por Schwarz (1978), e descrito como:

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

com $\hat{C}$ o modelo selecionado, p o número de parâmetros e n o número de observações para o conjunto de dados.

Nota-se que tanto o AIC quanto o BIC se baseiam na função de log-verossimilhança diferindo, no entanto, nas diferentes penalizações.

3) RAIZ DO ERRO QUADRÁTICO MÉDIO (RMSE) Esse critério consiste em verificar o quanto os valores preditos se aproximam dos valores verdadeiros. O RMSE pode ser calculado como:

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (X_k - \hat{X}_k)^2},$$

com X_k , $\hat{X}_k$ representando os valores observados e valores preditos no instante k , respectivamente. Dessa forma, quanto mais próximo o seu valor se aproximar de 0, menor será o valor do erro, mostrando assim que os valores preditos são próximos dos valores ajustados. Logo, o modelo apresenta um bom desempenho, de acordo com o conjunto de dados trabalhado.

Na sequência são apresentados os modelos desenvolvidos na literatura que serão utilizados para comparar com os dois novos propostos.

5.1.2 Processo EAR(2)

Trata-se de uma extensão do modelo AR(2) com distribuição marginal exponencial com 3 parâmetros em sua estrutura: dois parâmetros, α_1 e α_2 , que representam os termos autorregressivos e um parâmetro λ , que compõe a distribuição marginal. Assim, o modelo EAR(2) é uma generalização do modelo EAR(1), proposto por Gaver e Lewis (1980). O processo EAR(2), por definição é dado por:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t,$$

em que ϵ_t é uma sequência de variáveis independente e identicamente distribuída (iid) e α_1 e α_2 são os parâmetros do modelo AR(2). A função log-verossimilhança dos processos seguem

a seguinte estrutura:

$$l(\lambda, \alpha_1, \alpha_2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \sum_{t=3}^N \left[\frac{(y_t - \mu_{x_t})^2}{2\sigma^2} \right], \quad (5.1)$$

em que μ_{x_t} representa a média e $\sigma_{x_t}^2$ a variância do EAR(2), formadas pela junção dos modelos exponencial e AR(2). Desse modo, temos:

$$\mu_{x_t} = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \frac{1}{\lambda}(1 - \alpha_1 - \alpha_2)$$

e

$$\sigma_{x_t}^2 = \frac{(1 - \alpha_2)}{\lambda^2(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)}.$$

Substituindo na Equação (5.1), temos que a função log-verossimilhança gaussiana do processo EAR(2) é da forma:

$$\begin{aligned} l(\lambda, \alpha_1, \alpha_2) = & -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \left(\frac{(1 - \alpha_2)}{\lambda^2(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)} \right) \\ & - \frac{1}{2} \sum_{t=3}^n \left[\frac{\lambda^2(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)}{(1 - \alpha_2)} \right] \\ & \times \left[y_t - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - \frac{1}{\lambda}(1 - \alpha_1 - \alpha_2) \right]^2. \end{aligned}$$

5.1.3 Processo GAR(2)

Proposto por Gaver e Lewis (1980), o processo Gama autorregressivo de segunda ordem, assim como o modelo GLAR(2), também possui na sua composição dois parâmetros k_1 e λ , que correspondem aos parâmetros de forma e escala, respectivamente, além dos parâmetros autorregressivos: α_1 e α_2 . O valor esperado e a variância da distribuição gama são:

$$\mathbb{E}(X) = \frac{k_1}{\lambda} \quad \text{e} \quad \sigma_{x_t}^2 = \frac{k_1}{\lambda^2}.$$

Assim, a média e a variância do processo GAR(2) seguem, respectivamente, de acordo com as expressões:

$$\mu_{x_t} = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \frac{k}{\lambda}(1 - \alpha_1 - \alpha_2) \quad (5.2)$$

e

$$\sigma_{x_t}^2 = \frac{(1 - \alpha_2)k}{\lambda^2(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)}. \quad (5.3)$$

De modo análogo, como construído o processo anterior, substituindo as Equações (5.2), (5.3) em (5.1), a função de log-verossimilhança gaussiana é dada por:

$$l(k, \lambda, \alpha_1, \alpha_2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \left[\frac{(1 - \alpha_2)k}{\lambda^2(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)} \right] \\ - \frac{1}{2} \sum_{t=3}^n \left[\frac{\lambda^2(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)}{(1 - \alpha_2)k} \right] \\ \times \left[y_t - \alpha_1 y_{t-1} - \alpha_2 y_{t-2} - \frac{k}{\lambda}(1 - \alpha_1 - \alpha_2) \right]^2.$$

5.1.4 INGAR(2)

O processo AR(2) com distribuição marginal inversa gaussiana também possui quatro parâmetros em sua estrutura: 2 parâmetros μ e λ , que compõem a distribuição gaussiana inversa e os parâmetros α_1 e α_2 do modelo AR(2). Como uma extensão do processo autorregressivo de primeira ordem com distribuição marginal gaussiana proposta em Abraham e Balakrishna (1999). A fdp da distribuição inversa gaussiana é da forma:

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

em que x , μ e λ são positivos.

O valor esperado e a variância da inversa gaussiana são, respectivamente:

$$\mathbb{E}(x) = \mu \quad \text{e} \quad \text{Var}(X) = \frac{\mu^3}{\lambda}.$$

Assim, a log-verossimilhança gaussiana do processo INGAR(2) é da forma:

$$l(\lambda, \alpha_1, \alpha_2) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \left[\frac{\mu^3(1 - \alpha_2)}{\lambda(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)} \right] \\ - \frac{1}{2} \sum_{t=3}^N \left[\frac{\lambda(1 + \alpha_2)(1 - \alpha_1 - \alpha_2)(1 + \alpha_1 - \alpha_2)}{(1 - \alpha_2)\mu^3} \right] \\ \times [y_t - \alpha_1 y_{t-1} - \alpha_2 y_{t-2} - \mu(1 - \alpha_1 - \alpha_2)]^2.$$

Em relação ao processo EAR(2), o processo acima é bastante similar, mas com um parâmetro a mais na sua estrutura (μ).

5.2 RESULTADOS

Os novos modelos LAR(2) e GLAR(2) foram comparados de acordo com os dois casos em que foram desenvolvidos: Caso 1 que corresponde aos parâmetros autorregressivos iguais

e o *Caso 2* em que os parâmetros autorregressivos são diferentes. Quanto a análise sobre a eficiência dos modelos, o EAR(2) possui um parâmetro a menos em relação ao GLAR₂(2), o que pode influenciar positivamente nos resultados, uma vez que os critérios citados anteriormente penalizam quanto à quantidade de parâmetros adicionais ao modelo. No entanto, o GLAR₂(2) apresentou melhores resultados, conforme será analisado posteriormente.

Como foi utilizado duas bases de dados reais diferentes para analisar a capacidade de ajuste dos modelos a esses conjuntos de dados, tem-se que a Tabela 5 mostra os resultados da comparação dos modelos LAR₁(2) e GLAR₁(2) entre si, e com outros modelos desenvolvidos na literatura. Assim, as Tabelas 6 e 8 apresentam os resultados da comparação dos modelos para o caso geral, referente à base de dados sobre a cotação da bolsa de valores e a base de dados climáticos da capital da Índia, respectivamente. Da mesma maneira, as Tabela 7 e 8 apresentam os resultados da comparação dos modelos usando à base de dados climáticos vistos na seção de Aplicação, na construção de cada modelo.

Tabela 5 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados do mercado de ações - *Caso 1*

Modelo	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{k}$	$\hat{\mu}$	$\hat{\theta}$	AIC	BIC	RMSE
GLAR ₁ (2)	0.2750	0.2750	0.5	1	–	–	–	1185.89	1195.63	0.3119
LAR ₁ (2)	0.2750	0.2750	–	20	–	–	–	1212.02	1218.51	0.1906
INGAR(2)	0.3050	0.3550	–	20	–	1.75	–	1213.06	1225.08	0.1799
GAR(2)	0.4832	0.4661	–	20	0.95	–	–	1272.05	1285.03	0.1869
EAR(2)	0.4800	0.5000	–	20	–	–	–	1962.03	1973.74	0.1875

Fonte: Elaboração Própria (2025)

Tabela 6 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados do mercado de ações - *Caso 2*

Modelo	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{k}$	$\hat{\mu}$	$\hat{\theta}$	AIC	BIC	RMSE
GLAR ₂ (2)	0.2750	0.6250	0.5	5	–	–	–	1202.02	1211.76	0.1407
LAR ₂ (2)	0.2750	0.6250	–	100	–	–	–	1210.88	1217.37	0.1476
EAR(2)	0.4500	0.5000	–	20	–	–	–	1272.54	1282.28	0.1490
INGAR(2)	0.2750	0.2550	–	30	–	–	1.75	1273.66	1287.58	0.1498
GAR(2)	0.2750	0.6025	–	20	0.95	–	–	1303.08	1312.97	0.1512

Fonte: Elaboração Própria (2025)

Tabela 7 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados em hidrologia- *Caso 1*

Modelo	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{k}$	$\hat{\mu}$	$\hat{\theta}$	AIC	BIC	RMSE
GLAR ₁ (2)	0.4661	0.4661	0.5	0.1	–	–	–	980.03	988.24	0.5087
LAR ₁ (2)	0.4661	0.4661	–	20	–	–	–	985.26	989.73	0.5655
INGAR(2)	0.3050	0.3550	–	20	–	0.20	–	1014.62	1022.83	0.5594
EAR(2)	0.4800	0.5000	–	20	–	–	–	1028.86	1034.33	0.5710
GAR(2)	0.4832	0.4661	–	20	0.10	–	–	1033.27	1041.48	1.3824

Fonte: Elaboração Própria (2025)

Tabela 8 – Parâmetros estimados e valores para AIC, BIC e RMSE a partir dos dados em hidrologia- *Caso 2*

Modelo	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{k}$	$\hat{\mu}$	$\hat{\theta}$	AIC	BIC	RMSE
GLAR ₂ (2)	0.4832	0.4661	0.85	5	–	–	–	786.54	797.49	0.1751
LAR ₂ (2)	0.4832	0.4661	–	100	–	–	–	790.86	800.60	0.1847
EAR(2)	0.7417	0.1583	–	100	–	–	–	818.80	828.54	0.1795
INGAR(2)	0.7192	0.1808	–	20	–	0.20	–	819.43	830.37	0.1742
GAR(2)	0.6900	0.2100	–	95	0.5	–	–	820.59	833.58	0.1953

Fonte: Elaboração Própria (2025)

De acordo com as Tabelas acima, em relação aos dois novos modelos propostos, o GLAR(2), mesmo com um número maior de parâmetros em sua estrutura, obteve melhor desempenho em relação ao LAR(2), tanto para o *Caso 1* quanto para o *Caso 2*. Além disso, para todos os casos observados, de acordo com as bases de dados reais analisadas, o GLAR(2) mostrou ser eficiente em relação a todos outros modelos testados. O modelo LAR(2) foi superior aos modelos EAR(2), GAR(2) e INGAR(2). Sendo assim, de um modo geral, em relação aos critérios observados, tanto o modelo LAR(2), quanto o GLAR(2), apresentaram bons resultados, mostrando ser bons competidores em relação aos existentes na literatura.

Assim, os modelos LAR(2) e GLAR(2) podem consolidar-se como uma ferramentas relevantes na área do mercado financeiro, no âmbito dos ativos de moedas fiduciárias. Uma vez que as suas estruturas permitem capturar padrões complexos de volatilidade de mercado, com os bitcoins, por exemplo. Isso pode auxiliar investidores na tomada de decisões estratégicas, contribuindo para um entendimento profundo dos mecanismos de volatilidade em mercados financeiros.

Nessa perspectiva, os modelos propostos também podem ser úteis na área de Hidrologia, ao captar séries climáticas, pode-se fazer previsão sobre eventos climáticos diversos eventos climáticos.

6 CONCLUSÃO E TRABALHOS FUTUROS

Nesta Tese dois novos modelos autorregressivos $AR(2)$ não condicionais com distribuições marginais Lindley e Gama-Lindley foram desenvolvidos. Os resultados demonstram o potencial desses modelos para o ajuste em ST financeiras e climáticas. Como perspectiva futura, propõe-se essa abordagem para modelos de ordem p , o que permitiria maior flexibilidade na captura de dependências temporais. Entretanto, essa generalização enfrentaria desafios significativos como:

1. Aumento exponencial da complexidade computacional na estimação dos parâmetros.
2. Os modelos propostos $AR(2)$ geram uma mistura de gamas e exponenciais, a partir daí, analisar comportamento da estrutura para modelos $AR(3)$ e $AR(4)$, por exemplo, onde às dificuldades matemáticas tornam-se complexas, à medida que aumenta-se a ordem do modelo.
3. Dado que um processo $AR(2)$ possui uma equação característica, que representa um polinômio de grau 3, e cujo oposto das raízes são parâmetros da distribuição exponencial que corresponde a um dos componentes da distribuição de inovação, o desafio consiste em como generalizar processos $AR(p)$, de acordo com esse polinômio característico.
4. Na generalização de modelos $AR(p)$, a análise do polinômio característico requer estudo aprofundado sobre as condições de estacionariedade, bem como as regiões que garantam raízes reais e preservem as propriedades do modelo.

REFERÊNCIAS

- ABRAHAM, B.; BALAKRISHNA, N. Inverse gaussian autoregressive models. *Journal of time series analysis*, Wiley Online Library, v. 20, n. 6, p. 605–618, 1999.
- AKAIKE, H. Maximum likelihood identification of gaussian autoregressive moving average models. *Biometrika*, Oxford University Press, v. 60, n. 2, p. 255–265, 1973.
- AL-MUTAIRI, D.; GHITANY, M.; KUNDU, D. Inferences on stress-strength reliability from lindley distributions. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 42, n. 8, p. 1443–1463, 2013.
- ALIZADEH, S.; BRANDT, M. W.; DIEBOLD, F. X. Range-based estimation of stochastic volatility models. *The Journal of Finance*, Wiley Online Library, v. 57, n. 3, p. 1047–1091, 2002.
- ANTUNES, J. L. F.; CARDOSO, M. R. A. Uso da análise de séries temporais em estudos epidemiológicos. *Epidemiologia e Serviços de Saúde*, SciELO Brasil, v. 24, p. 565–576, 2015.
- ARFKEN, G. B.; WEBER, H. J.; HARRIS, F. E. *Mathematical methods for physicists: a comprehensive guide*. [S.l.]: Academic press, 2011.
- ARSLAN, T.; ACITAS, S.; SENOGLU, B. Generalized lindley and power lindley distributions for modeling the wind speed data. *Energy Conversion and Management*, Elsevier, v. 152, p. 300–311, 2017.
- BAKOUCH, H. S.; AL-ZAHRANI, B. M.; AL-SHOMRANI, A. A.; MARCHI, V. A.; LOUZADA, F. An extended lindley distribution. *Journal of the Korean Statistical Society*, Springer, v. 41, p. 75–85, 2012.
- BAKOUCH, H. S.; POPOVIĆ, B. V. Lindley first-order autoregressive model with applications. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 45, n. 17, p. 4988–5006, 2016.
- BALBINOT, E.; SCOTTON, J. W.; CEREZER, S. M.; MARTINAZZO, C. A. Modelos de séries temporais aplicados a previsão de radiação solar. *Revista Perspectiva*, 2017.
- BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. *Time series analysis: forecasting and control*. [S.l.]: John Wiley & Sons, 2015.
- BROCKWELL, P. J.; DAVIS, R. A. *Time series: theory and methods*. [S.l.]: Springer science & business media, 1991.
- CALLEN, J. L.; MOREL, M. Linear accounting valuation when abnormal earnings are ar (2). *Review of Quantitative Finance and Accounting*, Springer, v. 16, p. 191–204, 2001.
- CHIANG, M.-H.; WANG, L.-M. Volatility contagion: A range-based volatility approach. *Journal of Econometrics*, Elsevier, v. 165, n. 2, p. 175–189, 2011.
- CORDEIRO, G. M.; PRATAVIERA, F.; LIMA, M. do C. S.; ORTEGA, E. M. The marshall-olkin extended flexible weibull regression model for censored lifetime data. *Model Assisted Statistics and Applications*, SAGE Publications Sage UK: London, England, v. 14, n. 1, p. 1–17, 2019.

DEY, S.; NASSAR, M.; KUMAR, D. Alpha power transformed inverse lindley distribution: A distribution with an upside-down bathtub-shaped hazard function. *Journal of Computational and Applied Mathematics*, Elsevier, v. 348, p. 130–145, 2019.

GAVER, D. P.; LEWIS, P. A. First-order autoregressive gamma sequences and point processes. *Advances in Applied Probability*, Cambridge University Press, v. 12, n. 3, p. 727–745, 1980.

GHITANY, M. E.; AL-MUTAIRI, D. K.; BALAKRISHNAN, N.; AL-ENEZI, L. Power lindley distribution and associated inference. *Computational Statistics & Data Analysis*, Elsevier, v. 64, p. 20–33, 2013.

GHITANY, M. E.; ATIEH, B.; NADARAJAH, S. Lindley distribution and its application. *Mathematics and computers in simulation*, Elsevier, v. 78, n. 4, p. 493–506, 2008.

HAMILTON, J. D. *Time series analysis*. [S.l.]: Princeton university press, 2020.

HUGUENY, B. Spatial synchrony in population fluctuations: extending the moran theorem to cope with spatially heterogeneous dynamics. *Oikos*, Wiley Online Library, v. 115, n. 1, p. 3–14, 2006.

JOSHI, S.; JOSE, K. Wrapped lindley distribution. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 47, n. 5, p. 1013–1021, 2018.

KLIMKO, L. A.; NELSON, P. I. On conditional least squares estimation for stochastic processes. *The Annals of statistics*, JSTOR, p. 629–642, 1978.

KRISHNA, H.; KUMAR, K. Reliability estimation in lindley distribution with progressively type ii right censored sample. *Mathematics and Computers in Simulation*, Elsevier, v. 82, n. 2, p. 281–294, 2011.

LAWRANCE, A.; LEWIS, P. The exponential autoregressive-moving average earma (p, q) process. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 42, n. 2, p. 150–161, 1980.

LIMA, E. L. *Curso de análise*. [S.l.]: Instituto de Matemática Pura e Aplicada, CNPq, 1981. v. 2.

LIMA, I. G. d. et al. *Processo Gompertz-ARMA e propriedades: uma aplicação à precificação do mercado financeiro*. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2021.

LIMA, M. C.; CORDEIRO, G. M.; ORTEGA, E. M.; NASCIMENTO, A. D. A new extended normal regression model: simulations and applications. *Journal of Statistical Distributions and Applications*, Springer, v. 6, p. 1–17, 2019.

LIMA, M. d. C. S. Mathematical properties of some generalized gamma models. Universidade Federal de Pernambuco, 2015.

LINDLEY, D. V. Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 102–107, 1958.

MAZUCHELI, J.; ACHCAR, J. A. The lindley distribution applied to competing risks lifetime data. *Computer methods and programs in biomedicine*, Elsevier, v. 104, n. 2, p. 188–192, 2011.

- MAZUCHELI, J.; BAPAT, S. R.; MENEZES, A. F. B. A new one-parameter unit-lindley distribution. *Chilean Journal of Statistics (ChJS)*, v. 11, n. 1, 2020.
- MELLO, A. B. V. d. et al. *A construção de um novo processo autorregressivo de primeira ordem com distribuição marginal Gama-Lindley*. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2021.
- METWALLY, I. Simulation of the impulse response of electrical machines. *IEEE transactions on energy conversion*, IEEE, v. 14, n. 4, p. 861–867, 1999.
- MILANI, E. A. Modelos para séries temporais utilizando as distribuições normal generalizada e log-normal generalizada. Universidade Federal de São Carlos, 2016.
- NADARAJAH, S.; BAKOUCH, H. S.; TAHMASBI, R. A generalized lindley distribution. *Sankhya B*, Springer, v. 73, p. 331–359, 2011.
- PEIXOTO, J. A.; RENNA, R. B. D.; CORRÊA, C. R. B. Tutorial sobre expansao de funções algébricas racionais em frações parciais aplicada a sistemas analógicos e digitais. 2013.
- PERCONTINI, A.; GOMES-SILVA, F. S.; RAMOS, M.; VENANCIO, R.; CORDEIRO, G. M. A distribuição gama weibull poisson aplicada a dados de sobrevivência. *TEMA (São Carlos)*, SciELO Brasil, v. 15, p. 165–176, 2014.
- RUSSO, S.; RODRIGUES, P. M.; CAMARGO, M. E. Aplicação de séries temporais na série teor de umidade da areia de fundição da indústria fundimisa. *Revista Gestão Industrial*, v. 2, n. 1, 2006.
- SANKARAN, M. 275. note: The discrete poisson-lindley distribution. *Biometrics*, JSTOR, p. 145–149, 1970.
- SANTOS, T. Felipe de M. S. Valor em risco auto-regressivo condicional: o caso de índices brasileiros. 2006.
- SCHWARZ, G. Estimating the dimension of a model. *The annals of statistics*, JSTOR, p. 461–464, 1978.
- SHANKER, R.; MISHRA, A. A quasi lindley distribution. *African Journal of Mathematics and Computer Science Research*, v. 6, n. 4, p. 64–71, 2013.
- SHARMA, V. K.; SINGH, S. K.; SINGH, U.; AGIWAL, V. The inverse lindley distribution: a stress-strength reliability model with application to head and neck cancer data. *Journal of Industrial and Production Engineering*, Taylor & Francis, v. 32, n. 3, p. 162–173, 2015.
- SHUMWAY, R. H.; STOFFER, D. S.; STOFFER, D. S. *Time series analysis and its applications*. [S.l.]: Springer, 2000. v. 3.
- SOUZA, J. A. N. de. Trajetória recente da dívida pública brasileira: Uma previsão a partir de um modelo de séries temporais. *Revista de Economia Regional, Urbana e do Trabalho*, v. 11, n. 1, p. 130–150, 2022.
- ZEGHDOUDI, H.; NEDJAR, S. Gamma lindley distribution and its application. *Journal of Applied probability and statistics*, v. 10, 2015.

ZOGRAFOS, K.; BALAKRISHNAN, N. On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical methodology*, Elsevier, v. 6, n. 4, p. 344–362, 2009.

APÊNDICE A –

Cálculos da Proposição 3.1.1 (*Caso 1*).

A transformada de Laplace da distribuição Lindley é da forma:

$$\phi_X(s) = \frac{\lambda^2(\lambda + s + 1)}{(1 + \lambda)(s + \lambda)^2}.$$

Para o modelo AR(2) segue:

$$\phi_X(s) = [(1 - \alpha_2)\phi_{X_{i-1}}(\alpha_1 s) + \alpha_2\phi_{X_{i-2}}(\alpha_2 s)]\phi_\epsilon(s).$$

Admitindo a estacionariedade do processo, podemos reescrever a expressão acima:

$$\phi_X(s) = [(1 - \alpha_2)\phi_X(\alpha_1 s) + \alpha_2\phi_X(\alpha_2 s)]\phi_\epsilon(s). \quad (\text{A.1})$$

Da equação (A.1), tem-se:

$$\phi_\epsilon(s) = \frac{\phi_X(s)}{[(1 - \alpha_2)\phi_X(\alpha_1 s) + \alpha_2\phi_X(\alpha_2 s)]}. \quad (\text{A.2})$$

Agora, calculando $\phi_X(\alpha_1 s)$ e $\phi_X(\alpha_2 s)$:

$$\phi_X(\alpha_1 s) = \frac{\lambda^2(\lambda + \alpha_1 s + 1)}{(\lambda + 1)(\lambda + \alpha_1 s)^2} \quad \text{e} \quad \phi_X(\alpha_2 s) = \frac{\lambda^2(\lambda + \alpha_2 s + 1)}{(\lambda + 1)(\lambda + \alpha_2 s)^2}. \quad (\text{A.3})$$

Substituindo (A.3) em (A.2) obtém-se:

$$\phi_\epsilon(s) = \frac{\frac{\lambda^2(\lambda+s+1)}{(\lambda+1)(\lambda+s)^2}}{\frac{(1-\alpha_2)\lambda^2(\lambda+\alpha_1 s+1)}{(\lambda+1)(\lambda+\alpha_1 s)^2} + \frac{\alpha_2\lambda^2(\lambda+\alpha_2 s+1)}{(\lambda+1)(\lambda+\alpha_2 s)^2}}. \quad (\text{A.4})$$

Reescrevendo a partir da Equação (A.4):

$$\phi_\epsilon(s) = \frac{\lambda^2(\lambda + s + 1)(\lambda + 1)(\lambda + \alpha_1 s)^2(\lambda + \alpha_2 s)^2}{(\lambda + 1)(\lambda + s)^2\lambda^2[(1 - \alpha_2)(\lambda + \alpha_1 s + 1)(\lambda + \alpha_2 s)^2 + \alpha_2(\lambda + \alpha_2 s + 1)(\lambda + \alpha_1 s)^2]}.$$

Após simplificação da expressão acima:

$$\phi_\epsilon(s) = \frac{(\lambda + s + 1)(\lambda + \alpha_1 s)^2(\lambda + \alpha_2 s)^2}{(\lambda + s)^2[(1 - \alpha_2)(\lambda + \alpha_1 s + 1)(\lambda + \alpha_2 s)^2 + \alpha_2(\lambda + \alpha_2 s + 1)(\lambda + \alpha_1 s)^2]}.$$

Para esse modelo vamos fazer uma restrição nos parâmetros autorregressivos de modo que se tenha:

$$\alpha_1 = \alpha_2 = a, \quad a \in (0, 1).$$

Podemos reescrever $\phi_\epsilon(s)$ da forma:

$$\phi_\epsilon(s) = \frac{(\lambda + s + 1)(\lambda + as)^2}{(s + \lambda)^2(\lambda + as + 1)}.$$

Com o intuito de obter a distribuição do erro, aplica-se em $\phi_\epsilon(s)$ a transformada inversa de Laplace, usando o método de frações parciais. Técnica essa, utilizada em Cálculo Diferencial que facilita a integração de funções racionais. Desse modo, considere $H(s) = \phi_\epsilon(s)$. Assim, a transformada inversa de Laplace correspondente será denotada por $h(x)$. Segue:

$$f_\epsilon(x) = h(x).$$

Dessa forma, para encontrar $h(x)$ é necessário decompor a expressão $H(s)$ em frações parciais.

$$\frac{(\lambda + s + 1)(\lambda + as)^2}{(\lambda + s)^2(\lambda + as + 1)} = \frac{A^*}{(\lambda + s)} + \frac{B^*}{(\lambda + s)^2} + \frac{C^*}{(\lambda + as + 1)}.$$

Fazendo $s = -\lambda$, temos:

$$B^* = \frac{-\lambda^2(a - 1)^2}{(\lambda - a\lambda + 1)}.$$

Considere $s = \frac{-(1+\lambda)}{a}$, então:

$$C^* = \frac{a(a - 1)(1 + \lambda)}{(\lambda - a\lambda + 1)^2}.$$

Por fim, avaliando em $s = 0$ e substituindo os valores dos coeficientes B^* e C^* encontrados:

$$A^* = -\frac{\frac{(\lambda+1)(\lambda-a\lambda)^2}{(\lambda-a\lambda+1)} - \lambda^2(\lambda+1) + \frac{a\lambda^2(a-1)(\lambda+1)}{(\lambda-a\lambda+1)^2}}{\lambda(\lambda+1)}.$$

Por conseguinte, nota-se que $H(s)$ pode ser reescrita do seguinte forma:

$$H(s) = \frac{A^*}{(s + \lambda)} + \frac{B^*}{(s + \lambda)^2} + \frac{C^*}{(\lambda - a\lambda + 1)}.$$

Uma vez que temos funções conhecidas, agora podemos aplicar a transformada inversa de Laplace.

Sabe-se que $h(x)$ é a transformada inversa de $H(s)$, logo:

$$h(x) = T_1^* \lambda e^{-\lambda x} + T_2^* x \lambda^2 e^{-\lambda x} + T_3^* \left(\frac{\lambda + 1}{a} \right) e^{-\left(\frac{1+\lambda}{a}\right)x},$$

com os pesos finais T_1^* , T_2^* e T_3^* da forma:

$$T_1^* = \left(\frac{A^*}{\lambda} \right) = \frac{a^3 \lambda + a^2 \lambda^2 - 3a^2 \lambda - 2a^2 - 2a \lambda^2 + a \lambda + 3a + \lambda^2 + \lambda}{(\lambda - a\lambda + 1)^2},$$

$$T_2^* = \left(\frac{B^*}{\lambda^2} \right) = \frac{(a - 1)^2}{(\lambda - a\lambda + 1)} \quad \text{e} \quad T_3^* = - \left(\frac{C^*}{1 + \lambda} \right) a = \frac{-a^2(a - 1)}{(\lambda - a\lambda + 1)^2}.$$

Expressando a soma em função dos parâmetros a e λ , tem-se que:

$$S(a, \lambda) = \frac{-a^3 + a^2\lambda^2 - 2a\lambda^2 - 2a\lambda + a + \lambda^2 + 2\lambda + 1}{(\lambda - a\lambda + 1)^2}.$$

Portanto, considerando a Proposição 3.1.1, para o modelo $\text{LAR}_1(2)$, $h(x)$ é uma mistura de distribuições gama $(2, \lambda)$ e exponenciais de parâmetros (λ) e $\left(\frac{1+\lambda}{a}\right)$, com a soma dos pesos:

$$\boxed{T_1^* + T_2^* + T_3^* = 1}.$$

APÊNDICE B –

Cálculos para o espaço paramétrico do parâmetro λ referente à Proposição 3.1.1.

Analisando o espaço paramétrico para o $\text{LAR}_1(2)$, temos que a soma (S) dos pesos é da forma:

$$S(a, \lambda) = \frac{a^2\lambda^2 - 2a^2 - 2a\lambda^2 - 2a\lambda + 2a + \lambda^2 + 2\lambda + 1}{(\lambda - a\lambda + 1)^2}. \quad (\text{B.1})$$

De acordo com a região de estabilidade para o novo modelo, o termo autorregressivo está no intervalo: $[0.1; 0.49]$.

Analisando a soma S da Equação (B.1) para $a = 0.1$, tem-se:

$$S(\lambda) = \frac{81\lambda^2 + 180\lambda + 118}{81\lambda^2 + 180\lambda + 100}.$$

Reescrevendo $S(\lambda)$ obtém-se:

$$S(\lambda) = 1 + \frac{18}{81\lambda^2 + 180\lambda + 100}.$$

Agora, é necessário investigar o comportamento de λ de modo que a expressão

$$S_1(\lambda) = \frac{18}{81\lambda^2 + 180\lambda + 100},$$

se torne próxima de 0. Usando a definição de limite, veja Lima (1981), tem-se que:

$$\forall \epsilon > 0, \quad \exists \lambda_0 > 0 \mid \lambda > \lambda_0 \quad \Rightarrow \mid S_1(\lambda) - 0 \mid < \epsilon.$$

A ideia é encontrar valores para λ para o qual $S_1(\lambda)$ seja menor que um valor arbitrário pequeno ϵ . Dessa forma, tem-se:

$$\begin{aligned} S_1(\lambda) &< \epsilon \\ \Leftrightarrow \frac{18}{81\lambda^2 + 180\lambda + 100} &< \epsilon \\ \Leftrightarrow 18 &< (\epsilon)(81\lambda^2 + 180\lambda + 100) \\ \Leftrightarrow 18 &< 81\epsilon\lambda^2 + 180\epsilon\lambda + 100\epsilon \\ \Leftrightarrow 81\epsilon\lambda^2 + 180\epsilon\lambda + 100\epsilon - 18 &> 0. \end{aligned}$$

Resolvendo a inequação quadrática em λ

$$81\epsilon\lambda^2 + 180\epsilon\lambda + 100\epsilon - 18 > 0,$$

as raízes em função de ϵ são da seguinte forma:

$$\lambda_1(\epsilon) = \frac{-10}{9} + \frac{\sqrt{2\epsilon}}{3\epsilon} \quad \text{e} \quad \lambda_2(\epsilon) = \frac{-10}{9} - \frac{\sqrt{2\epsilon}}{3\epsilon}.$$

Como o parâmetro λ é positivo, temos que a solução possível da inequação neste caso, será

$$S_1 = \{\lambda \in \mathcal{R} : \lambda > \lambda_1(\epsilon)\}.$$

Considerando agora $a = 0.49$, substituindo na Equação (B.1):

$$S_2(\lambda) = \frac{292846565769766502400\lambda^2 + 1148417904979476468399\lambda + 1688624680282567475200}{112589990684262400(2601\lambda^2 + 10200\lambda + 10000)}.$$

De modo análogo ao processo anterior:

$$\forall \epsilon > 0, \exists \lambda_0 > 0 | \lambda > \lambda_0, \Rightarrow |S_2(\lambda) - 1| < \epsilon.$$

A expressão $S_2(\lambda)$ pode ser decomposta da forma:

$$S_2(\lambda) = 1 + R,$$

com R indicando o resíduo, de modo que:

$$R = \frac{4399\lambda + 562724773439943575200}{112589990684262400(2601\lambda^2 + 10200\lambda + 10000)}.$$

Dessa forma, tem-se:

$$R < \epsilon$$

$$\Leftrightarrow \frac{4399\lambda + 562724773439943575200}{112589990684262400(2601\lambda^2 + 10200\lambda + 10000)} < \epsilon$$

$$\Leftrightarrow 4399\lambda + 562724773439943575200 < 112589990684262400(2601\lambda^2 + 10200\lambda + 10000)\epsilon.$$

Resolvendo a inequação quadrática, e utilizando o software matemático Matlab, tem-se que uma expressão a única solução possível para λ , desde que $\lambda > 0$ é:

$$\lambda_2(\epsilon) = \frac{\sqrt{(A\epsilon - B)^2 - C\epsilon(D\epsilon - E)} - A\epsilon + B}{F\epsilon},$$

com os coeficientes:

$$\begin{aligned} A &= 1.14841790497947648 \times 10^{21}, & B &= 4399, \\ C &= 1.171386263079066 \times 10^{21}, & D &= 1.125899906842624 \times 10^{21}, \\ E &= 5.627247734399435 \times 10^{20}, & F &= 5.85693131539533 \times 10^{20}. \end{aligned}$$

Logo, a solução considerando essa condição é da forma:

$$S_2 = \{\lambda \in \mathcal{R} : \lambda > \lambda_2(\epsilon)\}.$$

Agora, a solução consiste em encontrar a intersecção para os dois resultados desenvolvidos.

$$\lambda(\epsilon) = \lambda_1(\epsilon) \cap \lambda_2(\epsilon) = \lambda_2(\epsilon).$$

Dessa forma, pode-se obter o valor λ de acordo com o ϵ escolhido, conforme mostra a tabela a seguir.

Tabela 9 – Valores de λ para diferentes ϵ

ϵ	λ
0.1	2.42
0.01	11.90
0.001	41.87
0.0001	136.65
0.00001	436.39

Fonte: Elaboração Própria (2025)

Desse modo, para a construção do modelo $\text{LAR}_1(2)$, considerou-se $\epsilon = 0.01$.

APÊNDICE C –

Cálculos da Proposição 3.2.1.

A transformada de Laplace para um modelo AR(2) é da forma:

$$\phi_X(s) = [(1 - \alpha_2)\phi_{X_{i-1}}(\alpha_1 s) + \alpha_2\phi_{X_{i-2}}(\alpha_2 s)]\phi_\epsilon(s).$$

Para o caso geral, a transformada para o erro é:

$$\phi_\epsilon(s) = \frac{(\lambda + s + 1)(\lambda + \alpha_1 s)^2(\lambda + \alpha_2 s)^2}{(\lambda + s)^2[(1 - \alpha_2)(\lambda + \alpha_1 s + 1)(\lambda + \alpha_2 s)^2 + \alpha_2(\lambda + \alpha_2 s + 1)(\lambda + \alpha_1 s)^2]}.$$

Considerando $\alpha_2 = c - \alpha_1 = k$, podemos reescrever $\phi_\epsilon(s)$ da forma:

$$\phi_\epsilon(s) = \left[\frac{(\lambda + s + 1)(\lambda + as)^2(\lambda + ks)^2}{(\lambda + s)^2 z} \right],$$

em que z é determinado pelo polinômio de terceiro grau da forma:

$$z = (1 - k)(\lambda + as + 1)(\lambda + ks)^2 + k(\lambda + ks + 1)(\lambda + as)^2.$$

Com o intuito de obter a função que representa a distribuição do erro aleatório, aplica-se em $\phi_\epsilon(s)$ a transformada inversa de Laplace. Desse modo, considere $H(s) = \phi_\epsilon(s)$. Assim, a transformada inversa de Laplace correspondente será denotada por $h(x)$. Segue:

$$f_\epsilon(x) = h(x).$$

Dessa forma, para encontrar $h(x)$ é necessário decompor a expressão $H(s)$ em frações parciais.

$$\frac{(\lambda + s + 1)(\lambda + as)^2(\lambda + ks)^2}{(s + \lambda)^2 z} = \frac{A^*}{(s + \lambda)} + \frac{B^*}{(s + \lambda)^2} + \frac{C^*}{z}$$

$$(\lambda + s + 1)(\lambda + as)^2(\lambda + ks)^2 = A^*(s + \lambda)z + B^*z + C^*(s + \lambda)^2.$$

Seja $s = -\lambda$:

$$B^* = -\frac{\lambda^4(a - 1)^2(k - 1)^2}{\lambda^2(k - 1)^3(\lambda - a\lambda + 1) - k\lambda^2(a - 1)^2(\lambda - k\lambda + 1)}.$$

Fazendo $s = r$, com r raiz da expressão:

$$[(1 - k)(\lambda + as + 1)(\lambda + ks)^2 + k(\lambda + ks + 1)(\lambda + as)^2]$$

então:

$$C^* = \frac{(\lambda + r + 1)(\lambda + ar)^2(\lambda + kr)^2}{(r + \lambda)^2}.$$

Avaliando em $s = 0$ e substituindo os valores encontrados para B^* e C^* :

$$A^* = \frac{\lambda^4(\lambda + 1) + \frac{\lambda^6(a-1)^2(k-1)^2(\lambda+1)}{\lambda^2(k-1)^3(\lambda-a\lambda+1)-k\lambda^2(a-1)^2(\lambda-k\lambda+1)} - \frac{\lambda^2(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\lambda+r)^2}}{\lambda^3(\lambda + 1)}.$$

Finalmente, pode-se reescrever $H(s)$ da seguinte forma:

$$H(s) = \left\{ \frac{A^*}{(s + \lambda)} + \frac{B^*}{(s + \lambda)^2} + \frac{C^*}{z} \right\}.$$

Aplicando a transformada inversa de Laplace e considerando $h(x)$ a transformada inversa de $H(s)$, temos:

$$h(x) = T_1^* \lambda e^{-\lambda x} + T_2^* x \lambda^2 e^{-\lambda x} + T_{31}^* (-r_1) e^{-r_1 x} + T_{32}^* (-r_2) e^{-r_2 x} + T_{33}^* (-r_3) e^{-r_3 x},$$

com r_1 , r_2 e r_3 raízes negativas do polinômio de grau 3 ao se aplicar a inversa de Laplace em:

$$[(1 - k)(\lambda + as + 1)(\lambda + ks)^2 + k(\lambda + ks + 1)(\lambda + as)^2]^{-1},$$

cujo resultado é uma expressão do tipo:

$$\sum_{i=1}^3 \left(\frac{\exp(tr_i)}{k_i} \right), \quad (\text{C.1})$$

em que r_i representa cada uma das 3 raízes do polinômio:

$$\begin{aligned} z(s) = & (\alpha_1^2 k^2 + \alpha_1 k^2) s^3 + (2\alpha_1 k^2 \lambda + 2\alpha_1 k \lambda + \alpha_1^2 k \lambda + k^2 \lambda + \alpha_1^2 k + k^2) s^2 \\ & + (2\alpha_1 k \lambda^2 + 2\alpha_1 k \lambda + 2k \lambda + 2k \lambda^2 + k^2 \lambda^2 + \alpha_1 \lambda^2) s + \lambda^3 + \lambda^2 + k \lambda^3 + k \lambda^2. \end{aligned}$$

É possível mostrar que $r_i < 0$, $\forall i$, ou seja, podemos escrever $r_i = -\phi_i$, $\forall i \therefore \phi_i > 0$.

Manipulando para aparecer uma exponencial em F.2:

$$\sum_{i=1}^3 \left(\frac{\exp(-\phi_i t)}{k_i} \right). \quad (\text{C.2})$$

Para $i = 1$, por exemplo:

$$\begin{aligned} & \frac{-T_3^*}{k_1 \phi_1} \phi_1 \exp(-\phi_1 t) \\ & = T_{31}^* \phi_1 \exp(-\phi_1 t). \end{aligned}$$

De modo análogo ocorre para $i = 2$ e $i = 3$.

Dessa forma para $i = 1$, considera-se a primeira raiz de $z(s)$. De modo análogo para $i = 2$ e $i = 3$. Já os termos k_1 , k_2 e k_3 são dados por:

$$\left\{ \begin{array}{l} k_1 = (3a^2k^2 - 3ak^3 + 3ak^2)\phi_1^2 + 2k\lambda + \phi_1(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 \\ \quad + 4ak\lambda + 2a^2k\lambda) + (a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2), \\ k_2 = (3a^2k^2 - 3ak^3 + 3ak^2)\phi_2^2 + 2k\lambda + \phi_2(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 \\ \quad + 4ak\lambda + 2a^2k\lambda) + (a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2), \\ k_3 = (3a^2k^2 - 3ak^3 + 3ak^2)\phi_3^2 + 2k\lambda + \phi_3(2a^2k + 2k^2\lambda - 2k^3\lambda + 2k^2 - 2k^3 \\ \quad + 4ak\lambda + 2a^2k\lambda) + (a\lambda^2 + 2k\lambda^2 - 2k^2\lambda - k^2\lambda^2 + 2ak\lambda + ak\lambda^2). \end{array} \right.$$

Já os pesos são determinados por:

$$T_1^* = \frac{A^*}{\lambda} = \frac{\lambda^4(\lambda + 1) + \frac{\lambda^6(a-1)^2(k-1)^2(\lambda+1)}{\lambda^2(k-1)^3(\lambda-a\lambda+1)-k\lambda^2(a-1)^2(\lambda-k\lambda+1)} - \frac{\lambda^2(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\lambda+r)^2}}{\lambda^4(\lambda + 1)},$$

$$T_2^* = \frac{B^*}{\lambda^2} = -\frac{\lambda^2(a-1)^2(k-1)^2}{\lambda^2(k-1)^3(\lambda-a\lambda+1)-k\lambda^2(a-1)^2(\lambda-k\lambda+1)},$$

$$T_{31}^* = -\frac{(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\phi_1) \left[(\lambda+r)^2(2k\lambda+a\lambda^2+\phi_1^2(3a^2k^2-3ak^3+3ak^2)+2k\lambda^2-2k^2\lambda) + \phi_1(2a^2k+2k^2\lambda-2k^3\lambda+2k^2-2k^3+4ak\lambda+2a^2k\lambda) - k^2\lambda^2+2ak\lambda+ak\lambda^2 \right]},$$

$$T_{32}^* = -\frac{(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\phi_2) \left[(\lambda+r)^2(2k\lambda+a\lambda^2+\phi_2^2(3a^2k^2-3ak^3+3ak^2)+2k\lambda^2-2k^2\lambda) + \phi_2(2a^2k+2k^2\lambda-2k^3\lambda+2k^2-2k^3+4ak\lambda+2a^2k\lambda) - k^2\lambda^2+2ak\lambda+ak\lambda^2 \right]} \text{ e}$$

$$T_{33}^* = -\frac{(\lambda+kr)^2(\lambda+ar)^2(\lambda+r+1)}{(\phi_3) \left[(\lambda+r)^2(2k\lambda+a\lambda^2+\phi_3^2(3a^2k^2-3ak^3+3ak^2)+2k\lambda^2-2k^2\lambda) + \phi_3(2a^2k+2k^2\lambda-2k^3\lambda+2k^2-2k^3+4ak\lambda+2a^2k\lambda) - k^2\lambda^2+2ak\lambda+ak\lambda^2 \right]},$$

de modo que

$$T_{31}^* = -\frac{T_3^*}{k_1\phi_1}, \quad T_{32}^* = -\frac{T_3^*}{k_2\phi_2}, \quad T_{33}^* = -\frac{T_3^*}{k_3\phi_3},$$

em que:

$$T_3^* = \frac{(\lambda + kr)^2(\lambda + ar)^2(\lambda + r + 1)}{(\lambda + r)^2}.$$

Assim, a distribuição da v.a ϵ_t corresponde a uma mistura do tipo:

$$h(x) = T_1^* \lambda e^{-\lambda x} + T_2^* x \lambda^2 e^{-\lambda x} + T_{31}^* (\phi_1) e^{-\phi_1 x} + T_{32}^* (\phi_2) e^{-\phi_2 x} + T_{33}^* (\phi_3) e^{-\phi_3 x}.$$

Com a soma dos pesos da forma:

$$\boxed{T_1^* + T_2^* + T_{31}^* + T_{32}^* + T_{33}^* = 1}.$$

APÊNDICE D –

Cálculos Gama-Lindley (Caso 1)

Considerando a Proposição 4.2.1, que corresponde ao modelo $GLAR_1(2)$, assim como no caso anterior, tem-se que a transformada de Laplace é dada por:

$$\phi_X(s) = [(1 - \alpha_2)\phi_{X_{i-1}}(\alpha_1 s) + \alpha_2\phi_{X_{i-2}}(\alpha_2 s)]\phi_\epsilon(s).$$

A próxima etapa será adaptar a expressão acima usando a transformada da Gama-Lindley de parâmetros (λ, β) do tipo:

$$\phi_X(s) = \frac{\lambda^2(\beta\lambda + \beta + s)}{\beta(1 + \lambda)(s + \lambda)^2}.$$

Temos que a transformada de Laplace para o erro aleatório é:

$$\phi_\epsilon(s) = \frac{\phi_X(s)}{(1 - \alpha_2)\phi_{X_{i-1}}(\alpha_1 s) + \alpha_2\phi_{X_{i-2}}(\alpha_2 s)}.$$

Admitindo a estacionariedade do processo, vamos calcular $\phi_X(\alpha_1 s)$ e $\phi_X(\alpha_2 s)$.

$$\phi_X(\alpha_1 s) = \frac{\lambda^2(\beta\lambda + \beta + \alpha_1 s)}{\beta(1 + \lambda)(\alpha_1 s + \lambda)^2} \quad \text{e} \quad \phi_X(\alpha_2 s) = \frac{\lambda^2(\beta\lambda + \beta + \alpha_2 s)}{\beta(1 + \lambda)(\alpha_2 s + \lambda)^2}.$$

Substituindo na expressão da transformada de Laplace para ϵ_s :

$$\phi_\epsilon(s) = \frac{\frac{\lambda^2(\beta\lambda + \beta + s)}{\beta(1 + \lambda)(s + \lambda)^2}}{\frac{(1 - \alpha_2)\lambda^2(\beta\lambda + \beta + \alpha_1 s)}{\beta(1 + \lambda)(\alpha_1 s + \lambda)^2} + \frac{\alpha_2\lambda^2(\beta\lambda + \beta + \alpha_2 s)}{\beta(1 + \lambda)(\alpha_2 s + \lambda)^2}}.$$

Fazendo as simplificações acima, chegamos na expressão:

$$\phi_\epsilon(s) = \left[\frac{(\beta\lambda + \beta + s)(\alpha_1 s + \lambda)^2(\alpha_2 s + \lambda)^2}{(s + \lambda)^2[(1 - \alpha_2)(\beta\lambda + \beta + \alpha_1 s)(\alpha_2 s + \lambda)^2 + \alpha_2(\beta\lambda + \beta + \alpha_2 s)(\alpha_1 s + \lambda)^2]} \right].$$

Considerando também a restrição para:

$$\alpha_1 = \alpha_2 = a, \quad a \in (0, 1).$$

Reescrevendo a expressão acima, temos:

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

Simplificando, chega-se a expressão:

$$\phi_\epsilon(s) = \left[\frac{(\beta\lambda + \beta + s)(as + \lambda)^2}{(s + \lambda)^2(\beta\lambda + \beta + as)} \right].$$

Com o objetivo de obter a distribuição do erro aleatório, aplica-se em $\phi_\epsilon(s)$ a transformada inversa de Laplace, usando o método de frações parciais, assim como no caso anterior. Desse modo, considere $I(s) = \phi_\epsilon(s)$. Assim, a transformada inversa de Laplace correspondente será denotada por $i(x)$. Segue:

$$f_\epsilon(x) = i(x).$$

Dessa forma, para encontrar $i(x)$ é necessário decompor a expressão $I(s)$ em frações parciais:

$$\frac{(\beta\lambda + \beta + s)(\lambda + as)^2}{(\lambda + s)^2(\lambda + as + 1)} = \frac{T_1}{(\lambda + s)} + \frac{T_2}{(\lambda + s)^2} + \frac{T_3}{(\beta\lambda + \beta + as)}.$$

Fazendo $s = -\lambda$, tem-se:

$$T_2 = \frac{\lambda^2(a-1)^2(\beta + \beta\lambda - \lambda)}{(\beta + \beta\lambda - a\lambda)}.$$

Seja $s = \frac{-\beta(1+\lambda)}{a}$, logo, tem-se:

$$T_3 = \frac{(\beta - \lambda + \beta\lambda)^2(\beta - \frac{\beta+\beta\lambda}{a} + \beta\lambda)}{(\lambda - \frac{\beta+\beta\lambda}{a})^2}.$$

Finalmente avaliando $s = 0$ e substituindo os valores encontrados dos termos T_2 e T_3 , obtém-se:

$$T_1 = -\frac{\frac{\beta\lambda^2(a-1)^2(\lambda+1)(\beta-\lambda+\beta\lambda)}{\beta-a\lambda+\beta\lambda} - \beta\lambda^2(\lambda+1) + \frac{a\beta\lambda^2(a-1)(\lambda+1)(\beta-\lambda+\beta\lambda)^2}{(\beta-a\lambda+\beta\lambda)^2}}{\lambda(\beta + \beta\lambda)}.$$

Desse modo, observa-se que $I(s)$ pode ser reescrita do seguinte modo:

$$I(s) = \frac{T_1}{(s + \lambda)} + \frac{T_2}{(s + \lambda)^2} + \frac{T_3}{(\beta\lambda + \beta + as)}.$$

Essa etapa consiste em aplicar a transformada inversa de Laplace em cada um dos termos separadamente, visto que temos funções conhecidas. Assim, considerando $i(x)$ a transformada inversa de $I(s)$, tem-se:

$$i(x) = T_1^* \lambda e^{-\lambda x} + T_2^* x \lambda^2 e^{-\lambda x} + T_3^* \left(\frac{\beta\lambda + \beta}{a} \right) e^{-\left(\frac{\beta\lambda + \beta}{a} \right) x}.$$

E os pesos finais são representados por:

$$T_1^* = \left(\frac{T_1}{\lambda} \right) = -\frac{\frac{\beta\lambda^2(a-1)^2(\lambda+1)(\beta-\lambda+\beta\lambda)}{\beta-a\lambda+\beta\lambda} - \beta\lambda^2(\lambda+1) + \frac{a\beta\lambda^2(a-1)(\lambda+1)(\beta-\lambda+\beta\lambda)^2}{(\beta-a\lambda+\beta\lambda)^2}}{\lambda^2(\beta + \beta\lambda)},$$

$$T_2^* = \left(\frac{T_2}{\lambda^2} \right) = \frac{(a-1)^2(\beta - \lambda + \beta\lambda)}{\beta - a\lambda + \beta\lambda} \quad \text{e} \quad T_3^* = \frac{T_3 a}{(\beta\lambda + \beta)} = \frac{a(\beta - \lambda + \beta\lambda)^2 \left(\beta - \frac{\beta+\beta\lambda}{a} + \beta\lambda \right)}{\left(\lambda - \frac{\beta+\beta\lambda}{a} \right)^2 (\beta + \beta\lambda)}.$$

Já a expressão da soma dos pesos em função dos parâmetros β , λ e a é da forma:

$$S(a, \beta, \lambda) = \frac{\begin{pmatrix} a^3\beta^2\lambda^2 + 2a^3\beta^2\lambda + a^3\beta^2 - 2a^3\beta\lambda^2 - 2a^3\beta\lambda + a^3\lambda^2 \\ - 2a^2\beta^2\lambda^2 - 4a^2\beta^2\lambda - 2a^2\beta^2 + 4a^2\beta\lambda^2 + 4a^2\beta\lambda - a^2\lambda^2 \\ + a\beta^2\lambda^2 + 2a\beta^2\lambda + a\beta^2 - 4a\beta\lambda^2 - 4a\beta\lambda + a\lambda^2 \\ + \beta^2\lambda^2 + 2\beta^2\lambda + \beta^2 \end{pmatrix}}{(\beta - a\lambda + \beta\lambda)^2}$$

Assim, $i(x)$ é uma mistura de distribuições Gama $(2, \lambda)$ e exponenciais de parâmetros (λ) e $\left(\frac{\beta\lambda + \beta}{a}\right)$, com a soma dos pesos:

$$\boxed{T_1^* + T_2^* + T_3^* = 1}.$$

APÊNDICE E –

Analisando o espaço paramétrico para o λ e β no Caso 1, Gama-Lindley.

Temos que a soma (S) dos pesos é uma função com três parâmetros: a , λ e β da forma:

$$S(a, \beta, \lambda) = \frac{a^3\beta\lambda^2 + a^3\beta\lambda - a^3\lambda^2 - 2a^2\beta^2\lambda^2 - 4a^2\beta^2\lambda - 2a^2\beta^2 + a^2\beta\lambda^2 + a^2\beta\lambda + 2a^2\lambda^2}{(\beta - a\lambda + \beta\lambda)^2} + \frac{3a\beta^2\lambda^2 + 6a\beta^2\lambda + 3a\beta^2 - 5a\beta\lambda^2 - 5a\beta\lambda + \beta\lambda^2 + \beta\lambda}{(\beta - a\lambda + \beta\lambda)^2} + \frac{(a-1)^2(\beta - \lambda + \beta\lambda)}{(\beta - a\lambda + \beta\lambda)} - \frac{a\beta^2(a-1)(\lambda+1)^2(\beta - \lambda + \beta\lambda)^2}{(\beta - a\lambda + \beta\lambda)^2}.$$

Como o modelo é estável para o termo autorregressivo a , no intervalo $[0.1; 0.49]$, considerou-se os valores extremos, ou seja, os limites inferiores e superiores da faixa, o que permite identificar a região de estabilidade para os parâmetros β e λ . Assim, seja $a = 0.1$, substituindo na expressão acima:

$$S(\beta, \lambda) = \frac{\frac{7\beta^2\lambda^2}{25} + \frac{14\beta^2\lambda}{25} + \frac{7\beta^2}{25} + \frac{511\beta\lambda^2}{1000} + \frac{511\beta\lambda}{1000} + \frac{19\lambda^2}{1000}}{(\beta - \frac{\lambda}{10} + \beta\lambda)^2} + \frac{\frac{81\beta}{100} - \frac{81\lambda}{100} + \frac{81\beta\lambda}{100}}{(\beta - \frac{\lambda}{10} + \beta\lambda)} - \frac{\frac{9\beta^2(\lambda+1)^2(\beta - \lambda + \beta\lambda)^2}{100}}{(\beta - \frac{\lambda}{10} + \beta\lambda)^2}.$$

Como a soma $S = 1$, resolvendo a equação em função de λ tem-se que as raízes para S são do tipo:

$$R_1 = \frac{-\beta}{(\beta - 1)}, \quad R_2 = \frac{-(\beta - 1i)}{\beta}, \quad R_3 = \frac{-(\beta + 1i)}{\beta}.$$

A partir daí, a única solução possível será R_1 , considerando apenas no contexto real, e como $\beta > 0$, para termos λ também positivo, tem-se que $\beta < 1$. Assim, a região será $0 < \beta < 1$.

Considerando agora $a = 0.49$, dentro da região de estabilidade, temos que:

$$S = \begin{cases} \frac{2601\beta - 2601\lambda + 2601\beta\lambda}{10000(\beta - \frac{49\lambda}{100} + \beta\lambda)} + \frac{4949\beta^2\lambda^2 + 9898\beta^2\lambda + 4949\beta^2}{5000(\beta - \frac{49\lambda}{100} + \beta\lambda)^2} \\ - \frac{49190611965950516729\beta\lambda^2 + 49190611965950516729\beta\lambda - 16327845485028007417\lambda^2}{45035996273704960000(\beta - \frac{49\lambda}{100} + \beta\lambda)^2} \\ + \frac{2499\beta^2(\lambda+1)^2(\beta - \lambda + \beta\lambda)^2}{10000(\beta - \frac{49\lambda}{100} + \beta\lambda)^2} \end{cases}$$

De modo semelhante, como $S = 1$, isolando a expressão em função de λ tem-se o polinô-

mio:

$$\begin{aligned}
 p(\lambda) = & \lambda^3(11448664824217600\beta^3 - 11448664824217600\beta^2) \\
 & + \lambda^2(34345994472652800\beta^3 - 22897329648435200\beta^2) \\
 & + \lambda(34345994472652800\beta^3 - 11448664824217600\beta^2 \\
 & + 11448664824217600\beta) \\
 & - 11448664824217601 + 11448664824217600\beta + 11448664824217600\beta^3.
 \end{aligned}$$

Ao resolver $p(\lambda) = 0$, encontramos também como raízes:

$$R_1 = \frac{-\beta}{(\beta - 1)}, \quad R_2 = \frac{-(\beta - 1i)}{\beta}, \quad R_3 = \frac{-(\beta + 1i)}{\beta}.$$

Sendo assim, de forma análoga ao caso anterior, a única solução real possível é R_1 , com $0 < \beta < 1$. Dessa forma, para qualquer valor de a , dentro da região de estabilidade, a relação entre os parâmetros β e λ coincidem com a expressão determinada em R_1 . De modo geral, para qualquer valor do parâmetro autorregressivo a , as raízes de $f(\lambda)$ são sempre os termos R_1 , R_2 e R_3 .

APÊNDICE F –

Cálculos da Proposição 4.3.1

A transformada de Laplace para um modelo AR(2) é da forma:

$$\phi_X(s) = [(1 - \alpha_2)\phi_{X_{i-1}}(\alpha_1 s) + \alpha_2\phi_{X_{i-2}}(\alpha_2 s)]\phi_\epsilon(s).$$

Para o caso geral, a transformada para o erro aleatório da distribuição Gama-Lindley é:

$$\phi_\epsilon(s) = \left[\frac{(\beta\lambda + \beta + s)(\alpha_1 s + \lambda)^2(\alpha_2 s + \lambda)^2}{(s + \lambda)^2[(1 - \alpha_2)(\beta\lambda + \beta + \alpha_1 s)(\alpha_2 s + \lambda)^2 + \alpha_2(\beta\lambda + \beta + \alpha_2 s)(\alpha_1 s + \lambda)^2]} \right].$$

Assumindo que

$$\alpha_2 = c - \alpha_1 = k, \quad 0 < c < 1 < \alpha_1,$$

pode-se reescrever $\phi_\epsilon(s)$ da forma

$$\phi_\epsilon(s) = \left[\frac{(\beta\lambda + \beta + s)(\alpha_1 s + \lambda)^2(ks + \lambda)^2}{(s + \lambda)^2[(1 - k)(\beta\lambda + \beta + \alpha_1 s)(ks + \lambda)^2 + k(\beta\lambda + \beta + ks)(\alpha_1 s + \lambda)^2]} \right].$$

Com o objetivo de obter a função que representa a distribuição do erro aleatório, utiliza-se em $\phi_\epsilon(s)$ a transformada inversa de Laplace. Dessa maneira, considere $I(s) = \phi_\epsilon(s)$. Assim, a transformada inversa de Laplace correspondente será denotada por $i(x)$. Segue:

$$f_\epsilon(x) = i(x).$$

Dessa maneira, para encontrar $i(x)$ é necessário decompor a expressão $I(s)$ em frações parciais.

$$\frac{(\beta\lambda + s + \beta)(\lambda + as)^2(\lambda + ks)^2}{(s + \lambda)^2 z} = \frac{D^*}{(s + \lambda)} + \frac{E^*}{(s + \lambda)^2} + \frac{F^*}{z}$$

$$(\beta\lambda + \beta + s)(\lambda + as)^2(\lambda + ks)^2 = D^*(s + \lambda)z + E^*z + F^*(s + \lambda)^2.$$

O termo z é determinado por:

$$(1 - k)(\beta\lambda + \beta + as)(ks + \lambda)^2 + k(\beta\lambda + \beta + ks)(as + \lambda)^2.$$

Agora, vamos encontrar os coeficientes D^* , E^* e F^* .

Fazendo $s = -\lambda$:

$$(\lambda - k\lambda)^2(\lambda - a\lambda)^2(\beta - \lambda + \beta\lambda) = E^*z$$

$$E^* = -\frac{(\lambda - k\lambda)^2(\lambda - a\lambda)^2(\beta - \lambda + \beta\lambda)}{(k(\lambda - a\lambda)^2(\beta + \beta\lambda - k\lambda) - (\lambda - k\lambda)^2(k - 1)(\beta - a\lambda + \beta\lambda))}.$$

Fazendo $s = r$, sendo r raíz de:

$$[(1 - k)(\beta\lambda + \beta + as)(ks + \lambda)^2 + k(\beta\lambda + \beta + ks)(as + \lambda)^2],$$

então:

$$F^* = \frac{(\beta\lambda + \beta + r)(\lambda + ar)^2(\lambda + kr)^2}{(r + \lambda)^2}.$$

Avaliando em $s = 0$ e substituindo os valores encontrados para B^* e C^* , tem-se:

$$\begin{aligned} D^* = & - \left(\frac{(\lambda - k\lambda)^2 \left(k\lambda^2(\beta + \beta\lambda) - \lambda^2(k - 1)(\beta + \beta\lambda) \right) (\lambda - a\lambda)^2 (\beta - \lambda + \beta\lambda)}{k(\lambda - a\lambda)^2 (\beta + \beta\lambda - k\lambda) - (\lambda - k\lambda)^2 (k - 1)(\beta - a\lambda + \beta\lambda)} \right. \\ & \left. - \lambda^4 (\beta + \beta\lambda) + \frac{\lambda^2 (\lambda + kr)^2 (\lambda + ar)^2 (\beta + r + \beta\lambda)}{(\lambda + r)^2} \right) \\ & \div \left(\lambda^2 \left(k\lambda^2(\beta + \beta\lambda) - \lambda^2(k - 1)(\beta + \beta\lambda) \right) \right). \end{aligned}$$

Por fim, pode-se reescrever $I(s)$ como uma expressão da tipo:

$$I(s) = \left\{ \frac{D^*}{(s + \lambda)} + \frac{E^*}{(s + \lambda)^2} + \frac{F^*}{z} \right\}.$$

Aplicando a transformada inversa de Laplace e considerando $i(x)$ a transformada inversa de $H(s)$, temos:

$$i(x) = T_1 \lambda e^{-\lambda x} + T_2 x \lambda^2 e^{-\lambda x} + T_{31} \varphi_1 e^{-\varphi_1 x} + T_{32} \varphi_2 e^{-\varphi_2 x} + T_{33} \varphi_3 e^{-\varphi_3 x},$$

com $r_j = -\varphi_j$, as raízes negativas do polinômio de grau 3 ao se aplicar a inversa de laplace em:

$$(1 - k)(\beta\lambda + \beta + as)(ks + \lambda)^2 + k(\beta\lambda + \beta + ks)(as + \lambda)^2,$$

em que o resultado possui uma estrutura semelhante ao encontrado no modelo $LAR_2(2)$, da forma:

$$\sum_{j=1}^3 \left(\frac{\exp(tr_j)}{k_j} \right), \quad (F.1)$$

com r_j representando as raízes do polinômio de grau 3 em s :

$$\begin{aligned} z(s) = & (a^2 k^2 - ak^3 + ak^2) s^3 + (\beta k^2 - \beta k^3 + 2ak\lambda + a^2 \beta k + \beta k^2 \lambda - \beta k^3 \lambda + a^2 \beta k \lambda) s^2 \\ & + (a\lambda^2 + k^2 \lambda^2 - 2\beta k^2 \lambda^2 + 2\beta k \lambda - ak\lambda^2 + 2\beta k \lambda^2 - 2\beta k^2 \lambda + 2a\beta k \lambda^2 + 2a\beta k \lambda) s \\ & + \beta \lambda^3 + \beta \lambda^2. \end{aligned}$$

De modo análogo ao LAR_2 , prova-se $r_j < 0$, $\forall i$, ou seja, podemos escrever $r_j = -\varphi_i$, $\forall i$ $\therefore \varphi_i > 0$. Manipulando para aparecer uma exponencial em F.1:

$$\sum_{j=1}^3 \left(\frac{\exp(-\varphi_j t)}{k_j} \right). \quad (\text{F.2})$$

Para $j = 1$, por exemplo:

$$\begin{aligned} & \frac{-T_3}{k_1 \varphi_1} \exp(-\varphi_1 t) \\ &= T_{31} \varphi_1 \exp(-\varphi_1 t). \end{aligned}$$

De modo análogo ocorre para $j = 2$ e $j = 3$.

Os termos k_1 , k_2 e k_3 possuem a seguinte estrutura:

$$k_1 = (\lambda + r)^2 w^*, \quad k_2 = (\lambda + r)^2 l^* \quad \text{e} \quad k_3 = (\lambda + r)^2 m^*,$$

com:

$$\left\{ \begin{aligned} w^* &= \varphi_1 \left(2\beta k^2 - 2\beta k^3 + 4ak\lambda + 2a^2\beta k + 2\beta k^2\lambda - 2\beta k^3\lambda + 2a^2\beta k\lambda \right) + a\lambda^2 \\ &\quad + \varphi_1^2 \left(3a^2k^2 - 3ak^3 + 3ak^2 \right) + k^2\lambda^2 - 2\beta k^2\lambda^2 + 2\beta k\lambda - ak\lambda^2 + 2\beta k\lambda^2 - 2\beta k^2\lambda, \\ l^* &= \varphi_2 \left(2\beta k^2 - 2\beta k^3 + 4ak\lambda + 2a^2\beta k + 2\beta k^2\lambda - 2\beta k^3\lambda + 2a^2\beta k\lambda \right) + a\lambda^2 \\ &\quad + \varphi_2^2 \left(3a^2k^2 - 3ak^3 + 3ak^2 \right) + k^2\lambda^2 - 2\beta k^2\lambda^2 + 2\beta k\lambda - ak\lambda^2 + 2\beta k\lambda^2 - 2\beta k^2\lambda, \\ m^* &= \varphi_3 \left(2\beta k^2 - 2\beta k^3 + 4ak\lambda + 2a^2\beta k + 2\beta k^2\lambda - 2\beta k^3\lambda + 2a^2\beta k\lambda \right) + a\lambda^2 \\ &\quad + \varphi_3^2 \left(3a^2k^2 - 3ak^3 + 3ak^2 \right) + k^2\lambda^2 - 2\beta k^2\lambda^2 + 2\beta k\lambda - ak\lambda^2 + 2\beta k\lambda^2 - 2\beta k^2\lambda. \end{aligned} \right.$$

Os pesos são dados por:

$$\begin{aligned} T_1 = \frac{D^*}{\lambda} = & - \left(\frac{(\lambda - k\lambda)^2 \left(k\lambda^2(\beta + \beta\lambda) - \lambda^2(k - 1)(\beta + \beta\lambda) \right) (\lambda - a\lambda)^2(\beta - \lambda + \beta\lambda)}{k(\lambda - a\lambda)^2(\beta + \beta\lambda - k\lambda) - (\lambda - k\lambda)^2(k - 1)(\beta - a\lambda + \beta\lambda)} \right. \\ & \left. - \lambda^4(\beta + \beta\lambda) + \frac{\lambda^2(\lambda + kr)^2(\lambda + ar)^2(\beta + r + \beta\lambda)}{(\lambda + r)^2} \right) \\ & \div \left(\lambda^2 \left(k\lambda^2(\beta + \beta\lambda) - \lambda^2(k - 1)(\beta + \beta\lambda) \right) \right), \end{aligned}$$

$$T_2 = \frac{E^*}{\lambda^2} = \frac{(\lambda - k\lambda)^2(\lambda - a\lambda)^2(\beta - \lambda + \beta\lambda)}{\lambda^2 \left(k(\lambda - a\lambda)^2(\beta + \beta\lambda - k\lambda) - (\lambda - k\lambda)^2(k - 1)(\beta - a\lambda + \beta\lambda) \right)},$$

$$T_3 = F^* = \frac{(\lambda + kr)^2(\lambda + ar)^2(\beta + r + \beta\lambda)}{(\lambda + r)^2},$$

$$T_{31} = -\frac{T_3}{k_1\varphi_1} = -\frac{(\lambda + kr)^2(\lambda + ar)^2(\beta + r + \beta\lambda)}{\varphi_1(\lambda + r)^2w^*},$$

$$T_{32} = -\frac{T_3}{k_2\varphi_2} = -\frac{(\lambda + kr)^2(\lambda + ar)^2(\beta + r + \beta\lambda)}{\varphi_2(\lambda + r)^2l^*}$$

e

$$T_{33} = -\frac{T_3}{k_3\varphi_3} = -\frac{(\lambda + kr)^2(\lambda + ar)^2(\beta + r + \beta\lambda)}{\varphi_3(\lambda + r)^2m^*}.$$

Assim, a distribuição da v.a ϵ_t para o modelo $\text{GLAR}_2(2)$ corresponde a uma mistura do tipo:

$$h(x) = T_1\lambda e^{-\lambda x} + T_2x\lambda^2 e^{-\lambda x} + T_{31}\varphi_1 e^{-\varphi_1 x} + T_{32}\varphi_2 e^{-\varphi_2 x} + T_{33}\varphi_3 e^{-\varphi_3 x}.$$

Com a soma dos pesos é da forma:

$$\boxed{T_1 + T_2 + T_{31} + T_{32} + T_{33} = 1}.$$