
MODELOS INAR SAZONAIS E
DE RAÍZES UNITÁRIAS

MARCELO BOURGUIGNON PEREIRA

Orientador: Prof. Dr. Klaus Leite Pinto Vasconcellos

Co-orientador: Prof. Dr. Valdério Anselmo Reisen

Área de Concentração: Estatística Matemática

Dissertação submetida como requerimento parcial para obtenção do grau
de Mestre em Estatística pela Universidade Federal de Pernambuco

Recife, fevereiro de 2011.

Catálogo na fonte
Bibliotecária Jane Souto Maior, CRB4-571

Pereira, Marcelo Bourguignon
Modelos inar sazonais e de raízes unitárias / Marcelo
Bourguignon Pereira - Recife: O Autor, 2011.
xii, 68 folhas : il., fig., tab.

Orientador: Klaus Leite Pinto Vasconcellos.
Dissertação (mestrado) Universidade Federal de
Pernambuco. CCEN. Estatística, 2011.

Inclui bibliografia e apêndice.

1. Séries temporais. I. Vasconcellos, Klaus Leite Pinto
(orientador). II. Título.

519.232

CDD (22. ed.)

MEI2011 – 022

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

17 de fevereiro de 2011

Nós recomendamos que a dissertação de mestrado de autoria de

Marcelo Bourguignon Pereira

intitulada

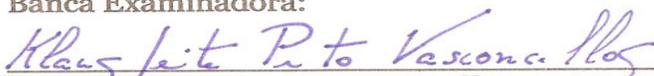
“Modelos INAR Sazonais e de raízes unitárias”

seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.



Coordenador da Pós-Graduação em Estatística

Banca Examinadora:



Klaus Leite Pinto Vasconcellos

orientador



Pedro Alberto Morettin (USP)



Renato José de Sobral Cintra

Este documento será anexado à versão final da dissertação.

Dedico este trabalho a minha mãe Orly e meu pai Alberto.

Agradecimentos

Quero agradecer em primeiro lugar a Deus, ser supremo, fonte de força inesgotável por ter me dado força, perseverança e paciência, qualidades mais do que necessárias para concluir o mestrado.

Agradeço muito especialmente aos meus pais pelo interesse em me ver concluir o curso e por serem minha eterna inspiração para continuar na luta por um objetivo na vida.

Ao professor Dr. Valdério Anselmo Reisen por ter me conduzido até aqui, ter confiado em meu potencial, pela dedicação na minha formação acadêmica e por todo o apoio dos últimos cinco anos.

Ao meu orientador Professor Dr. Klaus Leite Pinto Vasconcellos pela orientação, paciência, dedicação, sugestões e valiosas recomendações que tornaram possível este trabalho.

Ao professor Dr. Domingos do departamento de matemática da Universidade Federal do Espírito Santo pela confiança depositada em mim.

Aos meus amigos de Pernambuco, com destaque para os amigos Agrinaldo, Josivandro, Dona Dada, Jeremias, Iván, Lutemberg, Manoel, Natasha, Poema, Josimar e Helton, que se tornaram a minha família nesses dois anos de mestrado e com certeza foram fundamentais para a conclusão do mestrado, gostaria de expressar minha profunda gratidão.

Aos demais colegas da pós-graduação e classe que de alguma forma contribuíram para o desfecho meritório deste mestrado: Davis, Cícero, Laércio, Diego, Silvio, Francisco e Marcela.

A Valéria Bittencourt, secretária da pós-graduação em estatística, pela competência, carinho, amizade e atenção.

A minha namorada Gleicielle pela paciência, apoio e incentivo constantes.

Aos amigos de infância pelo apoio enquanto estava em outro estado.

Aos amigos Alessandro e Giovanni, pessoas nas quais sempre estiveram dispostas a me ajudar desde a época de graduação.

Ao amigo Fabio Fajardo pela troca de conhecimento e apoio.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro a este projeto e ao PIBIC-CNPq pela bolsa de iniciação científica importante para o surgimento do interesse em pesquisas.

Ao estado de Pernambuco, em especial à cidade do Recife.

A todos aqueles que acreditam em sonhos.

Não sou nada

Nunca serei nada

Não posso querer ser nada

À parte isso, tenho em mim todos os sonhos do mundo.

(Fernando Pessoa).

Resumo

Séries temporais de contagem têm chamado a atenção pela importância em aplicações nas diversas áreas de conhecimento. Os processos estocásticos usuais assumem que as marginais são contínuas e, em geral, não são adequados para modelar séries de contagem. Portanto, surge a necessidade de investigar metodologias apropriadas para séries temporais com distribuições marginais discretas. Em particular, o estudo da presença de raízes unitárias e o comportamento sazonal do processo de valores inteiros motivam uma vertente de pesquisa de grande interesse para aplicações práticas e são os principais objetivos desta pesquisa.

Nesse contexto, apresentamos o teste de Dikey & Fuller (1979) e verificamos o comportamento do teste, através de ensaios de Monte Carlo, em processos autorregressivos de valores inteiros de ordem um, quando o processo apresenta raiz unitária. Os pontos críticos empíricos da estatística de teste do teste de Dickey-Fuller, para vários valores do percentil α , são calculados quando o teste é utilizado em processos INAR(1) com erros Poisson, para diversos valores do parâmetro λ . Comparações entre a utilização do teste de Dickey-Fuller em processos com marginais contínuas e discretas também são abordadas.

No que tange à sazonalidade em processos de contagem, é proposto um modelo de valores inteiros com estrutura sazonal baseado no modelo de Al-Osh & Alzaid (1987). As principais propriedades do modelo proposto são derivadas, tais como os momentos, a função de autocovariância e a função de autocorrelação. Ensaios de Monte Carlo são realizados para comparar os vícios e erros quadráticos médios de três estimadores para os parâmetros do modelo proposto. Como motivação do uso da metodologia sugerida, a série do índice da qualidade do ar da cidade de Cariacica-ES foi analisada.

Palavras-chave: Distribuição de Poisson; processos de contagem; processos INAR; teste de Dickey-Fuller; sazonalidade.

Abstract

Time series of counts have been the focus of attention due to their importance in several areas of knowledge. The usual stochastic processes assume that the marginals are continuous and, in general, are not adequate for modeling series of counts. Therefore, there arises the need to investigate methodologies for time series with discrete marginal distributions. In particular, the study of the presence of unit roots and the seasonal behavior of integer-valued processes motivate an important area of research to practical applications and are the main objectives of this work.

In this context, we present the Dickey & Fuller test (1979) and verify the behavior of the test through Monte Carlo simulations in integer-valued autoregressive processes of order one, when the process has a unit root. The empirical critical points of the test statistic of Dickey-Fuller test for various values of percentile α are computed when the test is used in INAR(1) processes with Poisson errors for many values of the parameter λ . Comparisons between the use of the Dickey-Fuller test processes with discrete and continuous marginals are also discussed.

Regarding seasonality in counting processes, we propose an integer-valued model with seasonal structure based on the model by Al-Osh & Alzaid (1987). The main properties of the proposed model are derived, such as the moments, the autocovariance function, and the autocorrelation function. Monte Carlo simulations are performed in order to compare the bias and mean squared errors of three estimators for the parameters of the model. To motivate the proposed methodology, the series of air quality index in the city of Cariacica-ES was analyzed.

Keywords: Poisson distribution; counting process; INAR process; Dickey-Fuller test; seasonality.

Lista de Figuras	x
Lista de Tabelas	xii
1 Introdução	1
1.1 Preliminares	1
1.2 Objetivos da dissertação	7
1.3 Organização da dissertação	7
1.4 Suporte computacional	8
2 Processo Autorregressivo de Valores Inteiros	9
2.1 Introdução	9
2.2 O processo INAR	10
2.3 Propriedades	12
2.4 Estimação dos parâmetros	15
2.4.1 Estimador de Yule-Walker	15
2.4.2 Estimador de mínimos quadrados condicional	16
2.4.3 Estimador de máxima verossimilhança condicional	18
2.5 Previsão	20
3 Teste de Dickey-Fuller em processos INAR	21

3.1	Introdução	21
3.2	Raízes unitárias e o teste de Dickey-Fuller	22
3.3	Teste de Dickey-Fuller em processos INAR	24
3.4	Resultados numéricos	26
4	Processo Autorregressivo de Valores Inteiros com Estrutura Sazonal	36
4.1	Introdução	36
4.2	O processo SINAR: definição e propriedades	37
4.2.1	Estimador de Yule-Walker	40
4.2.2	Estimador de mínimos quadrados condicional	40
4.2.3	Estimador de máxima verossimilhança condicional	41
4.3	Previsão	42
4.4	Resultados numéricos	44
4.5	Aplicação	47
5	Considerações finais	54
	Apêndice	56
	Referências bibliográficas	63

Lista de Figuras

1.1	Realização de processos independentes e identicamente distribuídos com distribuição $\mathcal{N}(1, 1)$ e $\mathcal{P}o(1)$ ambos com tamanho amostral $n = 300$	3
1.2	Realização de processos independentes e identicamente distribuídos com distribuição $\mathcal{P}o(20)$ com tamanho amostral $n = 300$	4
2.1	Função de autocorrelação amostral do processo INAR(1).	14
3.1	Densidade empírica dos testes sob \mathcal{H}_0 , $n = 100$ e normal padrão.	27
3.2	Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $n = 100$ com λ variando.	28
3.3	Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $\lambda = 0.5$ com n variando.	28
3.4	Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $\lambda = 3$ com n variando.	29
3.5	Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $\lambda = 10$ e da normal padrão.	29
4.1	Série simulada de um processo SINAR(1) ₁₂ , com $P = 0.9$, $\lambda = 1$ e tamanho amostral $n = 100$	38
4.2	Função de autocorrelação amostral do processo SINAR(1) ₁₂	39
4.3	Gráfico de evolução da série índice de qualidade do ar da cidade de Cariacica-ES.	49
4.4	ACF da série temporal do índice de qualidade do ar da cidade de Cariacica-ES.	49
4.5	Histograma e box-plot do índice de qualidade do ar da cidade de Cariacica-ES.	50
4.6	Box-plot dos resíduos.	51

4.7	ACF dos resíduos.	52
4.8	Série observada e série ajustada.	53

Lista de Tabelas

3.1	Valores críticos de $\hat{\tau}_\lambda$ e $\hat{\tau}_\mu$ para o nível de significância de 5%	30
3.2	Percentis de $\hat{\tau}_\lambda$	31
3.3	Percentis de $\hat{\tau}_\lambda$	32
3.4	Percentis de $\hat{\tau}_\lambda$	33
3.5	Percentis de $\hat{\tau}_\lambda$	34
3.6	Taxa de rejeição utilizando $\hat{\lambda}$	35
4.1	Viés dos estimadores no caso $\lambda = 1$	44
4.2	EQM dos estimadores no caso $\lambda = 1$	45
4.3	Viés dos estimadores no caso $\lambda = 5$	46
4.4	EQM dos estimadores no caso $\lambda = 5$	47
4.5	Medidas de posição e dispersão.	50
4.6	Estimativas dos parâmetros do modelo.	51
4.7	Previsões para a série IQA com origem em $n = 80$, utilizando o modelo SINAR(1) ₇ dado por (4.15).	52

1.1 Preliminares

Eventos da natureza, índices econômicos, consequências do desenvolvimento de uma sociedade, entre outros fatos, produzem informações que necessitam de uma análise sofisticada de dados para entender a dinâmica de tais fenômenos e, assim, realizar inferências que possibilitam a tomada de decisões com certa precisão. Em quase todas as áreas do conhecimento, existem fenômenos de interesse que se desenvolvem e variam ao longo do tempo. A análise de dados observados em diferentes instantes de tempo, em geral, conduz a problemas novos que não possuem uma metodologia padrão para serem solucionados. A simples introdução de correlação pela amostragem de pontos adjacentes no tempo pode restringir gravemente a aplicabilidade dos métodos estatísticos mais convencionais que supõem independência das observações. A análise de séries temporais é composta pelo conjunto de ferramentas apropriadas para trabalhar no contexto de observações correlacionadas no tempo.

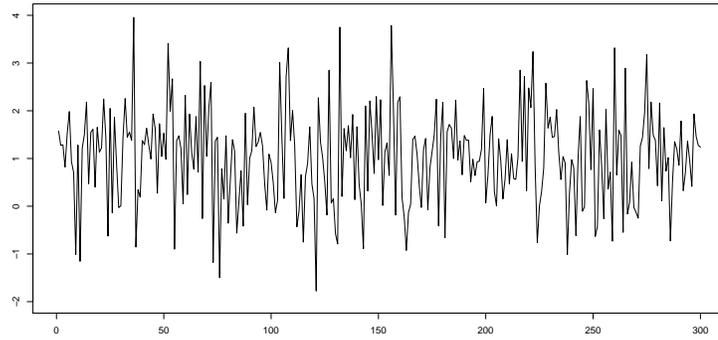
Na literatura, uma série temporal é definida como um conjunto de observações y_t , em que cada uma é coletada em um específico tempo t . Com o intuito de caracterizar a natureza dessas observações, é usual supor que cada valor da série temporal y_t é uma realização de uma variável aleatória Y_t . Isto é, a série temporal y_t é uma realização da família de variáveis aleatórias ou do processo estocástico Y_t . Uma série temporal é dita ser

contínua quando as observações são feitas continuamente no tempo, ou seja, se o conjunto dos tempos em que as observações são feitas, T , é contínuo. Uma série temporal é discreta se o conjunto dos tempos em que as observações são feitas, T , é discreto. Neste trabalho são consideradas séries temporais discretas com $T = \mathbb{Z}$. Note que esses termos não se referem à variável observada y_t , essa pode assumir valores discretos ou contínuos. Quando as variáveis aleatórias que compõem o processo seguem uma distribuição contínua, diz-se que o processo é de marginal contínua como, por exemplo, os processos ARMA (Box, Jenkins & Reinsel (1994)). Se as variáveis aleatórias que compõem o processo seguirem uma distribuição discreta, o processo é dito de marginal discreta e são conhecidos como *processos de contagem* ou *processos de valores inteiros*.

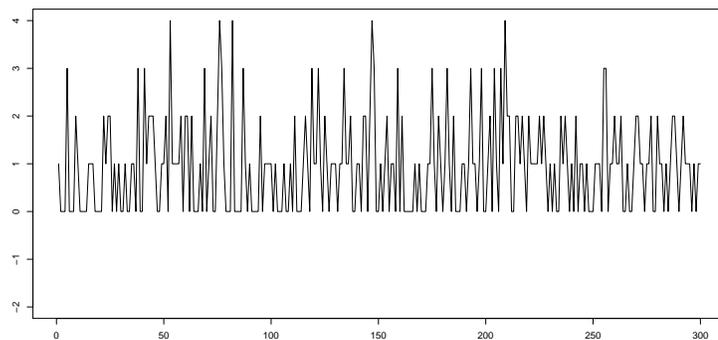
Séries temporais de valores inteiros (processos de contagem) têm chamado a atenção pela importância em aplicações nas diversas áreas do conhecimento, nas quais destacam-se as áreas de Medicina e Economia (Franke & Seligmann (1993), Freeland & McCabe (2004a)). Em ambas áreas, séries temporais de contagem aparecem de forma natural, como por exemplo, na Medicina, o número de pessoas internadas por causas epidemiológicas, o número de internações por causas respiratórias entre outras. Em Economia, é comum estudar os números de ações vendidas em um período de tempo, de patentes registradas, de incidentes causados em navios, de acidentes aéreos, de carros roubados, de passageiros transportados por uma empresa entre outros exemplos que constituem informações cujo interesse ou curiosidade nos motivam a buscar um tratamento estatístico adequado para sua análise.

Apesar do crescente reconhecimento da necessidade de modelar e simular processos de contagem, poucos trabalhos foram publicados sobre esse tipo de processo até o final dos anos setenta. Nas últimas duas décadas, tem havido diversos estudos para desenvolver uma classe de modelos adequados para esse tipo de processo. Na atualidade, o estudo de processos de contagem tem sido uma das áreas de grande destaque em diferentes campos de atuação. Em geral, as pesquisas dedicam-se à formulação e propriedades de modelos (Al-Osh & Alzaid (1987)), estimação (Jung, Ronning & Tremayne (2005)), testes e distribuições assintóticas dos estimadores dos modelos para diferentes distribuições marginais discretas (Freeland & McCabe (2005)), o que resulta em uma linha de pesquisa bastante interessante com impacto científico.

As metodologias desenvolvidas para séries temporais discretas com marginais contínuas estão sendo consideradas para a modelagem de séries temporais de contagem. Os modelos usuais de séries temporais, tais como o processo autorregressivo de média móvel (ARMA), assumem que as marginais são contínuas e, em geral, de distribuição gaussiana. Em alguns casos, os modelos de marginais contínuas podem ser uma aproximação adequada. Porém, essa suposição pode ser inapropriada para modelar séries temporais com marginais discretas. Portanto, surge a necessidade de investigar metodologias apropriadas para séries temporais registradas no tempo discreto, com distribuições marginais discretas.



(a) Realização de processos i.i.d. com distribuição $\mathcal{N}(1, 1)$.



(b) Realização de processos i.i.d. com distribuição $\mathcal{P}o(1)$.

Figura 1.1: Realização de processos independentes e identicamente distribuídos com distribuição $\mathcal{N}(1, 1)$ e $\mathcal{P}o(1)$ ambos com tamanho amostral $n = 300$.

A Figura 1.1 apresenta o comportamento de séries temporais discretas com marginais contínuas e discretas, mais especificamente, em processos com marginal gaussiana e em processos com marginal de Poisson. A notação $\mathcal{N}(1, 1)$ indica que o processo segue distribuição normal com média 1 e variância 1 e a notação $\mathcal{P}o(1)$ indica que o processo segue distribuição de Poisson com parâmetro igual a 1. A Figura 1.1(a) apresenta a realização de um processo independente e identicamente distribuído (i.i.d.) com distribuição gaussiana, com média e variância iguais a 1; a Figura 1.1(b) apresenta a realização de um processo i.i.d. gerado com distribuição de Poisson com parâmetro unitário.

Pela Figura 1.1, observa-se o quanto um processo estocástico de marginal contínua difere de um processo com mesmas média e variância, porém de marginal discreta. Através do exemplo fica ilustrado que, embora os dois processos apresentem a mesma média e variância, o processo de marginal contínua assume valores negativos, o que não ocorre quando se trabalha com dados inteiros. Além disso, em processos de contagem existe a possibilidade de observar um grande número de zeros e valores repetidos. Assim, observamos que essas restrições inviabilizam, em muitas situações, a utilização de modelos com marginais contínuas na modelagem de séries de contagem.

No caso em que as magnitudes dos valores distintos da série temporal com marginal discreta e a variância do processo são grandes, pode-se utilizar séries temporais com marginais contínuas para analisar séries de contagem, como mostra a Figura 1.2.

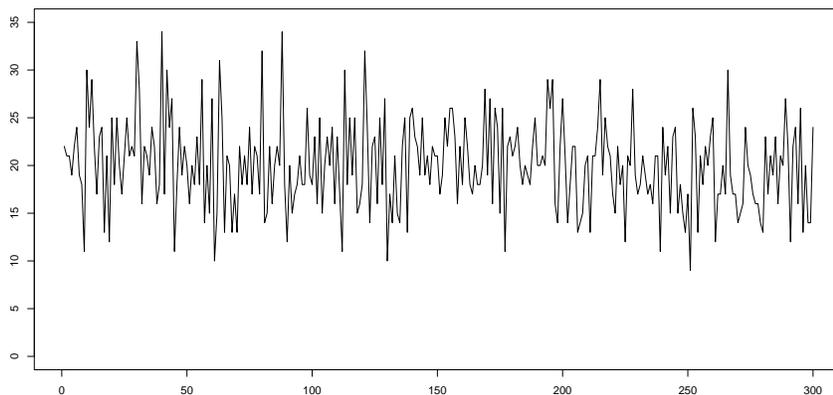


Figura 1.2: Realização de processos independentes e identicamente distribuídos com distribuição $\mathcal{P}o(20)$ com tamanho amostral $n = 300$.

Portanto, a utilização de processos de valores inteiros é adequada para a modelagem de dados discretos cuja magnitude dos valores observados e a variância do processo são pequenas, e que apresentam um grande número de valores repetidos.

Nos últimos anos, têm havido várias tentativas de propor processos que se assemelham à estrutura e propriedades dos usuais modelos lineares estacionários (ARMA). A primeira abordagem foi o modelo discreto autorregressivo e de médias móveis (DARMA) (Jacobs & Lewis (1978a, b)), que é obtido por uma combinação probabilística de uma sequência de variáveis aleatórias independentes e identicamente distribuídas.

Os usuais modelos lineares para séries temporais não são adequados para modelar processos de contagem, pois o produto de uma constante real por uma variável aleatória de valor inteiro pode produzir uma variável aleatória real. A solução encontrada por Al-Osh & Alzaid (1987) para essa dificuldade foi substituir a multiplicação por uma operação chamada *thinning*, proposta por Steutel & Van Harn (1979). Baseados nessa solução, os modelos autorregressivos de primeira ordem para valores inteiros (INAR(1)), apresentados por Al-Osh & Alzaid (1987), representam uma modificação da metodologia usual para analisar dados de contagem. Os processos INAR se assemelham aos usuais modelos lineares estacionários em estrutura e propriedade.

Os processos INAR têm algumas vantagens sobre os processos DARMA (Jacobs & Lewis (1978a, b)). Por exemplo, considere o processo discreto autorregressivo de ordem 1 (DAR(1)), definido por

$$y_t = \begin{cases} y_{t-1}, & \text{com probabilidade } \tau \\ x_t, & \text{com probabilidade } 1 - \tau, \end{cases}$$

em que $\{x_t\}_{t \in \mathbb{Z}}$ é uma sequência de variáveis aleatórias de valores inteiros não-negativos i.i.d. Observa-se que o processo DAR(1) sofre com a desvantagem de que y_t pode ser igual a y_{t-1} . Em contraste, se $y_t = x_t$, a memória do processo antes do tempo t é perdida, o que não ocorre quando se trabalha com o processo INAR(1).

Os processos autorregressivos de valores inteiros propostos por Al-Osh & Alzaid (1987) se assemelham em estrutura e propriedades aos modelos lineares usuais (ARMA). Portanto, o estudo da presença de raízes unitárias e do comportamento sazonal do processo INAR motivam uma vertente de pesquisa de grande interesse para aplicações práticas.

Estacionariedade é uma das suposições mais frequentes que se faz a respeito de uma

série temporal. Entretanto, as aplicações dos testes de raízes unitárias mostraram que essa suposição não é, em geral, satisfeita em séries reais.

O teste de Dickey & Fuller (1979) tem sido um dos mais populares procedimentos para testar a presença de raiz unitária em processos autorregressivos (AR). Na literatura, existem inúmeros artigos e livros publicados com dedicação a essa vertente de investigação (Said & Dickey (1984), Maddala & Kim (1998)). No contexto de processos de contagem, a literatura sobre o estudo de raiz unitária ainda é bastante restrita. Hellström (2001) é a referência pioneira na área. O autor investigou a distribuição empírica da estatística de teste do teste Dickey-Fuller no modelo INAR(1).

Séries temporais que contêm componente sazonal também são bastante comuns em várias situações práticas, especialmente em economia, em sociologia, em ciências e em engenharia. Em séries temporais econômicas, a análise da variação sazonal vem sendo analisada desde o final da década de 20, como por exemplo, nos trabalhos de Mitchell (1927), de Macaulay (1938) e de Burns & Mitchell (1946). Na metodologia Box-Jenkins, o estudo de séries temporais sazonais tem sido bastante explorado em diversos artigos e livros. Por exemplo, Dickey, Hasza & Fuller (1984) propuseram o teste de raízes unitárias sazonais. Trabalhos como Porter-Hudak (1990) e Hassler (1994) apresentam propostas para a estimação do parâmetro de diferenciação sazonal em processos sazonais com a característica de longa dependência.

Ao contrário, o estudo de séries temporais de contagem sazonais não tem recebido muita atenção, até o momento, na literatura. Freeland & McCabe (2004a) ajustaram uma série temporal de contagens mensais do número de requerentes de um benefício por perda salarial causada por acidentes de trabalho. Para a modelagem, Freeland & McCabe (2004a) ajustaram para o conjunto de dados um modelo Poisson autorregressivo com estrutura sazonal, em que a estrutura sazonal era imposta aos erros do processo através de uma função de ligação logarítmica.

Em situações práticas, exemplos de aplicação de séries de contagem com sazonalidade são diversos; por exemplo, o número de automóveis vendidos em um ano, que decai nos meses de agosto e setembro devido à chegada de novos modelos ao mercado, assim como o número de brinquedos vendidos que cresce no mês de dezembro devido às festas natalinas, são exemplos de séries de valores inteiros que têm a componente sazonal.

Devido à importância e necessidade de modelar processos de contagem em diversas áreas e a carência de investigação de processos INAR com estrutura sazonal e de raiz unitária, neste trabalho é proposto um modelo de valores inteiros com estrutura sazonal baseado no modelo apresentado por Al-Osh & Alzaid (1987) e estuda-se o comportamento do teste de Dickey-Fuller em processos autorregressivos de valores inteiros com raiz unitária.

1.2 Objetivos da dissertação

As contribuições deste trabalho se concentram em dois pontos principais. O primeiro se dedica ao estudo do processo autorregressivo de valores inteiros de ordem 1 (INAR(1)), introduzido por Al-Osh & Alzaid (1987), quando o processo apresenta raiz unitária, e tem como objetivo verificar o comportamento do teste de Dickey-Fuller no modelo através do estudo de Monte Carlo.

O segundo consiste em propor o modelo de valores inteiros com estrutura sazonal e verificar, através de estudo de Monte Carlo, o comportamento dos estimadores em diversos cenários.

1.3 Organização da dissertação

Esta dissertação encontra-se dividida em cinco capítulos. No segundo capítulo, descrevemos uma introdução do processo autorregressivo de valores inteiros de ordem 1 (INAR(1)), necessária para o entendimento dos demais capítulos. Além disso, realizamos um levantamento bibliográfico das suas propriedades e das principais metodologias utilizadas para a estimação dos parâmetros do modelo.

No terceiro capítulo, descrevemos o teste de Dickey & Fuller (1979) e verificamos o comportamento do teste em processos autorregressivos de valores inteiros. Comparações entre a utilização do teste de Dickey-Fuller em processos contínuos e processos discretos também são abordadas neste capítulo.

No quarto capítulo, apresentamos a principal contribuição deste trabalho, que é a proposta do processo autorregressivo de valores inteiros com estrutura sazonal (SINAR). Neste capítulo, será exposta a estrutura do modelo proposto, assim como estimadores

para os parâmetros do modelo. Ensaio de Monte Carlo foram realizados para comparar os vícios e os erros quadráticos médios dos estimadores. Como motivação do uso da metodologia proposta, a série do índice da qualidade do ar da cidade de Cariacica-ES, no período de 1 de janeiro de 2007 a 21 de março de 2007, é analisada.

Finalmente, no quinto capítulo, são apresentadas conclusões, comentários e sugestões para futuras pesquisas.

1.4 Suporte computacional

A análise de séries temporais é atualmente impensável sem o auxílio computacional. Por esse motivo, destacamos que nesta dissertação todas as figuras e o processo de programação foram realizadas através do ambiente de programação R, versão 2.10.0 para a plataforma Windows. R é um ambiente integrado que possui grandes facilidades para a manipulação de dados, a geração de gráficos e a modelagem estatística em geral. A linguagem e seus pacotes podem ser obtidos gratuitamente no endereço <http://www.r-project.org>. Mais detalhes podem ser encontrados em Ihaka & Gentleman (1996), Cribari-Neto & Zarkos (1999) e Venables *et al.* (2009).

Processo Autorregressivo de Valores Inteiros

Este capítulo apresenta a forma geral do processo autorregressivo de valores inteiros de ordem 1 (INAR(1)), proposto por Al-Osh & Alzaid (1987), bem como suas propriedades e métodos utilizados para a estimação dos parâmetros do modelo. O estudo do modelo INAR(1) é justificado por ser um dos modelos de interesse na pesquisa proposta por esta dissertação, além de auxiliar no entendimento dos demais capítulos da dissertação.

2.1 Introdução

Recentemente, tem havido um crescente interesse no estudo de séries temporais de valor inteiro não-negativo e, em especial, de séries de contagens. Modelos para dados de contagem são amplamente empregados nas mais diversas áreas de estudo, para a modelagem de diversos fenômenos. Em Controle de Qualidade, por exemplo, usualmente, utiliza-se uma distribuição binomial na modelagem do número de peças defeituosas de uma linha de produção (Montgomery (2004)). Em Epidemiologia, é comum utilizar uma distribuição de Poisson para modelar o número de indivíduos que sofrem de uma determinada doença (Zheng & Basawa (2008)). Fenômenos desse tipo são, geralmente, caracterizados por uma evolução temporal. Para esse tipo de dados, a estrutura temporal

deve ser levada em conta na modelagem.

Vários modelos para processos estacionários com distribuição marginal discreta têm sido propostos (McKenzie (1988)). Um desses modelos, particularmente usado para séries de contagem, é o processo autorregressivo de valores inteiros, denotado por INAR. Nos últimos anos, o modelo INAR(1), proposto por Al-Osh & Alzaid (1987), tem recebido uma importante aceitação pela comunidade científica devido a sua semelhança com os habituais modelos lineares estacionários e por não sofrer das desvantagens do processo DARMA.

Du & Li (1991) generalizaram o processo INAR para ordem p , isto é, INAR(p). Gauthier & Latour (1994) generalizaram o conceito de operador *thinning* permitindo que séries de contagem sigam qualquer distribuição discreta. Du & Li (1991) e Latour (1998) apresentaram a condição de estacionariedade do processo INAR(p). Latour (1997) propôs o INAR multivariado. Utilizando as propriedades do operador *thinning*, Silva & Oliveira (2005) obtiveram as expressões dos momentos e cumulantes do processo INAR(p). Recentes publicações, Subba Rao (1994), Jung & Tremayne (2003) e Jung, Ronning & Tremayne (2005), apresentam teoria, procedimentos e aplicações do processo de contagem. Na sua maioria, as metodologias para modelar séries temporais de contagem são baseadas no domínio do tempo e os erros seguem distribuição de Poisson. Estudos de metodologias com abordagem no domínio da frequência, no contexto de processos autorregressivos de valores inteiros, ainda são bastante limitados. Stoffer (1987) apresenta resultados estatísticos nessa direção.

2.2 O processo INAR

Nesta seção, são resumidas as propriedades do processo autorregressivo de valores inteiros de ordem 1 (INAR(1)) e metodologias para estimação dos parâmetros do modelo. Como o processo INAR(1), apresentado por Al-Osh & Alzaid (1987), baseia-se no operador *thinning* binomial proposto por Steutel & Van Harn (1979). Antes de definir o processo INAR(1), vamos definir abaixo o operador *thinning* binomial.

Definição 2.1. *Seja X uma variável aleatória de valor inteiro não-negativo e p um número real tal que $p \in [0, 1]$. O operador *thinning* binomial, denotado por “ \circ ”, é definido por Steutel & Van Harn (1979) como*

$$p \circ X = \sum_{j=1}^X z_j, \quad (2.1)$$

em que $\{z_j\}_{j=1}^X$, é uma sequência de variáveis aleatórias Bernoulli independentes e identicamente distribuídas (i.i.d), independentes de X , tal que $P(z_j = 1) = 1 - P(z_j = 0) = p$.

Em decorrência da definição do operador *thinning* binomial, temos o seguinte lema.

LEMA 2.1. *Sejam X e Y variáveis aleatórias de valores inteiros não-negativos, e p e q números reais tal que $p \in [0, 1]$ e $q \in [0, 1]$. O operador *thinning* binomial definido em (2.1) apresenta as seguintes propriedades*

1. $0 \circ X = 0$,
2. $1 \circ X = X$,
3. $p \circ (q \circ X) \stackrel{d}{=} (pq) \circ X$, em que a notação $X \stackrel{d}{=} Y$ é utilizada para indicar que X tem a mesma distribuição de Y ,
4. $p \circ (X + Y) \stackrel{d}{=} p \circ X + p \circ Y$,
5. $\mathbb{E}(p \circ X | X) = pX$,
6. $\mathbb{E}(p \circ X) = p\mathbb{E}(X)$,
7. $\text{Var}(p \circ X | X) = p(1 - p)X$,
8. $\text{Var}(p \circ X) = p^2 \text{Var}(X) + p(1 - p)\mathbb{E}(X)$,
9. $\text{Cov}(p \circ X, q \circ Y) = pq \text{Cov}(X, Y)$, se as variáveis z_i incluídas em $p \circ X$ são independentes das variáveis z_j incluídas em $q \circ Y$.

Gauthier & Latour (1994) generalizaram o conceito de operação *thining*, permitindo que as séries de contagem sigam qualquer distribuição discreta. Com a definição do operador *thinning* binomial, podemos definir o modelo autorregressivo de primeira ordem para valores inteiros proposto por Al-Osh & Alzaid (1987).

Definição 2.2. *Um processo estocástico discreto de valor inteiro não-negativo, $\{y_t\}_{t \in \mathbb{Z}}$ diz-se um processo INAR(1) se satisfaz a seguinte equação*

$$y_t = p \circ y_{t-1} + \epsilon_t, \quad (2.2)$$

em que $p \in [0, 1]$, $\{\epsilon_t\}_{t \in \mathbb{Z}}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas de valores inteiros não-negativos, com $\mathbb{E}(\epsilon_t) = \mu$ e $\text{Var}(\epsilon_t) = \sigma^2$. Se $p = 1$ o processo é um passeio aleatório e se $p = 0$ o processo será a sequência $\{\epsilon_t\}_{t \in \mathbb{Z}}$.

Podemos observar pela Definição 2.2 que a realização de $\{y_t\}_{t \in \mathbb{Z}}$ tem duas componentes aleatórias: o total de elementos sobreviventes no tempo $t - 1$, $p \circ y_{t-1}$, cada um com probabilidade de sobrevivência p e os elementos que entram no processo no intervalo $(t - 1, t]$, chamados de elementos de entrada ϵ_t . O modelo definido em (2.2) pode ser interpretado da seguinte maneira

$$\underbrace{y_t}_{\text{População no tempo } t} = \underbrace{p \circ y_{t-1}}_{\text{Sobreviventes do tempo } t-1} + \underbrace{\epsilon_t}_{\text{Imigrantes}}.$$

Nota-se que o processo INAR(1) tem forte apelo interpretativo, visto que seus termos são variáveis aleatórias de uso frequente em aplicações. Com essas interpretações, o processo INAR(1) se aplica em muitas situações práticas. Como exemplo, y_t pode descrever o número de clientes, ϵ_t pode descrever os novos clientes e $y_{t-1} - p \circ y_{t-1}$ pode descrever os clientes que foram perdidos no último período. Brännäs, Hellström & Nordström (2002) utilizaram essa abordagem para modelar o número de noites de hóspedes em hotéis.

Entre as distribuições discretas abordadas na análise de séries temporais, a de Poisson é a mais observada na teoria e na prática (Gallager (2009)). Diante desse fato, ao longo deste trabalho vamos assumir que a sequência $\{\epsilon_t\}_{t \in \mathbb{Z}}$ segue distribuição de Poisson com parâmetro λ , que denotamos por $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$.

2.3 Propriedades

Geralmente, as metodologias utilizadas para modelar séries temporais necessitam de algumas suposições com respeito à estrutura de correlação do processo estocástico suposto. A principal delas é a estacionariedade fraca, nesta pesquisa, por simplicidade, denominada estacionariedade. Essa propriedade estabelece que os momentos de primeira

e segunda ordens do processo não variam sob translações do tempo. Mais precisamente, seja $Y_t, t \in \mathbb{Z}$, um processo estocástico com esperança $\mu_t = \mathbb{E}(Y_t)$ e autocovariância $\gamma^t(k) = Cov(Y_t, Y_{t+k})$, para a defasagem (*lag*) k . O processo Y_t é dito estacionário se $\mu_t = \mu$ e $\gamma^t(k) = \gamma(k)$ são constantes com respeito a $t, \forall k \in \mathbb{Z}$. Brockwell & Davis (1991) e Priestley (1981) são referências importantes sobre esse tipo de processo. Nesse contexto, Du & Li (1991) e Latour (1998) mostraram que, se $p < 1$, o processo INAR(1) definido em (2.2) é estacionário.

A distribuição marginal do processo INAR(1) pode ser expressa em termos da sequência de inovação $\{\epsilon_t\}_{t \in \mathbb{Z}}$ e é dada por (Al-Osh & Alzaid (1987))

$$y_t \stackrel{d}{=} \sum_{j=0}^{\infty} p^j \circ \epsilon_{t-j}. \quad (2.3)$$

Al-Osh & Alzaid (1987) mostraram que a média e a variância do processo INAR(1) são dadas respectivamente por

$$\mathbb{E}(y_t) = p^t \mathbb{E}(y_0) + \mu \sum_{j=0}^{t-1} p^j \quad (2.4)$$

e

$$\text{Var}(y_t) = p^{2t} \text{Var}(y_0) + (1-p) \sum_{j=1}^t p^{2j-1} \mathbb{E}(y_{t-j}) + \sigma^2 \sum_{j=1}^t p^{2(j-1)} \quad (2.5)$$

É fácil ver que, se $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$, a média e a variância do processo são dadas, respectivamente, por

$$\mathbb{E}(y_t) = \frac{\lambda}{1-p} \quad (2.6)$$

e

$$\text{Var}(y_t) = \frac{\lambda}{1-p}. \quad (2.7)$$

McKenzie (1988) mostrou que a distribuição do processo INAR(1), quando $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$, é Poisson com média $\lambda/(1-p)$. Freeland (1998), de maneira simples, fornece uma prova desse mesmo resultado utilizando funções geradoras de probabilidades.

A investigação de séries temporais consiste, essencialmente, das etapas de identificação, estimação e de diagnóstico do ajuste. Para a identificação e a estimação, no contexto de processos estacionários, destacam-se como principais ferramentas as funções

de autocorrelação (ACF) e de autocorrelação parcial (PACF). As funções de autocovariância e de autocorrelação do processo proposto por Al-Osh & Alzaid (1987) são dadas respectivamente por

$$\gamma(k) = \text{Cov}(y_{t-k}, y_t) = p^k \gamma(0), \quad k = 0, 1, 2, \dots \quad (2.8)$$

e

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = p^k, \quad k = 0, 1, 2, \dots \quad (2.9)$$

Como no processo AR(1), a função de autocorrelação decai exponencialmente em função de k . Entretanto, diferentemente dos processos AR, a função de autocorrelação assume somente valores positivos. A Figura 2.1 mostra o gráfico da função de autocorrelação amostral de um processo INAR(1) simulado com $p = 0.9$, $\lambda = 1$ e tamanho amostral $n = 100$.

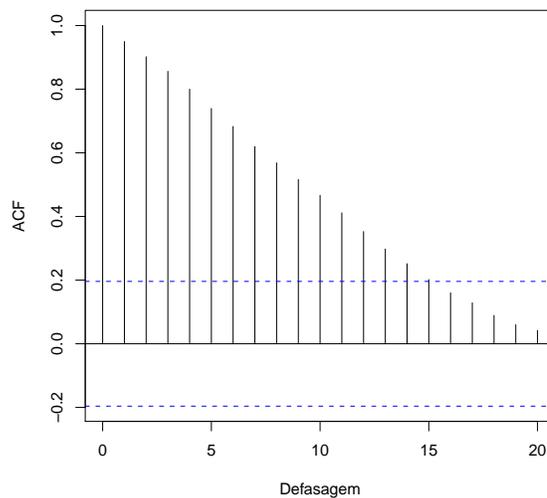


Figura 2.1: Função de autocorrelação amostral do processo INAR(1).

A próxima seção descreve os três métodos de estimação usualmente utilizados para estimar os parâmetros do processo INAR(1) e investiga suas propriedades assintóticas.

2.4 Estimação dos parâmetros

A estimação dos parâmetros do modelo é uma das principais etapas no ajuste de séries temporais. Existem muitos procedimentos sugeridos na literatura com essa finalidade, e o objetivo desta seção é apresentar alguns procedimentos que são utilizados na estimação dos parâmetros do processo INAR(1). Os estimadores discutidos nesta seção têm sido, em geral, propostos no contexto do processo INAR(1) para distribuição marginal de Poisson.

O problema relacionado a estimação dos parâmetros do processo INAR(1) é mais complicado do que para o AR(1), devido ao fato de que a distribuição condicional de y_t dado y_{t-1} no processo INAR(1) é a convolução da distribuição de $\{\epsilon_t\}_{t \in \mathbb{Z}}$ e de uma binomial com parâmetros y_{t-1} e p .

A estimação do vetor de parâmetros do modelo INAR(1) é feita utilizando métodos estatísticos convencionais. Tal como na modelagem de Box-Jenkins, as equações de Yule-Walker podem ser utilizadas para estimar os parâmetros através das autocorrelações amostrais para os diferentes *lags* (Al-Osh & Alzaid (1987)). Há também a alternativa de utilizar o método de mínimos quadrados condicional, assim como a estimação por máxima verossimilhança. Para o método de máxima verossimilhança, as estimativas obtidas pelo método de Yule-Walker podem ser utilizadas como valores iniciais para o procedimento de estimação.

Considerando o modelo INAR(1) como completamente especificado através do vetor de parâmetros $\boldsymbol{\theta} = (p, \lambda)^\top$, a seguir são apresentados os estimadores de Yule-Walker, mínimos quadrados condicional e de máxima verossimilhança condicional.

2.4.1 Estimador de Yule-Walker

Através de uma amostra y_1, \dots, y_n de um processo estacionário $\{y_t\}_{t \in \mathbb{Z}}$, a função de autocorrelação amostral é dada por

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2},$$

em que $\hat{\gamma}(k)$ é a função de autocovariância amostral e $\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t$ é a média aritmética da série.

Como $\rho(k) = p^k$ em (2.9), o estimador de Yule-Walker (YW) para o parâmetro p é dado por

$$\hat{p} = \hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-1} (y_t - \bar{y})(y_{t+1} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}. \quad (2.10)$$

Uma estimativa para λ é baseada no primeiro momento de $\{y_t\}_{t \in \mathbb{Z}}$, que é dado por $\mathbb{E}(y_t) = \lambda/(1-p)$. Assim, um estimador para λ é dado por

$$\hat{\lambda} = (1 - \hat{p}) \frac{1}{n} \sum_{t=1}^n y_t, \quad (2.11)$$

em que \hat{p} é definido em (2.10).

Du *et al.* Li (1991) e Gauthier *et al.* Latour (1994) mostraram que os estimadores usuais da média e das funções de autocovariância e autocorrelação, $\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t$, $\hat{\gamma}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})$ e $\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)}$, $0 \leq k \leq n-1$, respectivamente, são consistentes.

2.4.2 Estimador de mínimos quadrados condicional

O estimador de mínimos quadrados condicional (MQC) proposto por Klimko *et al.* Nelson (1978) consiste em minimizar a soma dos quadrados das distâncias de cada observação, y_t , ao valor esperado condicional dadas as observações anteriores. Al-Osh *et al.* Alzaid (1987), Du *et al.* Li (1991) e Latour (1998) aplicaram esse método de estimação em modelos INAR.

Seja $\{y_t\}_{t \in \mathbb{Z}}$ o processo INAR(1) com $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$. A esperança condicional de y_t dado y_{t-1} é dada por

$$\mathbb{E}[y_t | y_{t-1}] = p y_{t-1} + \lambda = g(\boldsymbol{\theta}, y_{t-1}),$$

em que $\boldsymbol{\theta} = (p, \lambda)^\top$ é o vetor de parâmetros a ser estimado.

Dada uma amostra de tamanho n , o MQC consiste em minimizar a soma dos quadrados

$$Q_n(\boldsymbol{\theta}) = \sum_{t=2}^n [y_t - g(\boldsymbol{\theta}, y_{t-1})]^2. \quad (2.12)$$

Derivando $Q_n(\boldsymbol{\theta})$ em relação a p e λ e igualando as duas equações a zero, obtemos

$$\hat{p} = \frac{\sum_{t=2}^n y_t y_{t-1} - \frac{1}{n-1} \sum_{t=2}^n y_t \sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2 - \frac{1}{n-1} (\sum_{t=2}^n y_{t-1})^2} \quad (2.13)$$

e

$$\widehat{\lambda} = \frac{1}{n-1} \left(\sum_{t=2}^n y_t - \widehat{p} \sum_{t=2}^n y_{t-1} \right). \quad (2.14)$$

Pode-se verificar que as funções $g, \partial g/\partial p, \partial g/\partial \lambda$ e $\partial^2 g/\partial p \partial \lambda$ satisfazem todas as condições de regularidade do Teorema (3.1) de Klimko & Nelson (1978), que são dadas por

- (i) $\lim_{n \rightarrow \infty} \sup_{\delta \rightarrow 0} (|T_n(\boldsymbol{\theta})_{ij}|/n\delta) < \infty, i \leq 2, j \leq 2$, em que $T_n(\boldsymbol{\theta}) = (\partial^2 Q_n(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2 - V_n)$ e $V_n = (\partial^2 Q_n(\boldsymbol{\theta})/\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j)$,
- (ii) $(2n)^{-1} V_n \rightarrow \partial^2 Q_n(\boldsymbol{\theta})/\partial \boldsymbol{\theta}^2$,
- (iii) $n^{-1} \partial Q_n(\boldsymbol{\theta})/\partial \boldsymbol{\theta}_i \rightarrow 0, i \leq 2$,

e, conseqüentemente, os estimadores dados em (2.13) e (2.14) são consistentes. Se $\mathbb{E}(|\epsilon_t|^3) < \infty$, que é o caso de uma distribuição de Poisson, pelo Teorema (3.2) de Klimbo & Nelson (1978), $(\widehat{p}, \widehat{\lambda})^\top$ possuem assintoticamente distribuição normal. Esse resultado é formalizado no seguinte teorema.

TEOREMA 2.1. (Klimko & Nelson (1978)). *Se $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$ em (2.2) e o processo $\{y_t\}_{t \in \mathbb{Z}}$ é estacionário, a distribuição assintótica do estimador de mínimos quadrados condicional de $\boldsymbol{\theta}$ é dada por*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}(0, V^{-1} W V^{-1}),$$

em que $\widehat{\boldsymbol{\theta}} = (\widehat{p}, \widehat{\lambda})^\top$, $\boldsymbol{\theta} = (p, \lambda)^\top$ e V é uma matriz 2×2 com os elementos

$$V_{ij} = \mathbb{E} \left(\frac{\partial g(\boldsymbol{\theta}, y_{t-1})}{\partial \boldsymbol{\theta}_i} \cdot \frac{\partial g(\boldsymbol{\theta}, y_{t-1})}{\partial \boldsymbol{\theta}_j} \right) \quad i, j = 1, 2,$$

e os elementos da matriz W são da forma

$$W_{ij} = \mathbb{E} \left(u_t^2(\boldsymbol{\theta}) \frac{\partial g(\boldsymbol{\theta}, y_{t-1})}{\partial \boldsymbol{\theta}_i} \cdot \frac{\partial g(\boldsymbol{\theta}, y_{t-1})}{\partial \boldsymbol{\theta}_j} \right) \quad i, j = 1, 2,$$

com $u_t(\boldsymbol{\theta}) = y_t - g(\boldsymbol{\theta}, y_{t-1})$ e a notação \xrightarrow{d} indica convergência em distribuição.

COROLÁRIO 2.1. (Freeland & McCabe (2005)). Se $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$ em (2.2) e o processo $\{y_t\}_{t \in \mathbb{Z}}$ é estacionário, então a distribuição assintótica do estimador de mínimos quadrados condicional de $\boldsymbol{\theta} = (p, \lambda)^\top$ é dada por

$$\sqrt{n} \begin{pmatrix} \widehat{p} - p \\ \widehat{\lambda} - \lambda \end{pmatrix} \xrightarrow{d} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{p(1-p)^2}{\lambda} + (1-p)^2 & -(1+p)\lambda \\ -(1+p)\lambda & \lambda + \frac{1+p}{1-p}\lambda^2 \end{pmatrix} \right).$$

Observa-se, para n grande, que o estimador de MQC de $\boldsymbol{\theta} = (p, \lambda)^\top$, dados em (2.13) e (2.14), respectivamente, e o estimador de YW para $\boldsymbol{\theta} = (p, \lambda)^\top$, dados em (2.10) e (2.11), respectivamente, são equivalentes.

2.4.3 Estimador de máxima verossimilhança condicional

Al-Osh & Alzaid (1987) propõem duas alternativas para a obtenção das estimativas dos parâmetros do modelo INAR(1) através da maximização da função de verossimilhança: estimadores condicionais e não condicionais.

Para o processo INAR(1) com $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$, a distribuição condicional de y_t dado y_{t-1} , denotada por $f(y_t|y_{t-1})$, é a convolução de uma distribuição binomial resultante do operador *thinning* binomial, $p \circ y_{t-1}$ e de uma distribuição de Poisson resultante do processo de chegada $\{\epsilon_t\}_{t \in \mathbb{Z}}$ (Sprott (1983)). Assim, se $*$ denota a convolução e

$$f_1(i) = \binom{y_{t-1}}{i} p^i (1-p)^{(y_{t-1})-i}, \quad f_2(i) = \frac{\lambda^i}{i!} e^{-\lambda},$$

então

$$\begin{aligned} f(y_t|y_{t-1}) &= f_1 * f_2 = \sum_{i=0}^{\infty} f_1(i) f_2(y_t - i) \\ &= e^{-\lambda} \sum_{i=0}^{\min(y_t, y_{t-1})} \frac{\lambda^{(y_t)-i}}{((y_t) - i)!} \binom{y_{t-1}}{i} p^i (1-p)^{(y_{t-1})-i}. \end{aligned}$$

Sendo $\mathbf{y} = (y_0, \dots, y_n)$ uma amostra de $n+1$ observações, a função de verossimilhança incondicional é dada por

$$L(p, \lambda; \mathbf{y}) = \frac{[\lambda/(1-p)]^{y_0}}{y_0!} \exp\left(\frac{\lambda}{1-p}\right) \prod_{t=1}^n f(y_t|y_{t-1}),$$

e a função de verossimilhança condicional do processo INAR(1), dado y_0 , é definida como

$$L(p, \lambda; \mathbf{y}|y_0) = \prod_{t=1}^n f(y_t|y_{t-1}).$$

A função de log-verossimilhança condicional é da forma

$$\ell(p, \lambda; y|y_0) = \log \left(\prod_{t=1}^n f(y_t|y_{t-1}) \right) = \sum_{t=1}^n \log(f(y_t|y_{t-1})). \quad (2.15)$$

A função escore condicional é dada por

$$\begin{aligned} \frac{\partial \ell(p, \lambda; y|y_0)}{\partial p} &= \sum_{t=1}^n \frac{y_{t-1} [f(y_t - 1|y_{t-1} - 1) - f(y_t|y_{t-1})]}{1 - p} \frac{1}{f(y_t|y_{t-1})} \\ \frac{\partial \ell(p, \lambda; y|y_0)}{\partial \lambda} &= \sum_{t=1}^n \frac{[f(y_t - 1|y_{t-1}) - f(y_t|y_{t-1})]}{f(y_t|y_{t-1})}. \end{aligned}$$

O estimador de máxima verossimilhança condicional (MVC) de $\boldsymbol{\theta} = (p, \lambda)^\top$ é o valor $\hat{\boldsymbol{\theta}} = (\hat{p}, \hat{\lambda})^\top$ que maximiza a função de verossimilhança $L(p, \lambda; y|y_0)$. O valor de $\boldsymbol{\theta}$ que maximiza a função verossimilhança condicional $L(p, \lambda; y|y_0)$, também maximiza $\ell(p, \lambda; y|y_0)$. Assim, os estimadores de máxima verossimilhança condicional de p e λ são obtidos através da maximização do logaritmo da função de verossimilhança condicional definido em (2.15) recorrendo, em geral, a métodos numéricos, uma vez que as equações de log-verossimilhança conduzem-nos a um sistema de equações não-lineares. A distribuição assintótica do estimador MVC para os parâmetros do modelo INAR(1) é estabelecida no seguinte teorema.

TEOREMA 2.2. (*Franke & Seligmann (1993)*). *Os estimadores de máxima verossimilhança condicional de p e λ tem a seguinte distribuição assintótica*

$$\sqrt{n} \begin{pmatrix} \hat{p} - p \\ \hat{\lambda} - \lambda \end{pmatrix} \xrightarrow{d} \mathcal{N}(0, \mathbf{K}^{-1}),$$

em que a matriz \mathbf{K} é a matriz de informação de Fisher, isto é, $-\mathbf{K}$ é a esperança das segundas derivadas do logaritmo da função de verossimilhança condicional.

Um estudo empírico foi realizado em Al-Osh & Alzaid (1987) para determinar qual o melhor entre os métodos de estimação apresentados quando a sequência de inovações $\{\epsilon_t\}_{t \in \mathbb{Z}}$ são provenientes da distribuição de Poisson. O resultado obtido demonstra que o método de máxima verossimilhança condicional é aquele que fornece as melhores propriedades para os estimadores.

2.5 Previsão

Um dos principais objetivos da análise de séries temporais é fazer previsões a partir do modelo estimado. Estudos relacionados com previsão de modelos INAR apareceram recentemente na literatura em Freeland & McCabe (2004b) e Brännäs (1994) para o modelo INAR(1).

Considere a previsão de um valor futuro y_{n+h} , dado que temos observado a série até o momento n . Assumindo que o processo definido em (2.2) é estacionário, após sucessivas substituições podemos escrever y_{n+h} da seguinte maneira

$$y_{n+h} = p^h \circ y_n + \sum_{j=1}^h p^{h-j} \circ \epsilon_{n+j}, \quad h = 1, 2, 3, \dots, \quad (2.16)$$

Partindo-se da origem em n , e supondo que o objetivo é prever a série em um período futuro h , \hat{y}_{n+h} representa a previsão para um período $n+h$ feita em n . O valor de y_{n+h} , que não se conhece no tempo n , é substituído pela sua previsão \hat{y}_{n+h} . A previsão que apresenta o erro quadrático médio mínimo é a esperança condicional de y_{n+h} , dadas as observações passadas da série, que no caso do INAR(1) é dada por

$$\mathbb{E}[y_{n+h}|y_1, \dots, y_n] = p^h \left(y_n - \frac{\lambda}{1-p} \right) + \frac{\lambda}{1-p}, \quad h = 1, 2, 3, \dots$$

Logo um estimador para a previsão y_{n+h} é dado por

$$\hat{y}_{n+h} = \hat{p}^h \left(y_n - \frac{\hat{\lambda}}{1-\hat{p}} \right) + \frac{\hat{\lambda}}{1-\hat{p}}, \quad h = 1, 2, 3, \dots, \quad (2.17)$$

onde o erro de previsão é dado por $e_{n+h} = y_{n+h} - \hat{y}_{n+h}$ e é fácil verificar que $\mathbb{E}(e_{n+h}) = 0$.

Podemos observar que as previsões dadas por (2.17) não vão gerar valores inteiros. A fim de obter previsões coerentes para y_{n+h} , Freeland & McCabe (2004b) sugerem a utilização do valor que minimiza o erro esperado absoluto dada a amostra, isto é, minimizar $\mathbb{E}(|y_{n+h} - \tilde{y}_{n+h}| | y_1, \dots, y_n)$, em que \tilde{y}_{n+h} é a mediana h -passos à frente da distribuição condicional $f(y_{n+h} | y_1, \dots, y_n)$.

Diante do que foi exposto neste capítulo, os próximos capítulos mostrarão as contribuições deste trabalho no contexto de séries de contagem com dependência temporal.

Teste de Dickey-Fuller em processos INAR

O objetivo deste capítulo é resumir alguns conceitos básicos a respeito do teste de Dickey & Fuller (1979) e verificar a utilização do teste no contexto de processos autorregressivos de valores inteiros.

3.1 Introdução

Na modelagem Box-Jenkins, o número de diferenças necessárias para que uma série se torne estacionária é conhecido como ordem de integração da série. Os testes de raízes unitárias são capazes, em geral, de detectar se a série foi suficientemente diferenciada para se tornar estacionária. Para tanto, testa-se a hipótese nula de que a série não é estacionária, ou seja, possui raiz unitária, contra a alternativa de que a série é estacionária. Os testes clássicos mais utilizados para se testar essa hipótese foram, inicialmente, estudados por Fuller (1976) e Dickey & Fuller (1979). Outros trabalhos foram publicados posteriormente, como os testes apresentados por Said & Dickey (1984) e Phillips & Perron (1988). A próxima seção descreve o teste de Dickey & Fuller (1979), que será o teste utilizado neste trabalho.

3.2 Raízes unitárias e o teste de Dickey-Fuller

Embora a teoria exposta nesta seção seja diminuta comparada ao que se tem disponível na literatura sobre o assunto, descreveremos de forma concisa, porém sem prejuízo do rigor matemático, a metodologia para compreensão do teste de Dickey-Fuller (DF) em processos autorregressivos, que será necessária para a aplicação do teste em processos autorregressivos de valores inteiros.

Suponha que $\{a_t\}_{t \in \mathbb{Z}}$ é um processo ruído branco com média 0 e variância σ^2 . Um processo $\{x_t\}_{t \in \mathbb{Z}}$ é chamado de processo autoregressivo de ordem p (AR(p)) se

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + a_t$$

ou

$$(1 - \phi_1 B - \cdots - \phi_p B^p)x_t = \phi(B)x_t = a_t,$$

em que $\phi(B)$ é o polinômio autorregressivo e B é o operador de translação para o passado ($B^k x_t = x_{t-k}$). Se $\phi^{-1}(B)$ é estável (somável) x_t é estacionária. Pode-se mostrar que o processo AR(p) será estacionário se $\phi^{-1}(B)$ convergir para $|B| \leq 1$ (Box, Jenkins & Reinsel (1994)).

Sejam m_i^{-1} , $i = 1, \dots, p$, as raízes da equação característica $\phi(B) = 0$; então podemos escrever

$$\phi(B) = (1 - m_1 B)(1 - m_2 B) \dots (1 - m_p B)$$

e, expandindo em frações parciais, temos que

$$\phi^{-1}(B) = \sum_{i=1}^p \frac{A_i}{1 - m_i B}.$$

Se $\phi^{-1}(B)$ convergir para $|B| \leq 1$ devemos ter que $|m_i| < 1$, $i = 1, \dots, p$. Esta condição é equivalente a que a equação característica $\phi(B) = 0$ tenha raízes fora do círculo unitário. Ou seja, para o processo AR(p) ser estacionário, as raízes de $\phi(B) = 0$ devem ser menores que um em valor absoluto.

Nesse contexto, para verificar se o processo é estacionário, verificamos se o processo possui raiz unitária, e esta é a razão (interesse) do estudo de raízes unitárias em processos autorregressivos.

Em processos autorregressivos e de médias móveis (ARMA), o problema de raiz unitária aparece quando o polinômio autorregressivo apresenta pelo menos uma raiz sobre o círculo unitário. Isso implica que se deve tomar um número de diferenças apropriado da série original antes de ajustar o modelo (Moretтин & Toloι (2004)). Pode-se ter raízes unitárias também no polinômio de médias móveis, indicando que os dados foram superdiferenciados.

Os testes de raízes unitárias mais simples e amplamente utilizados foram desenvolvidos por Fuller (1976) e Dickey & Fuller (1979). O teste DF tem como hipótese nula a existência de uma raiz unitária, ou seja, $\mathcal{H}_0 : d = 1$, em que d é o número de diferenças necessárias para tornar a série estacionária.

Seja $\{x_t\}_{t \in \mathbb{Z}}$ um processo univariado autorregressivo de ordem 1 (AR(1)), ou seja, definido como

$$x_t = \phi x_{t-1} + a_t, \quad (3.1)$$

em que $\{a_t\}_{t \in \mathbb{Z}}$ é um processo ruído branco com média 0 e variância σ^2 e $\phi \in \mathbb{R}$. Se $\phi = 1$, o processo $\{x_t\}_{t \in \mathbb{Z}}$ não é estacionário e é conhecido como passeio aleatório. Em contraste, se $|\phi| < 1$ o processo $\{x_t\}_{t \in \mathbb{Z}}$ é estacionário.

Considere o modelo definido em (3.1) e que queremos testar a hipótese $\mathcal{H}_0 : \phi = 1$ contra $\mathcal{H}_1 : \phi < 1$. Através de uma amostra x_1, \dots, x_n o teste DF é baseado no estimador de mínimos quadrados

$$\hat{\phi} = \frac{\sum_{t=2}^n x_{t-1} x_t}{\sum_{t=2}^n x_{t-1}^2}. \quad (3.2)$$

Assim, para testar $\mathcal{H}_0 : \phi = 1$ contra $\mathcal{H}_1 : \phi < 1$ no modelo definido em (3.1), utiliza-se a estatística proposta por Dickey & Fuller (1979)

$$\hat{\tau} = \frac{\hat{\phi} - 1}{s(\hat{\phi})}, \quad (3.3)$$

em que

$$s(\hat{\phi}) = \frac{S}{(\sum_{t=2}^n x_{t-1}^2)^{1/2}},$$

com

$$S^2 = \frac{1}{n-2} \sum_{t=2}^n (x_t - \hat{\phi} x_{t-1})^2.$$

A quantidade $\hat{\phi}$ é o estimador de mínimos quadrados, $s(\hat{\phi})$ é o erro padrão de $\hat{\phi}$ e S^2 é o estimador de σ^2 (mais detalhes em Said & Dickey (1984) e Chang & Park (2003)).

Assintoticamente, quando $|\phi| < 1$, temos que $\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{d} \mathcal{N}(0, (1 - \phi^2))$. Segundo Dickey & Fuller (1979), a distribuição da estatística definida em (3.3), sob \mathcal{H}_0 , é tal que

$$\hat{\tau} \xrightarrow{d} \frac{\frac{1}{2}([W(1)]^2 - 1)}{\left(\int_0^1 [W(r)]^2 dr\right)^{1/2}}, \quad (3.4)$$

onde $W(r)$ é o movimento browniano padrão, ou seja, para cada t , $W(t) \sim \mathcal{N}(0, t)$.

Caso o processo $\{x_t\}_{t \in \mathbb{Z}}$ seja definido da forma $x_t = \mu + \phi(x_{t-1} - \mu) + a_t$, com $\mu \neq 0$, tem-se

$$\hat{\tau}_\mu = \frac{\hat{\phi}_\mu - 1}{s(\hat{\phi}_\mu)} \quad (3.5)$$

e

$$\hat{\tau}_\mu \xrightarrow{d} \frac{\frac{1}{2}([W(1)]^2 - 1) - W(1) \int_0^1 W(r) dr}{\left[\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r) dr\right)^2\right]^{1/2}}.$$

Nesse contexto, a distribuição da estatística de teste afasta-se mais da distribuição normal do que no caso em que $\mu = 0$. As distribuições dessas estatísticas estão tabeladas em Fuller (1976).

Phillips (1987) e Phillips & Perron (1988) generalizaram os resultados de Dickey & Fuller (1979) para o caso em que os erros $\{a_t\}_{t \in \mathbb{Z}}$ são correlacionados e, possivelmente, heterocedásticos; mais detalhes em, por exemplo, Hamilton (1994).

Outros testes de raízes unitárias são também bastante utilizados, como, por exemplo, o teste de Dickey-Fuller aumentado (ADF) (Said & Dickey (1984) e Phillips & Perron (1998)). Kwiatkowski, Phillips, Schmidt & Shin [KPSS] (1992) propõem um teste de estacionariedade cuja hipótese alternativa considera a raiz unitária. Para um estudo desses testes, ver os Capítulos 3 e 4 de Maddala & Kim (1998).

3.3 Teste de Dickey-Fuller em processos INAR

Nesta seção, será apresentado o teste de Dickey-Fuller aplicado em processos INAR, ou seja, em vez de usar a metodologia descrita na seção anterior para processos AR(1), vamos utilizá-la no processo INAR(1).

Como vimos no Capítulo 1, o processo INAR(1) pode ser interpretado como um processo de nascimento-morte, e uma raiz unitária no modelo implica que indivíduos que entraram têm uma vida infinitamente longa e que a população (o número de empresas, número de pessoas, etc.) vai se tornar infinitamente grande.

Considere o modelo definido em (2.2) e que queremos testar a hipótese $\mathcal{H}_0 : p = 1$ contra $\mathcal{H}_1 : p < 1$. Na seção anterior, observamos que através de uma amostra y_1, \dots, y_n o teste DF é baseado no estimador de mínimos quadrados. No caso do processo INAR(1), em vez de utilizarmos o estimador de mínimos quadrados no teste DF, vamos utilizar o estimador de mínimos quadrados condicional definido no capítulo anterior.

Assim, para testar $\mathcal{H}_0 : p = 1$ contra $\mathcal{H}_1 : p < 1$ no modelo definido em (2.2), utilizaremos a estatística

$$\hat{\tau}_\lambda = \frac{\hat{p} - 1}{s(\hat{p})}, \quad (3.6)$$

em que

$$s(\hat{p}) = \frac{S}{[\sum_{t=2}^n (y_{t-1} - \bar{y})^2]^{1/2}}$$

e

$$S^2 = \frac{1}{n-2} \sum_{t=2}^n (y_t - \hat{\lambda} - \hat{p}y_{t-1})^2,$$

com

$$\hat{p} = \frac{\sum_{t=2}^n y_t y_{t-1} - \frac{1}{n-1} \sum_{t=2}^n y_t \sum_{t=2}^n y_{t-1}}{\sum_{t=2}^n y_{t-1}^2 - \frac{1}{n-1} (\sum_{t=2}^n y_{t-1})^2}$$

e

$$\hat{\lambda} = \frac{1}{n-1} \left(\sum_{t=2}^n y_t - \hat{p} \sum_{t=2}^n y_{t-1} \right).$$

De acordo com o Corolário 2.1, quando $p < 1$ e $n \rightarrow \infty$, temos que $\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, (p(1-p)^2/\lambda + (1-p^2)))$. Diante desse resultado, podemos observar que a distribuição, sob \mathcal{H}_1 , da estatística de teste do teste de Dickey-Fuller utilizado em processos autorregressivos de valores inteiros é diferente quando comparada com a distribuição da

estatística de teste do teste DF utilizada em processos autorregressivos usuais. O Teorema 3.1 abaixo, é uma contribuição do trabalho. O Teorema 3.1 mostra, sob certas condições, que a distribuição da estatística do teste DF tem a mesma distribuição quando é aplicada em processos INAR(1) e AR(1).

TEOREMA 3.1. *Seja $\{y_t\}_{t \in \mathbb{Z}}$ o processo INAR(1) com $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$ e $\{x_t\}_{t \in \mathbb{Z}}$ o processo AR(1) com $\{a_t\}_{t \in \mathbb{Z}} \sim \mathcal{N}(\lambda, \lambda)$. Seja \hat{p} o estimador de mínimos quadrados condicional de p e $\hat{\phi}$ o estimador de mínimos quadrados de ϕ . Se $p < 1$ e $|\phi| < 1$ em (2.2) e (3.1), respectivamente, tem-se que*

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, (1 - p^2)), \text{ se } n, \lambda \rightarrow \infty.$$

O Teorema 3.1 mostra que, assintoticamente, com relação a λ , a estatística do teste de Dickey-Fuller, sob \mathcal{H}_1 , possui a mesma distribuição assintótica quando é aplicada em processos autorregressivos usuais e processos autoregressivos de valores inteiros. Este é um resultado esperado, uma vez que quando $\lambda \rightarrow \infty$, a distribuição de Poisson tende para uma distribuição normal.

A próxima seção apresenta um estudo de Monte Carlo com o intuito de verificar o funcionamento do teste DF em processos autorregressivos de valores inteiros e comparar o teste entre os processos com marginais contínuas e discretas.

3.4 Resultados numéricos

Esta seção se dedica à investigação empírica do teste DF com objetivo de verificar a distribuição do teste em séries temporais de contagem com marginal de Poisson, mais especificamente em processos INAR(1) com a hipótese nula de raiz unitária. No estudo, investiga-se o comportamento do teste sob diferentes tamanhos amostrais e valores do parâmetro λ do modelo INAR(1).

Para o estudo empírico, amostras foram simuladas dos modelos AR(1) e INAR(1), isto é, $x_t = \lambda + px_{t-1} + a_t$ e $y_t = p \circ y_{t-1} + \epsilon_t$, respectivamente. Nos modelos, $\{a_t\}_{t \in \mathbb{Z}}$ é um processo i.i.d $\mathcal{N}(0, \lambda)$ e o processo $\{\epsilon_t\}_{t \in \mathbb{Z}}$ é um processo i.i.d com distribuição de Poisson com parâmetro λ . Considerar a mesma variância em ambos os processos permite também o estudo do efeito desse parâmetro no desempenho do teste (Hellström (2001)).

Foram utilizadas amostras de tamanho $n = 50, 100, 250$ e 500 . Em todos os testes foram consideradas as hipóteses $\mathcal{H}_0 : p = 1$ contra $\mathcal{H}_1 : p < 1$. Foram realizadas 10000 replicações de Monte Carlo para gerar a distribuição empírica dos testes.

As próximas figuras mostram o comportamento da distribuição empírica das estatísticas $\hat{\tau}_\mu$ e $\hat{\tau}_\lambda$ definidas em 3.5 e 3.6, além da normal padrão. Essas figuras elucidam o comportamento dos testes e contribuem para uma melhor compreensão dos resultados apresentados nas tabelas.

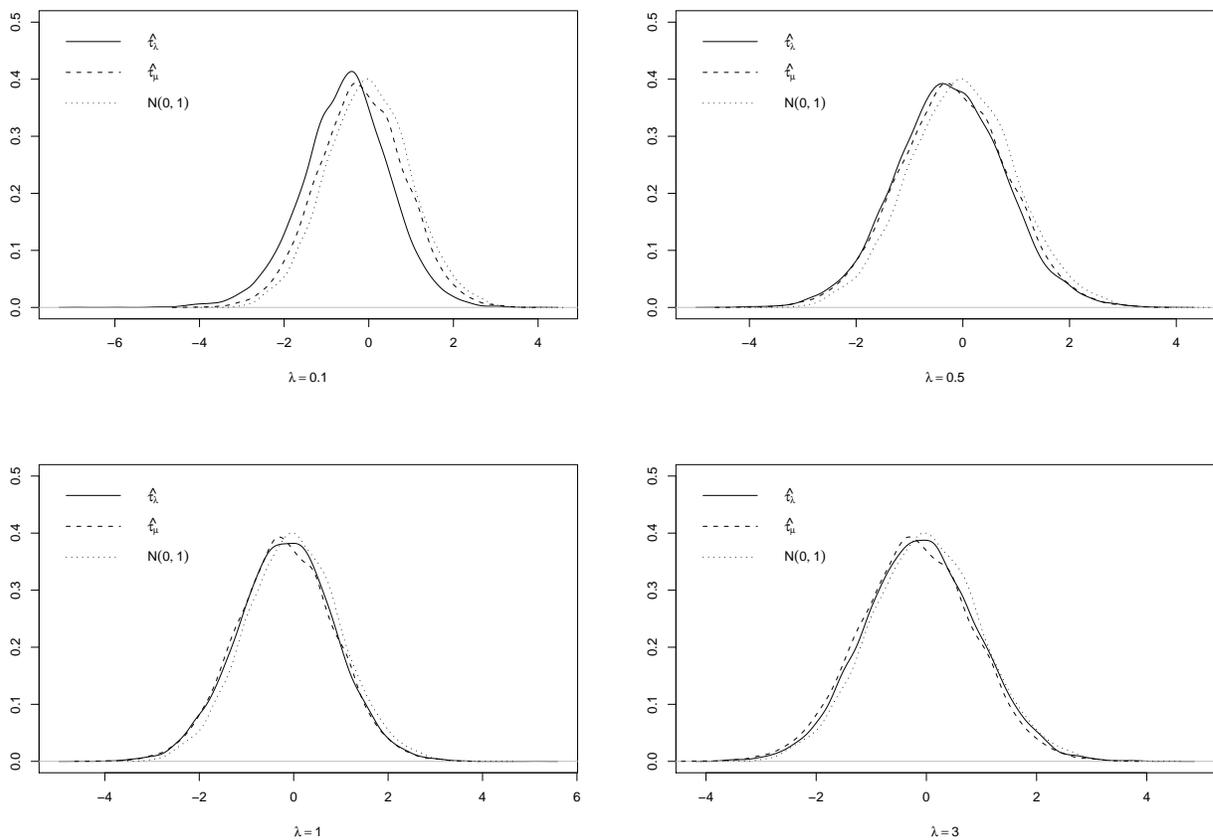


Figura 3.1: Densidade empírica dos testes sob \mathcal{H}_0 , $n = 100$ e normal padrão.

Através da Figura 3.1, percebe-se que a distribuição de $\hat{\tau}_\lambda$ se desloca para à direita lentamente à medida que o valor do parâmetro λ aumenta. Além disso, podemos ver que a distribuição de $\hat{\tau}_\mu$ não se modifica à medida que o valor do parâmetro λ aumenta. O deslocamento da distribuição de $\hat{\tau}_\lambda$ pode ser verificado com maior facilidade nas Figuras 3.2, 3.3, 3.4 e 3.5.

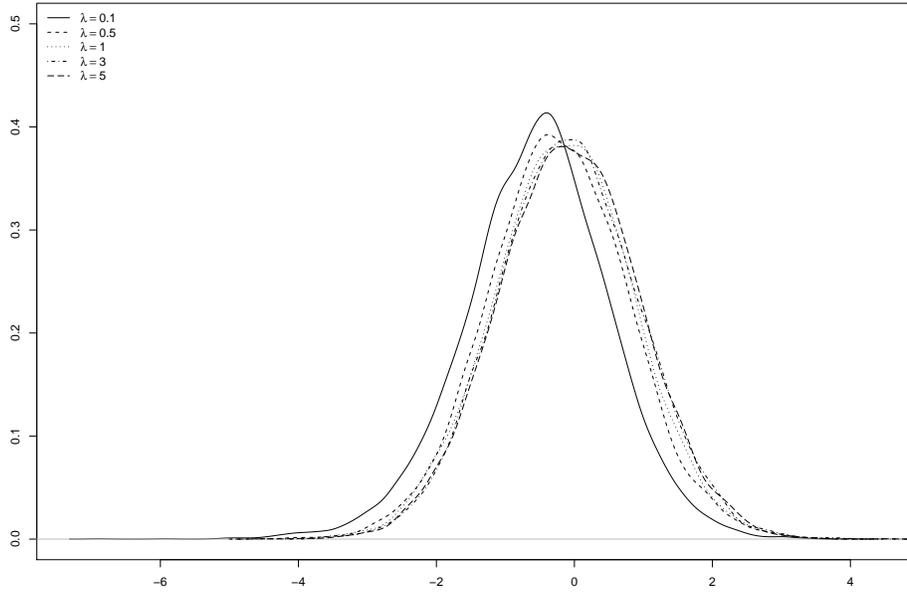


Figura 3.2: Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $n = 100$ com λ variando.

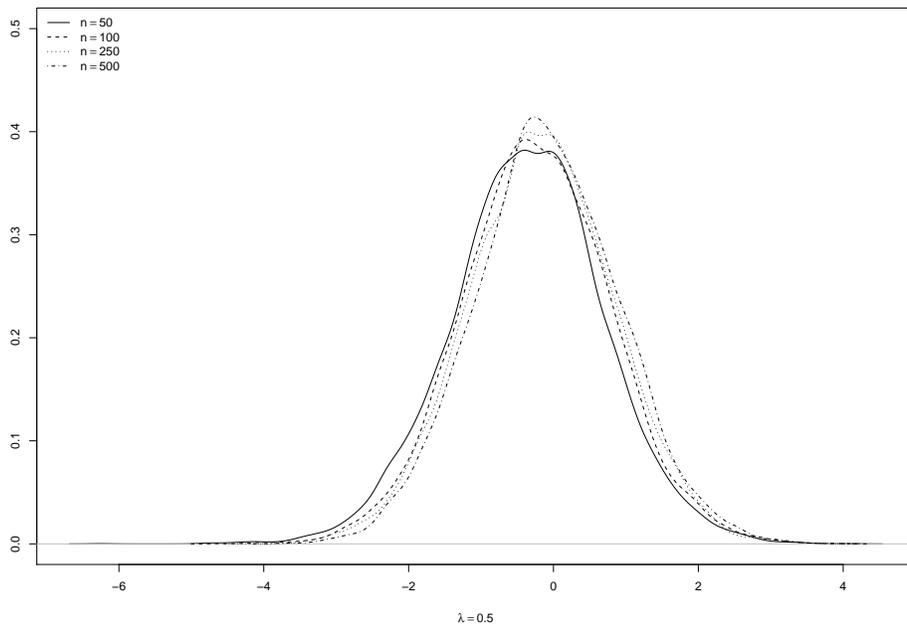


Figura 3.3: Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $\lambda = 0.5$ com n variando.

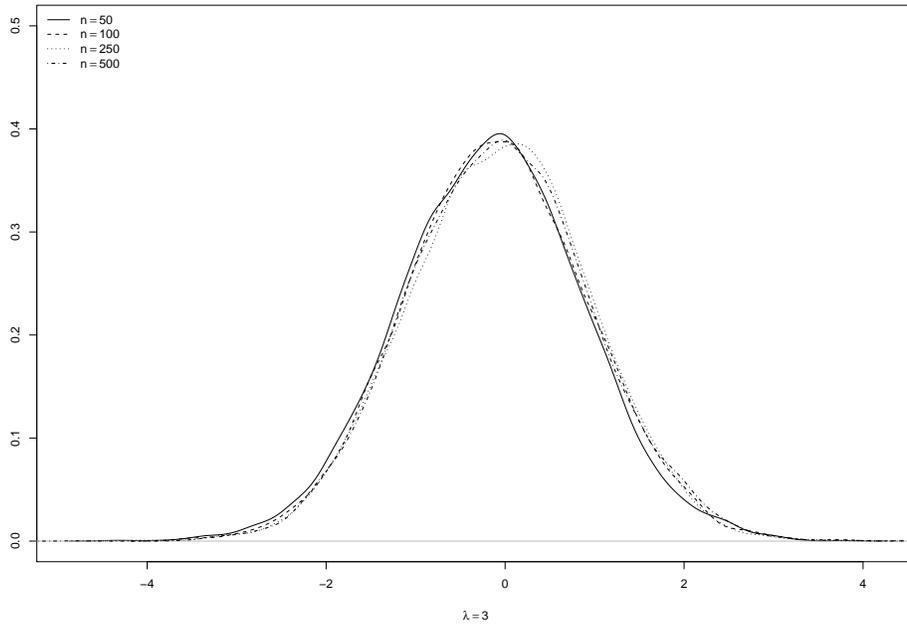


Figura 3.4: Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $\lambda = 3$ com n variando.

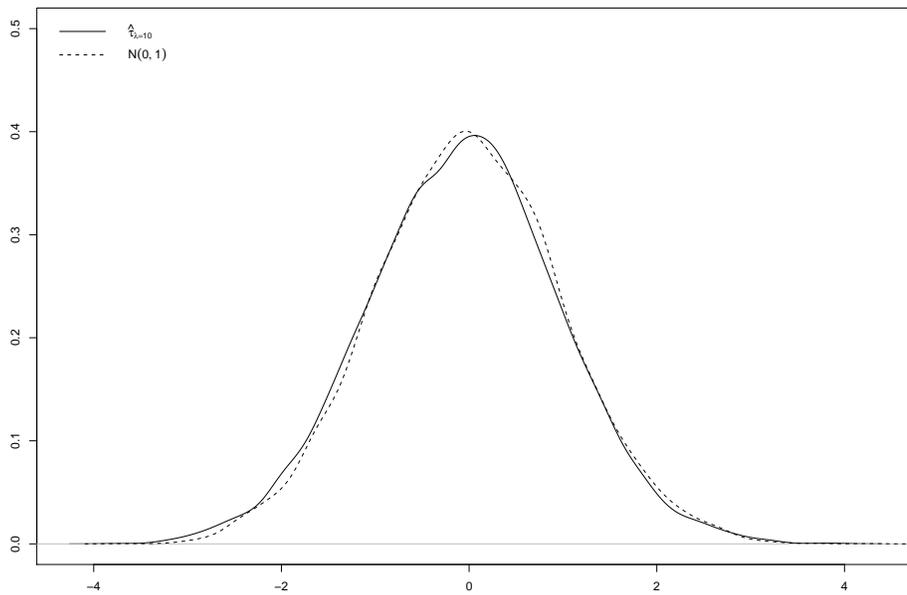


Figura 3.5: Densidade empírica de $\hat{\tau}_\lambda$ sob \mathcal{H}_0 , $\lambda = 10$ e da normal padrão.

Diante das Figuras 3.1, 3.2, 3.3, 3.4 e 3.5, observamos que o deslocamento da distribuição de $\hat{\tau}_\lambda$ também acontece à medida que o tamanho da amostra aumenta (Figuras 3.3 e 3.4). Note que para $\lambda = 10$ (Figura 3.5), a distribuição de $\hat{\tau}_\lambda$ é bastante próxima da distribuição normal padrão. A Tabela 3.1 apresenta os valores críticos de $\hat{\tau}_\lambda$ e $\hat{\tau}_\mu$ para o nível de significância de 5% e para $\lambda = 0.1, 0.5, 1, 3, 5$ e 10. Vale ressaltar que os valores críticos obtidos para $\hat{\tau}_\mu$ foram encontrados pela densidade empírica de $\hat{\tau}_\mu$.

Tabela 3.1: Valores críticos de $\hat{\tau}_\lambda$ e $\hat{\tau}_\mu$ para o nível de significância de 5%

λ	n	$\hat{\tau}_\lambda$	$\hat{\tau}_\mu$
0.1	50	-2.29	-1.93
	100	-2.30	-1.85
	400	-1.92	-1.77
	700	-1.87	-1.72
0.5	50	-2.16	-1.95
	100	-1.89	-1.83
	400	-1.81	-1.77
	700	-1.73	-1.72
1	50	-1.95	-1.94
	100	-1.87	-1.84
	400	-1.73	-1.76
	700	-1.70	-1.72
3	50	-1.84	-1.95
	100	-1.75	-1.86
	400	-1.71	-1.77
	700	-1.65	-1.72
5	50	-1.86	-1.95
	100	-1.77	-1.85
	400	-1.68	-1.77
	700	-1.70	-1.72
10	50	-1.82	-1.93
	100	-1.75	-1.85
	400	-1.69	-1.77
	700	-1.66	-1.72
$\mathcal{N}(0, 1)$		-1.65	-1.65

Pela Tabela 3.1, podemos ver que os percentis da distribuição de $\hat{\tau}_\lambda$ mudam à medida que λ aumenta, enquanto que os percentis da distribuição $\hat{\tau}_\mu$ permanecem praticamente inalterados à medida que λ aumenta. Os resultados sugerem que a maioria dos quantis diferem mais para valores pequenos de λ . Podemos também observar que quando

aumentamos os valores de n e/ou λ , a distribuição $\hat{\tau}_\lambda$ tende à distribuição normal padrão.

Diante dos resultados empíricos apresentados até o momento, nota-se que a utilização do teste de Dickey-Fuller usual, em processos INAR(1), pode resultar em conclusões errôneas, uma vez que a distribuição do teste DF, sob \mathcal{H}_0 , aplicado em processos INAR(1), é diferente da distribuição do teste DF, sob \mathcal{H}_0 , aplicado em processos AR(1).

Como já mencionado, é bastante comum na prática considerar a metodologia para séries temporais com marginais contínuas em séries temporais de contagem, ou seja, geralmente o tratamento desse tipo de dado é realizado como se o processo fosse de marginal contínua (Box, Jenkins & Reinsel (1994)) e podemos ver que nesse caso não distinguir os processos pode resultar em conclusões erradas.

Para valores de n e/ou λ suficientemente grandes, os resultados empíricos mostram que a distribuição do teste DF aplicado em processos INAR se aproxima da distribuição $\mathcal{N}(0, 1)$. Logo, podemos utilizar os percentis da distribuição $\mathcal{N}(0, 1)$ na aplicação do teste em processos INAR. Entretanto, quando temos um processo onde n e/ou λ não são grandes, como podemos utilizar o teste DF em processos INAR? Nesse contexto, assim como Dickey & Fuller (1979), tabelamos abaixo os pontos críticos do teste DF, aplicado em processo INAR(1), para vários valores do percentil α e vários valores do parâmetro λ .

Tabela 3.2: Percentis de $\hat{\tau}_\lambda$

		Percentis								
λ	n	0.01	0.025	0.05	0.10	0.50	0.9	0.95	0.975	0.99
0.1	50	-3.52	-3.01	-2.29	-2.07	-0.64	0.61	1.00	1.35	1.70
	100	-3.23	-2.69	-2.30	-1.86	-0.51	0.74	1.13	1.47	1.89
	250	-2.73	-2.33	-2.01	-1.61	-0.33	0.95	1.34	1.67	2.09
	500	-2.67	-2.27	-1.92	-1.53	-0.24	1.02	1.41	1.71	2.13
0.2	50	-3.31	-2.73	-2.30	-1.88	-0.54	0.72	1.10	1.48	1.92
	100	-2.86	-2.43	-2.10	-1.71	-0.41	0.87	1.24	1.62	2.05
	250	-2.70	-2.29	-1.95	-1.56	-0.25	1.07	1.42	1.76	2.15
	500	-2.52	-2.13	-1.83	-1.46	-0.19	1.13	1.49	1.83	2.21
$\mathcal{N}(0, 1)$		-2.33	-1.96	-1.65	-1.28	0	1.28	1.65	1.96	2.33

Tabela 3.3: Percentis de $\hat{\tau}_\lambda$

		Percentis								
λ	n	0.01	0.025	0.05	0.10	0.50	0.9	0.95	0.975	0.99
0.3	50	-3.02	-2.54	-2.21	-1.75	-0.43	0.88	1.26	1.62	2.00
	100	-2.79	-2.36	-2.00	-1.63	-0.32	0.96	1.31	1.63	2.06
	250	-2.66	-2.22	-1.87	-1.49	-0.18	1.11	1.44	1.79	2.20
	500	-2.53	-2.13	-1.83	-1.41	-0.12	1.15	1.52	1.84	2.14
0.4	50	-2.95	-2.53	-2.13	-1.70	-0.39	0.90	1.29	1.66	2.11
	100	-2.72	-2.32	-2.00	-1.60	-0.27	0.99	1.36	1.69	2.14
	250	-2.57	-2.19	-1.84	-1.45	-0.18	1.11	1.48	1.79	2.13
	500	-2.51	-2.11	-1.77	-1.41	-0.16	1.19	1.52	1.87	2.27
0.5	50	-2.89	-2.44	-2.16	-1.69	-0.35	0.95	1.34	1.70	2.12
	100	-2.66	-2.24	-1.89	-1.53	-0.25	1.05	1.43	1.82	2.23
	250	-2.57	-2.15	-1.84	-1.46	-0.17	1.11	1.49	1.79	2.13
	500	-2.45	-2.03	-1.73	-1.37	-0.10	1.20	1.56	1.92	2.32
0.6	50	-2.82	-2.37	-2.01	-1.61	-0.30	1.01	1.38	1.71	2.13
	100	-2.58	-2.22	-1.91	-1.52	-0.23	1.09	1.45	1.75	2.19
	250	-2.53	-2.17	-1.81	-1.43	-0.13	1.15	1.53	1.86	2.17
	500	-2.41	-2.06	-1.74	-1.40	-0.09	1.18	1.57	1.90	2.21
0.7	50	-2.82	-2.38	-2.02	-1.64	-0.28	1.01	1.37	1.71	2.09
	100	-2.62	-2.21	-1.85	-1.50	-0.19	1.11	1.48	1.80	2.20
	250	-2.52	-2.13	-1.79	-1.42	-0.14	1.15	1.52	1.86	2.19
	500	-2.42	-2.05	-1.74	-1.41	-0.10	1.19	1.57	1.91	2.29
0.8	50	-2.74	-2.33	-1.98	-1.57	-0.27	1.04	1.42	1.76	2.16
	100	-2.60	-2.21	-1.87	-1.49	-0.18	1.10	1.49	1.83	2.20
	250	-2.51	-2.11	-1.81	-1.42	-0.12	1.18	1.53	1.86	2.22
	500	-2.41	-2.05	-1.75	-1.38	-0.08	1.21	1.55	1.90	2.28
0.9	50	-2.73	-2.29	-1.95	-1.59	-0.26	1.09	1.47	1.80	2.21
	100	-2.53	-2.20	-1.87	-1.47	-0.17	1.10	1.48	1.80	2.19
	250	-2.51	-2.11	-1.78	-1.40	-0.10	1.20	1.55	1.87	2.25
	500	-2.40	-2.03	-1.74	-1.36	-0.09	1.21	1.57	1.88	2.27
$\mathcal{N}(0, 1)$		-2.33	-1.96	-1.65	-1.28	0	1.28	1.65	1.96	2.33

Tabela 3.4: Percentis de $\widehat{\tau}_\lambda$

		Percentis								
λ	n	0.01	0.025	0.05	0.10	0.50	0.9	0.95	0.975	0.99
1	50	-2.76	-2.33	-1.95	-1.56	-0.23	1.04	1.44	1.77	2.15
	100	-2.54	-2.17	-1.87	-1.47	-0.16	1.12	1.51	1.84	2.25
	250	-2.45	-2.07	-1.77	-1.40	-0.10	1.18	1.50	1.83	2.15
	500	-2.43	-2.05	-1.75	-1.35	-0.07	1.25	1.61	1.92	2.31
2	50	-2.65	-2.20	-1.87	-1.47	-0.17	1.10	1.51	1.85	2.26
	100	-2.48	-2.11	-1.77	-1.42	-0.13	1.18	1.55	1.87	2.24
	250	-2.44	-2.05	-1.73	-1.38	-0.08	1.19	1.55	1.855	2.21
	500	-2.45	-2.06	-1.69	-1.33	-0.07	1.23	1.63	1.96	2.35
3	50	-2.60	-2.17	-1.84	-1.46	-0.14	1.14	1.52	1.88	2.31
	100	-2.46	-2.07	-1.75	-1.40	-0.11	1.22	1.60	1.93	2.27
	250	-2.42	-2.04	-1.74	-1.38	-0.05	1.23	1.58	1.89	2.23
	500	-2.36	-2.03	-1.72	-1.36	-0.07	1.24	1.64	1.97	2.32
4	50	-2.59	-2.16	-1.84	-1.46	-0.11	1.19	1.55	1.87	2.35
	100	-2.45	-2.09	-1.77	-1.39	-0.07	1.23	1.59	1.93	2.32
	250	-2.46	-2.04	-1.72	-1.33	-0.07	1.20	1.59	1.94	2.35
	500	-2.41	-2.05	-1.71	-1.33	-0.06	1.24	1.61	1.98	2.39
5	50	-2.65	-2.23	-1.86	-1.46	-0.13	1.18	1.56	1.87	2.24
	100	-2.44	-2.08	-1.77	-1.39	-0.08	1.22	1.59	1.93	2.29
	250	-2.41	-2.04	-1.72	-1.36	-0.05	1.25	1.62	1.96	2.33
	500	-2.37	-2.00	-1.69	-1.33	-0.04	1.24	1.60	1.93	2.36
6	50	-2.64	-2.20	-1.84	-1.45	-0.11	1.21	1.59	1.92	2.37
	100	-2.44	-2.07	-1.74	-1.38	-0.08	1.22	1.58	1.94	2.28
	250	-2.42	-2.05	-1.71	-1.34	-0.06	1.27	1.61	1.95	2.27
	500	-2.35	-2.01	-1.70	-1.33	-0.04	1.24	1.61	1.95	2.33
7	50	-2.52	-2.16	-1.81	-1.41	-0.10	1.19	1.56	1.88	2.27
	100	-2.50	-2.11	-1.75	-1.38	-0.07	1.21	1.61	1.97	2.32
	250	-2.40	-2.03	-1.70	-1.32	-0.06	1.26	1.64	1.93	2.26
	500	-2.37	-2.01	-1.67	-1.32	-0.04	1.25	1.64	1.95	2.33
$\mathcal{N}(0, 1)$		-2.33	-1.96	-1.65	-1.28	0	1.28	1.65	1.96	2.33

Tabela 3.5: Percentis de $\hat{\tau}_\lambda$

		Percentis								
λ	n	0.01	0.025	0.05	0.10	0.50	0.9	0.95	0.975	0.99
8	50	-2.55	-2.16	-1.81	-1.39	-0.10	1.21	1.57	1.90	2.27
	100	-2.53	-2.09	-1.76	-1.36	-0.06	1.24	1.57	1.92	2.35
	250	-2.37	-2.00	-1.72	-1.34	-0.05	1.26	1.62	1.93	2.29
	500	-2.38	-2.01	-1.69	-1.31	-0.04	1.25	1.64	1.93	2.33
9	50	-2.48	-2.12	-1.83	-1.41	-0.08	1.21	1.58	1.89	2.33
	100	-2.49	-2.07	-1.74	-1.37	-0.05	1.25	1.61	1.94	2.40
	250	-2.35	-2.00	-1.70	-1.33	-0.05	1.27	1.63	1.94	2.29
	500	-2.37	-2.01	-1.68	-1.32	-0.04	1.25	1.64	1.95	2.37
10	50	-2.49	-2.13	-1.82	-1.41	-0.08	1.21	1.58	1.91	2.29
	100	-2.49	-2.06	-1.75	-1.37	-0.04	1.25	1.61	1.94	2.41
	250	-2.35	-2.02	-1.69	-1.32	-0.04	1.28	1.63	1.94	2.28
	500	-2.37	-2.02	-1.68	-1.32	-0.04	1.26	1.63	1.95	2.37
$\mathcal{N}(0, 1)$		-2.33	-1.96	-1.65	-1.28	0	1.28	1.65	1.96	2.33

As Tabelas 3.2, 3.3, 3.4 e 3.5 mostram os pontos críticos do teste DF em processos INAR(1). Observa-se que o teste depende do valor de λ , que na prática é desconhecido. Vale resaltar que Hellström (2001) considera uma regressão onde a variável resposta são os percentis da distribuição $\hat{\tau}_\lambda$, em vez de tabelar os pontos críticos. Entretanto, a regressão também depende do parâmetro desconhecido λ .

É plausível que o primeiro pensamento que vem à cabeça do leitor nesse momento é o de utilizar $\hat{\lambda}$ no lugar de λ , para assim poder desfrutar dos pontos críticos das tabelas acima. Entretanto, utilizando $\hat{\lambda}$ em vez de λ também poderíamos estar distorcendo o resultado do teste, uma vez que o estimador de λ pode ser um valor bem distante do valor real de λ .

A Tabela 3.6 mostra a distorção da taxa de rejeição do teste quando utilizamos $\hat{\lambda}$ em vez de λ na utilização das Tabelas 3.2, 3.3, 3.4 e 3.5.

Tabela 3.6: Taxa de rejeição utilizando $\hat{\lambda}$

λ	n	$\alpha = 0.05$	$\alpha = 0.10$
0.1	100	0.0960	0.1588
	250	0.0637	0.1165
	500	0.0590	0.1118
0.5	100	0.0617	0.1184
	250	0.0558	0.1021
	500	0.0554	0.0980
0.9	100	0.0576	0.1161
	250	0.0514	0.0982
	500	0.0497	0.0994
1	100	0.0537	0.1075
	250	0.0521	0.1032
	500	0.0477	0.1010
3	100	0.0491	0.1025
	250	0.0483	0.1004
	500	0.0493	0.0913
5	100	0.0517	0.0996
	250	0.0529	0.0931
	500	0.0513	0.1010

Nota-se na Tabela 3.6 que a taxa de rejeição do teste, utilizando $\hat{\lambda}$ em vez de λ , ficou próxima do que era esperado para valores maiores de λ . Para tamanho de amostra grande observa-se que a diferença é mínima e que a maior diferença encontrada é quando o valor de λ é pequeno. Portanto, em muitos casos, os percentis tabelados em 3.2, 3.3, 3.4 e 3.5 podem ser utilizados, fazendo o uso de $\hat{\lambda}$, quando o intuito é utilizar o teste de Dickey-Fuller em processos INAR.

Processo Autorregressivo de Valores Inteiros com Estrutura Sazonal

O objetivo deste capítulo é introduzir um modelo autorregressivo de valores inteiros com estrutura sazonal e estudar suas propriedades, além de comparar, através do estudo de simulações, diferentes estimadores para os parâmetros do modelo proposto. Além disso, aplicaremos a metodologia em um conjunto de dados reais. As provas de algumas proposições mais importantes podem ser encontradas no apêndice desta dissertação.

4.1 Introdução

Uma série temporal exhibe o fenômeno sazonal quando um determinado comportamento se repete depois de um período regular de tempo. O menor período de tempo, dentro de um ano, da repetição desse fenômeno é denominado período sazonal. Assim, quando as observações de uma série temporal que estão deslocadas “ s ” unidades de tempo exibem um comportamento similar, dizemos que a série apresenta sazonalidade com período s . Séries temporais sazonais, em geral, exibem intervalos de tempo de 1 mês e período $s = 12$ meses. No entanto, ocorrem casos com período $s = 4$ trimestres como em dados trimestrais, ou com $s = 7$ dias em dados diários (Box, Jenkins & Reinsel (1994)).

Um exemplo de dados sazonais são os preços dos produtos agrícolas. Esse tipo de produto tem sua oferta geralmente relacionada ao ciclo biológico das culturas agrícolas ou

das criações de animais, que ocorrem segundo um período de tempo determinado. Outro exemplo é o de séries relacionadas à demanda de produtos e serviços utilizados durante o calendário de festividades de um país, como por exemplo, o natal.

Como exemplos de modelos sazonais, pode-se citar Barsky & Miron (1989), que consideram um modelo sazonal determinístico; Hansen & Sargent (1993), que propõem um modelo sazonal periódico com periodicidade não alterada e Canova (1992), que considera um modelo sazonal como a soma de um processo determinístico com um processo estocástico estacionário. Outra forma de tratamento é o modelo sazonal não estacionário, que permite raízes unitárias sazonais (Box, Jenkins & Reinsel (1994)).

Em geral, não é fácil saber a priori qual aproximação produz a melhor descrição estatística dos dados e, sabendo que a sazonalidade é quantitativamente importante em muitas séries, a imposição de um tipo de sazonalidade quando outra está presente pode resultar em sérios vieses ou perda de informações. Portanto, é importante estabelecer qual tipo de sazonalidade está presente.

Na próxima seção, apresentaremos o processo autorregressivo de valores inteiros com estrutura sazonal (SINAR).

4.2 O processo SINAR: definição e propriedades

Nesta seção, são resumidas propriedades básicas do processo autorregressivo de valores inteiros de ordem 1 e período sazonal s (SINAR(1) $_s$) definido abaixo.

Definição 4.1. *Um processo estocástico discreto de valores inteiros não-negativos, $\{y_t\}_{t \in \mathbb{Z}}$, diz-se um processo sazonal autorregressivo de valores inteiros de ordem 1 e período sazonal s (SINAR(1) $_s$) se satisfaz a seguinte equação*

$$y_t = P \circ y_{t-s} + \epsilon_t, \tag{4.1}$$

onde $P \in [0, 1]$, $\{\epsilon_t\}_{t \in \mathbb{Z}}$ é uma sequência de variáveis aleatórias independentes e identicamente distribuídas de valores inteiros não-negativos, com $\mathbb{E}(\epsilon_t) = \mu$, $\text{Var}(\epsilon_t) = \sigma^2$ e $s \in \mathbb{N}$ é o período sazonal.

Assim como no processo INAR(1), a realização de $\{y_t\}_{t \in \mathbb{Z}}$ em (4.1) tem duas componentes aleatórias: o total de elementos sobreviventes no tempo $t - s$, $P \circ y_{t-s}$, cada um

com probabilidade de sobrevivência P , e os elementos que entram no processo no intervalo $(t-s, t]$, chamados de elementos de entrada ϵ_t . Além disso a condição de estacionariedade do processo $\text{SINAR}(1)_s$ é equivalente à do processo $\text{INAR}(1)$. Ou seja, se $P < 1$ o processo autorregressivo de valores inteiros de ordem 1 e período sazonal s é estacionário.

Pelos mesmos motivos do Capítulo 2, ao longo deste capítulo assumiremos que a sequência $\{\epsilon_t\}_{t \in \mathbb{Z}}$ segue distribuição de Poisson com parâmetro λ , ou seja, $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{Po}(\lambda)$.

A média e a variância do processo definido em (4.1) são dadas respectivamente por

$$\mathbb{E}(y_t) = \frac{\lambda}{1-P} \quad (4.2)$$

e

$$\text{Var}(y_t) = \frac{\lambda}{1-P}. \quad (4.3)$$

A distribuição marginal do modelo definido em (4.1) pode ser expressa em termos do processo de inovação $\{\epsilon_t\}_{t \in \mathbb{Z}}$ como

$$y_t \stackrel{d}{=} \sum_{j=0}^{\infty} P^j \circ \epsilon_{t-sj}. \quad (4.4)$$

A Figura 4.1 mostra o comportamento de um processo $\text{SINAR}(1)_{12}$ simulado com $P = 0.9$, $\lambda = 1$ e tamanho amostral $n = 100$. Observando o gráfico percebe-se que a série simulada apresenta um comportamento sazonal.

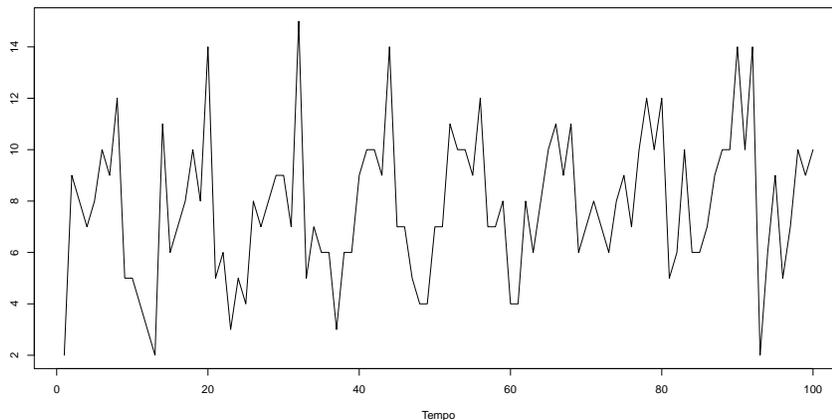


Figura 4.1: Série simulada de um processo $\text{SINAR}(1)_{12}$, com $P = 0.9$, $\lambda = 1$ e tamanho amostral $n = 100$.

PROPOSIÇÃO 4.1. *Seja $\{y_t\}_{t \in \mathbb{Z}}$ o processo $\text{SINAR}(1)_s$ e $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$. A função de autocovariância do processo é dada por*

$$\gamma(k) = \begin{cases} \gamma(0)P^{k/s}, & \text{se } k \text{ é múltiplo de } s, \\ 0, & \text{caso contrário,} \end{cases} \quad (4.5)$$

em que $\gamma(0)$ é a variância do processo $\{y_t\}_{t \in \mathbb{Z}}$.

De acordo com a Proposição 4.1, a função de autocorrelação do processo é definida como

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = \begin{cases} P^{k/s}, & \text{se } k \text{ é múltiplo de } s, \\ 0, & \text{caso contrário.} \end{cases} \quad (4.6)$$

Assim, a função de autocorrelação será não nula somente nos *lags* múltiplos de s . A Figura 4.2 mostra o gráfico da função de autocorrelação amostral de um processo $\text{SINAR}(1)_{12}$ com $P = 0.9$, $\lambda = 1$ e tamanho amostral $n = 100$. Nota-se, pelo gráfico, que o comportamento da sazonalidade é bem visível quando observamos a função de autocorrelação amostral.

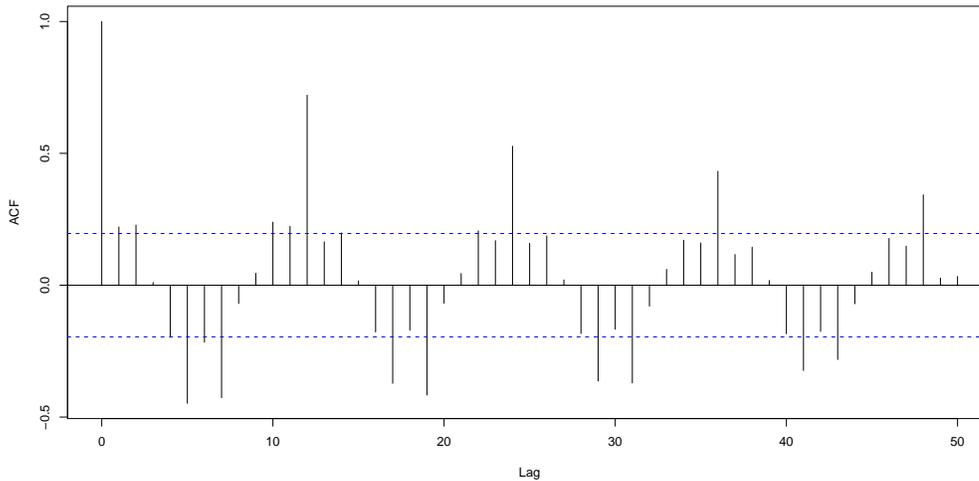


Figura 4.2: Função de autocorrelação amostral do processo $\text{SINAR}(1)_{12}$.

As próximas seções descrevem três métodos de estimação para os parâmetros do processo proposto: estimador de Yule-Walker, mínimos quadrados condicional e máxima verossimilhança condicional.

4.2.1 Estimador de Yule-Walker

Através de uma amostra y_1, \dots, y_n de um processo estacionário $\{y_t\}_{t \in \mathbb{Z}}$, a função de autocorrelação amostral é dada por

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2},$$

em que $\hat{\gamma}(k)$ é a função de autocovariância amostral e $\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t$ é a média aritmética da série.

Como $\rho(k) = P^{k/s}$ em (4.6), o estimador de Yule-Walker (YW) para o parâmetro P é dado por

$$\hat{P} = \hat{\rho}(s) = \frac{\hat{\gamma}(s)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{n-s} (y_t - \bar{y})(y_{t+s} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2}. \quad (4.7)$$

De maneira análoga ao processo INAR(1), uma estimativa para λ é baseada no primeiro momento de $\{y_t\}_{t \in \mathbb{Z}}$, que é dado por $\mathbb{E}(y_t) = \lambda/(1 - P)$. Assim, um estimador para λ é dado por

$$\hat{\lambda} = (1 - \hat{P}) \frac{1}{n} \sum_{t=1}^n y_t, \quad (4.8)$$

onde \hat{P} é dado em (4.7).

4.2.2 Estimador de mínimos quadrados condicional

Seja $\{y_t\}_{t \in \mathbb{Z}}$ o processo definido em (4.1) com $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$. A esperança condicional de y_t dado y_{t-s} é dada por

$$\mathbb{E}[y_t | y_{t-s}] = P y_{t-s} + \lambda = g(\Theta, y_{t-s}),$$

onde $\Theta = (P, \lambda)^\top$ é o vetor de parâmetros a ser estimado.

Dada uma amostra de y_1, \dots, y_n , a estimação dos parâmetros P e λ , através do método de MQC, consiste em minimizar a soma dos quadrados

$$Q_n(\Theta) = \sum_{t=s+1}^n [y_t - g(\Theta, y_{t-s})]^2. \quad (4.9)$$

Derivando $Q_n(\Theta)$ em relação a P e λ e igualando as duas equações a zero, obtemos

$$\hat{P} = \frac{\sum_{t=s+1}^n y_t y_{t-s} - \frac{1}{n-s-1} \sum_{t=s+1}^n y_t \sum_{t=s+1}^n y_{t-s}}{\sum_{t=s+1}^n y_{t-s}^2 - \frac{1}{n-s-1} \left(\sum_{t=s+1}^n y_{t-s} \right)^2} \quad (4.10)$$

e

$$\hat{\lambda} = \frac{1}{n-s-1} \left(\sum_{t=s+1}^n y_t - \hat{P} \sum_{t=s+1}^n y_{t-s} \right). \quad (4.11)$$

Pode-se verificar que as funções $g, \partial g / \partial P, \partial g / \partial \lambda$ e $\partial^2 g / \partial P \partial \lambda$ satisfazem todas as condições de regularidade do teorema (3.1) de Klimko & Nelson (1978). Consequentemente, os estimadores dados em (4.10) e (4.11) são consistentes.

PROPOSIÇÃO 4.2. *Se $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$ em (4.1) e o processo $\{y_t\}_{t \in \mathbb{Z}}$ é estacionário, então a distribuição assintótica do estimador de mínimos quadrados condicional de $\Theta = (P, \lambda)^\top$ é dada por*

$$\sqrt{n} \begin{pmatrix} \hat{P} - P \\ \hat{\lambda} - \lambda \end{pmatrix} \xrightarrow{d} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{P(1-P)^2}{\lambda} + (1-P^2) & -(1+P)\lambda \\ -(1+P)\lambda & \lambda + \frac{1+P}{1-P} \lambda^2 \end{pmatrix} \right).$$

4.2.3 Estimador de máxima verossimilhança condicional

Para o processo $\text{SINAR}(1)_s$ definido em (4.1) com $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{P}o(\lambda)$, a distribuição condicional de y_t dado y_{t-s} , denotada por $f(y_t | y_{t-s})$, é a convolução de uma distribuição binomial resultada do operador *thinning* binomial, $P \circ y_{t-s}$, e de uma distribuição de Poisson resultante do processo de chegada $\{\epsilon_t\}_{t \in \mathbb{Z}}$. Assim, se $*$ denota a convolução e

$$f_1(i) = \binom{y_{t-s}}{i} P^i (1-P)^{(y_{t-s})-i}, \quad f_2(i) = \frac{\lambda^i}{i!} e^{-\lambda},$$

então

$$\begin{aligned}
f(y_t|y_{t-s}) &= f_1 * f_2 = \sum_{i=0}^{\infty} f_1(i)f_2(y_t - i) \\
&= e^{-\lambda} \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{\lambda^{(y_t)-i}}{((y_t) - i)!} \binom{y_{t-s}}{i} P^i (1 - P)^{(y_{t-s})-i}.
\end{aligned}$$

Sendo $y = (y_0, \dots, y_n)$ uma amostra de $n+1$ observações, a função de verossimilhança condicional (MVC) do processo $\text{SINAR}(1)_s$, dados y_0, \dots, y_{s-1} , é definida como

$$L(P, \lambda; y|y_0, \dots, y_{s-1}) = \prod_{t=s}^n f(y_t|y_{t-s}).$$

A função de log-verossimilhança condicional é da forma

$$\ell(P, \lambda; y|y_0, \dots, y_{s-1}) = \log \left(\prod_{t=s}^n f(y_t|y_{t-s}) \right) = \sum_{t=s}^n \log(f(y_t|y_{t-s})). \quad (4.12)$$

A função escore condicional é dada por

$$\begin{aligned}
\frac{\partial \ell(P, \lambda; y|y_0, \dots, y_{s-1})}{\partial P} &= \sum_{t=s}^n \frac{y_{t-s}}{1 - P} \frac{[f(y_t - 1|y_{t-s} - 1) - f(y_t|y_{t-s})]}{f(y_t|y_{t-s})} \\
\frac{\partial \ell(P, \lambda; y|y_0, \dots, y_{s-1})}{\partial \lambda} &= \sum_{t=s}^n \frac{[f(y_t - 1|y_{t-s}) - f(y_t|y_{t-s})]}{f(y_t|y_{t-s})}.
\end{aligned}$$

O estimador de máxima verossimilhança condicional de $\Theta = (P, \lambda)^\top$ é o valor $\hat{\Theta} = (\hat{P}, \hat{\lambda})^\top$ que maximiza a função de log-verossimilhança $\ell(y, P, \lambda|y_0, \dots, y_{s-1})$. Assim, os estimadores de máxima verossimilhança condicional de P e λ são obtidos através da maximização do logaritmo da função de verossimilhança condicional definida em (4.12).

4.3 Previsão

Nesta seção, vamos considerar os cálculos das previsões para o processo sazonal de valores inteiros definido em (4.1).

Considere a previsão de um valor futuro y_{n+h} , dado que temos observado a série até o momento n . Assumindo que o processo $\text{SINAR}(1)_s$ é estacionário, após sucessivas substituições podemos escrever y_{n+h} da seguinte maneira

$$y_{n+h} = P^q \circ y_{n-r} + \sum_{j=0}^{q-1} P^j \circ \epsilon_{n+h-j_s}, \quad h = 1, 2, 3, \dots, \quad (4.13)$$

onde $h = qs - r$, $q = \lceil \frac{h}{s} \rceil$, $q \geq 1$, $r = 0, \dots, s - 1$ e $\lceil \frac{h}{s} \rceil$ retorna o menor inteiro que seja maior ou igual a $\frac{h}{s}$.

Partindo-se da origem em n , e supondo que o objetivo é prever a série em um período futuro h , \hat{y}_{n+h} representa a previsão para um período $n + h$ feita em n . O valor de y_{n+h} , que não se conhece no tempo n , é substituído pela sua previsão \hat{y}_{n+h} . A previsão que apresenta o erro quadrático médio mínimo é a esperança condicional de y_{n+h} , dadas as observações passadas da série, que no caso do $\text{SINAR}(1)_s$ é dada por

$$\mathbb{E}[y_{n+h}|y_1, \dots, y_n] = P^q \left(y_{n-r} - \frac{\lambda}{1-P} \right) + \frac{\lambda}{1-P}, \quad h = 1, 2, 3, \dots$$

Logo um estimador para a previsão y_{n+h} é dado por

$$\hat{y}_{n+h} = \hat{P}^q \left(y_{n-r} - \frac{\hat{\lambda}}{1-\hat{P}} \right) + \frac{\hat{\lambda}}{1-\hat{P}}, \quad h = 1, 2, 3, \dots, \quad (4.14)$$

onde o erro de previsão é dado por $e_{n+h} = y_{n+h} - \hat{y}_{n+h}$ e é fácil verificar que $\mathbb{E}(e_{n+h}) = 0$. Podemos observar que as previsões dadas por (4.14) não vão gerar valores inteiros. Esse mesmo problema também ocorre nos processos $\text{INAR}(1)$ discutido no Capítulo 2.

PROPOSIÇÃO 4.3. *O modelo $\text{SINAR}(1)_s$ satisfaz as seguintes propriedades*

1. $\mathbb{E}[y_{n+h}|y_1, \dots, y_n] = P^q \left(y_{n-r} - \frac{\lambda}{1-P} \right) + \frac{\lambda}{1-P}, \quad h = 1, 2, 3, \dots,$
2. $\text{Var}[y_{n+h}|y_1, \dots, y_n] = P^q (1 - P^q) y_{n-r} + \frac{\lambda(1-P^q)}{1-P}, \quad h = 1, 2, 3, \dots,$
3. $\lim_{h \rightarrow \infty} \mathbb{E}[y_{n+h}|y_1, \dots, y_n] = \lim_{h \rightarrow \infty} \text{Var}[y_{n+h}|y_1, \dots, y_n] = \frac{\lambda}{1-P}.$

A próxima seção apresenta um estudo de Monte Carlo com o intuito de comparar os vícios e os EQMs dos estimadores de Yule-Walker, mínimos quadrados condicional e máxima verossimilhança condicional na estimação dos parâmetros do modelo proposto. Uma aplicação em um conjunto de dados reais também é realizada na Seção 4.5 para ilustrar a metodologia proposta.

4.4 Resultados numéricos

O desempenho dos estimadores dos parâmetros do processo proposto em (4.1) através de métodos empíricos é a motivação do estudo apresentado nesta seção. Os resultados da simulação são expostos por meio de gráficos e tabelas.

Através de 1000 replicações de Monte Carlo, avaliamos os desempenhos dos estimadores de Yule-Walker, mínimos quadrados condicional e de máxima verossimilhança condicional dos parâmetros (P, λ) do processo $\text{SINAR}(1)_s$, com $s = 12$ e $\{\epsilon_t\}_{t \in \mathbb{Z}} \sim \mathcal{Po}(\lambda)$, em amostras de tamanho finito e sob diferentes cenários. Os tamanhos amostrais considerados foram $n = 100, 250$ e 500 . Os valores considerados para o parâmetro P são $P = 0.30, 0.50$ e 0.80 , e para o parâmetro $\lambda = 1$ e 5 .

Para o método de máxima verossimilhança condicional, as estimativas obtidas pelo método de Yule-Walker são utilizadas como valores iniciais para o procedimento de estimação. A função de log-verossimilhança (4.12) foi maximizada através do método de otimização não-linear BFGS com derivadas numéricas.

Para a análise dos resultados da estimação pontual foram calculados, para cada tamanho de amostra, o viés e o erro quadrático médio (EQM) dos estimadores para cada combinação dos parâmetros P e λ , usando as 1000 estimativas.

Tabela 4.1: Viés dos estimadores no caso $\lambda = 1$

n	P	Viés(\hat{P})			Viés($\hat{\lambda}$)		
		YW	MQC	MVC	YW	MQC	MVC
100	0.30	-0.0178	-0.0307	-0.0067	0.0315	0.0508	0.0040
	0.50	-0.0240	-0.0334	-0.0081	0.0500	0.0691	0.0064
	0.80	-0.0267	-0.0362	-0.0031	0.1385	0.1854	0.0113
250	0.30	-0.0115	-0.0156	-0.0067	0.0221	0.0282	0.0115
	0.50	-0.0106	-0.0146	-0.0029	0.0254	0.0337	0.0057
	0.80	-0.0143	-0.0166	-0.0016	0.0700	0.0823	0.0028
500	0.30	-0.0058	-0.0079	-0.0023	0.0063	0.0093	-0.0008
	0.50	-0.0033	-0.0056	-0.0007	0.0102	0.0148	0.0029
	0.80	-0.0086	-0.0098	-0.0003	0.0468	0.0500	0.0043

A Tabela 4.1 apresenta o viés para o modelo gerado com $\lambda = 1$. Nesse cenário, podemos observar que o estimador de MVC se destaca, apresentando em módulo, menor viés que os outros estimadores considerados. Também podemos observar que, para o parâmetro λ , o viés dos estimadores de YW e MQC aumentam quando o processo está próximo da região de não estacionariedade. Note que à medida que o tamanho da amostra aumenta, o viés tende para zero nos três casos, o que evidencia que os estimadores são assintoticamente não viesados.

Tabela 4.2: EQM dos estimadores no caso $\lambda = 1$

n	P	EQM(\hat{P})			EQM($\hat{\lambda}$)		
		YW	MQC	MVC	YW	MQC	MVC
100	0.30	0.0125	0.0133	0.0114	0.0353	0.0365	0.0291
	0.50	0.0100	0.0116	0.0063	0.0549	0.0560	0.0304
	0.80	0.0058	0.0078	0.0012	0.1583	0.1921	0.0289
250	0.30	0.0045	0.0044	0.0035	0.0130	0.0133	0.0104
	0.50	0.0037	0.0040	0.0023	0.0174	0.0184	0.0109
	0.80	0.0019	0.0022	0.0004	0.0511	0.0572	0.0113
500	0.30	0.0022	0.0022	0.0018	0.0055	0.0056	0.0045
	0.50	0.0018	0.0018	0.0010	0.0087	0.0089	0.0052
	0.80	0.0009	0.0009	0.0002	0.0237	0.0255	0.0055

Na Tabela 4.2 temos o EQM para o modelo gerado com $\lambda = 1$ e podemos observar que o estimador de MVC também apresenta menor EQM para as estimativas de P e λ em todos os tamanhos amostrais entre os estimadores considerados nas simulações. Observamos também que à medida que o tamanho da amostra aumenta, o EQM dos três estimadores diminui, tendendo para zero e evidenciando que os três estimadores são consistentes. Também nota-se que os EQMs dos estimadores de YW e MQC para o parâmetro λ aumentam quando o processo está próximo da região de não estacionariedade.

O fato de existir uma alta correlação negativa entre os estimadores de P e λ , quando P é grande, para os métodos de YW e MQC, como observamos nas Equações (4.8) e (4.10), pode explicar o motivo da piora dos estimadores quando P está perto da região de não estacionariedade.

O EQM do estimador de YW é, em geral, menor que o de MQC. No entanto, ao aumentar o tamanho da amostra reduzem-se e aproximam-se os valores dos EQMs. Portanto, os resultados empíricos indicam que o estimador de MQC e o de YW são assintoticamente equivalentes.

As poucas diferenças significantes encontradas entre os métodos de YW e MQC para a estimativa do parâmetro λ são explicadas pelo fato de que o estimador de YW é calculado utilizando tamanho de amostra n (Equação (4.8)) enquanto que o estimador de MQC é calculado baseado em uma amostra de tamanho $n - s - 1$ (Equação (4.11)). As Tabelas 4.3 e 4.4 apresentam o viés e o EQM, respectivamente, para o modelo gerado com $\lambda = 5$.

Tabela 4.3: Viés dos estimadores no caso $\lambda = 5$

n	P	Viés(\hat{P})			Viés($\hat{\lambda}$)		
		YW	MQC	MVC	YW	MQC	MVC
100	0.30	-0.0180	-0.0855	-0.0083	0.2048	0.6947	0.0765
	0.50	-0.0199	-0.0933	-0.0033	0.2620	1.0068	0.0335
	0.80	-0.0304	-0.1509	-0.0003	0.7953	3.7812	0.0017
250	0.30	-0.0082	-0.0300	-0.0034	0.0750	0.2348	0.0197
	0.50	-0.0083	-0.0307	-0.0025	0.1063	0.3314	0.0262
	0.80	-0.0135	-0.0406	-0.0011	0.3572	1.7114	0.0305
500	0.30	-0.0019	-0.0125	-0.0002	0.0304	0.1068	0.0077
	0.50	-0.0029	-0.0134	0.0009	0.0369	0.1426	-0.0109
	0.80	-0.0050	-0.0162	0.0000	0.1353	0.4147	0.0001

Nota-se que os resultados dos vícios e dos EQMs dos estimadores YW e MQC são maiores e menores, respectivamente, para a estimativa do parâmetro P comparados com os valores descritos nas Tabelas 4.1 e 4.2. Na estimativa do parâmetro λ há um acréscimo tanto no viés quanto no EQM dos três estimadores quando comparados com os valores descritos nas Tabelas 4.1 e 4.2. Com relação ao parâmetro λ , é constatado que o desempenho do estimador de MQC varia de acordo com λ . Entretanto, a interpretação do desempenho empírico dos estimadores é equivalente à do modelo com $\lambda = 1$.

Tabela 4.4: EQM dos estimadores no caso $\lambda = 5$

n	P	EQM(\hat{P})			EQM($\hat{\lambda}$)		
		YW	MQC	MVC	YW	MQC	MVC
100	0.30	0.0113	0.0216	0.0103	0.6722	1.3055	0.5793
	0.50	0.0097	0.0236	0.0060	1.1078	2.6607	0.6218
	0.80	0.0055	0.0516	0.0010	3.5483	32.314	0.6105
250	0.30	0.0041	0.0054	0.0035	0.2340	0.3073	0.1940
	0.50	0.0035	0.0049	0.0021	0.3669	0.5212	0.2121
	0.80	0.0019	0.0044	0.0003	1.2000	2.7742	0.2269
500	0.30	0.0019	0.0022	0.0017	0.1083	0.1236	0.0905
	0.50	0.0016	0.0019	0.0001	0.1681	0.1982	0.1059
	0.80	0.0009	0.0013	0.0002	0.5409	0.7884	0.1104

Como observado na Tabela 4.1, os resultados apresentados na Tabela 4.3 indicam que o estimador de MVC possui vício menor que o estimador de YW e MQC. Porém, para maiores tamanhos amostrais, o vício de ambos tende a zero.

Com a análise das tabelas podemos ver a superioridade do método de MVC tanto em termos de viés e EQM. Isso era esperado, uma vez que o estimador de MVC utiliza toda a informação da distribuição. O estudo de simulação também revela que não há nenhum ganho ao considerar o método YW sobre o método MQC no modelo gerado com $\lambda = 1$, sendo que as diferenças entre o viés e os seus erros quadrados médios são insignificantes. A magnitude do ganho em termos de viés e EQM, com o estimador de MVC em relação aos outros dois estimadores, faz valer os cálculos extras e custo computacional.

4.5 Aplicação

Como ilustração da metodologia proposta nas seções anteriores deste capítulo, os dados diários do índice de qualidade do ar, em relação ao dióxido de enxofre (SO_2) da cidade de Cariacica-ES, no período de 1 de janeiro de 2007 a 21 de março de 2007, serão analisados nesta seção. Esses dados foram fornecidos pelo instituto estadual de meio ambiente e recursos hídricos do Espírito Santo (IEMA).

O monitoramento da qualidade do ar é realizado para determinar o nível de concentração dos poluentes presentes na atmosfera. Os resultados não só permitem um acompanhamento sistemático da qualidade do ar na área monitorada, como também constituem elementos básicos para elaboração de diagnósticos da qualidade do ar, subsidiando ações governamentais para o controle das emissões. O Índice de Qualidade do Ar (IQA) é um parâmetro utilizado para transformar as concentrações medidas dos diversos poluentes em um único valor inteiro não negativo e adimensional que possibilita a comparação com os limites legais de concentração (padrões de qualidade do ar) para os diversos poluentes. O IQA tem como objetivo principal proporcionar à população o entendimento sobre a qualidade do ar local em relação a diversos poluentes atmosféricos amostrados nas estações de monitoramento.

O conhecimento prévio dos níveis dos poluentes na atmosfera de uma região pode ser útil para fornecer dados para ativar ações de emergência durante períodos de estagnação atmosférica, quando os níveis de poluentes na atmosfera podem representar risco à saúde pública. Para esse conhecimento prévio, várias estratégias de modelagens têm sido desenvolvidas e otimizadas no intuito de gerar previsões satisfatórias das concentrações dos poluentes em diversos cenários e, conseqüentemente, obter os índices de qualidade do ar. A utilização de modelos de previsão é uma metodologia importante para descrever, através de medidas de probabilidade, o comportamento dos dados, podendo dessa forma prever possíveis índices de qualidade do ar.

Nesse contexto, esta seção propõe um modelo para o IQA da cidade de Cariacica-ES utilizando a metodologia proposta nas seções anteriores. O primeiro objetivo é proceder a análise exploratória da série temporal relativa ao índice de qualidade do ar através do uso de gráficos e de algumas medidas descritivas. Em seguida é realizado o ajuste e a previsão do IQA.

A série do índice de qualidade do ar apresenta 80 observações e possui algumas características que dificultam a utilização de modelos gaussianos. Em primeiro lugar, a magnitude dos valores da série é extremamente pequena. Além disso, a série apresenta uma grande quantidade de valores iguais a 1, que correspondem a aproximadamente 41% das observações, além de não assumir valores negativos.

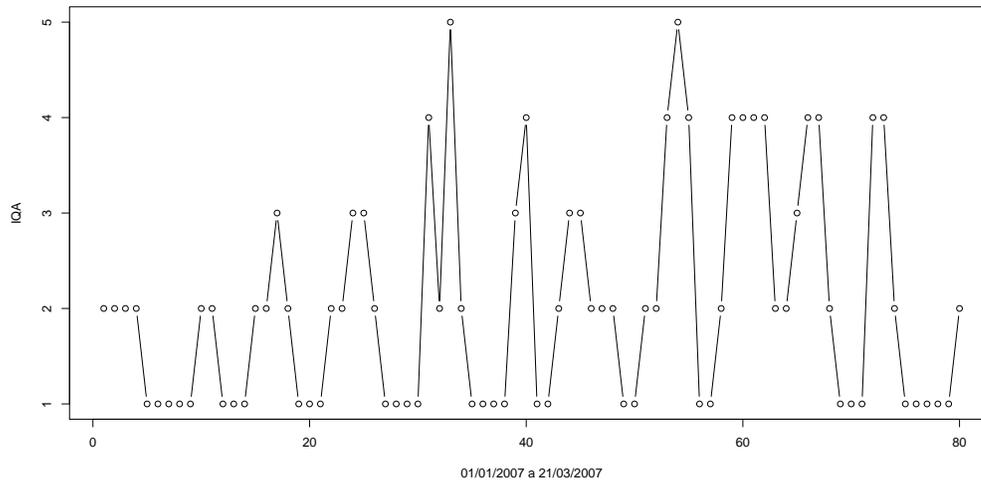


Figura 4.3: Gráfico de evolução da série índice de qualidade do ar da cidade de Cariacica-ES.

A Figura 4.3 sugere que a série é estacionária e, além disso, apresenta um comportamento sazonal que é confirmado pelo gráfico da função de autocorrelação amostral (ACF). Pela função de autocorrelação amostral da série, percebe-se indícios de sazonalidade com período sazonal $s = 7$.

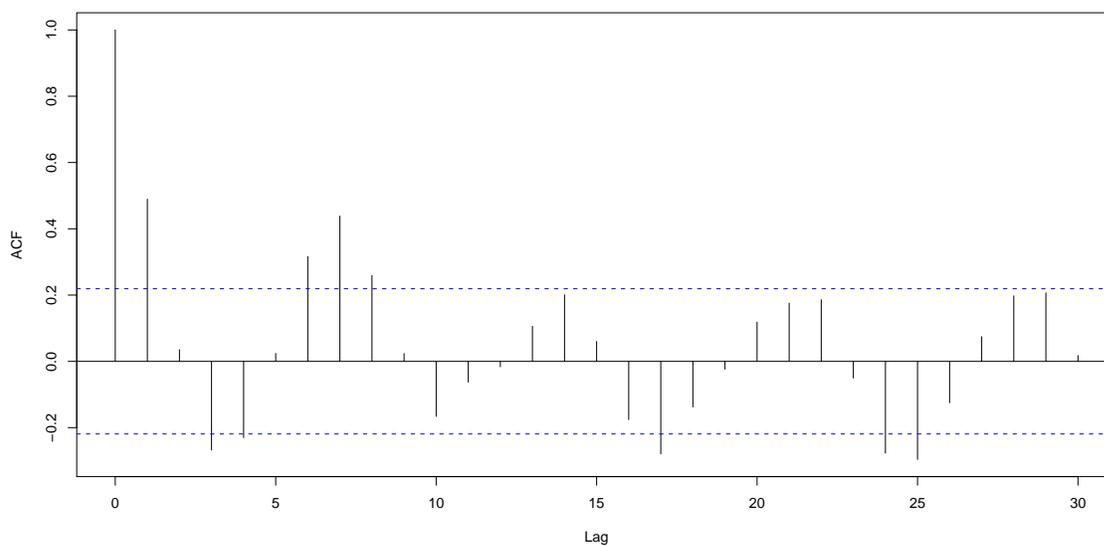


Figura 4.4: ACF da série temporal do índice de qualidade do ar da cidade de Cariacica-ES.

Na Tabela 4.5, são apresentadas algumas medidas de posição e dispersão.

Tabela 4.5: Medidas de posição e dispersão.

Medida	IQA
Mínimo	1.000
Média	2.050
Mediana	2.000
Desvio padrão	1.158
Máximo	5.000

Podemos observar na Tabela 4.5, que a série apresenta uma média amostral de 2.050 com o desvio padrão de 1.158. O maior valor observado é 5 e o valor mínimo é 1.

O histograma e o box-plot apresentados auxiliam na interpretação e análise da dispersão e distribuição dos dados. Pela Figura 4.5, constata-se que o conjunto de dados se distribui de maneira assimétrica e não apresenta possíveis *outliers*.

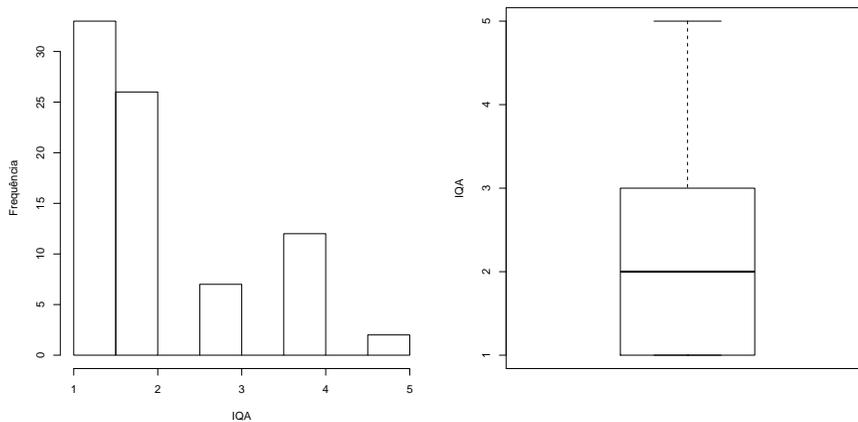


Figura 4.5: Histograma e box-plot do índice de qualidade do ar da cidade de Cariacica-ES.

A função de autocorrelação amostral (Figura 4.4) indica que a série do IQA da cidade de Cariacica-ES apresenta sazonalidade de período $s = 7$ e este trabalho adotou, pela proposta desta pesquisa, o modelo $\text{SINAR}(1)_7$. Entretanto, outros parâmetros talvez deveriam ser considerados no modelo para obter melhor ajuste. Assim, a modelagem do IQA vai se restringir ao $\text{SINAR}(1)_7$, que talvez não seja o modelo mais adequado; entretanto, a análise residual para este modelo parece indicar que o mesmo está bem ajustado.

A Tabela 4.6 mostra, para o modelo proposto, as estimativas dos parâmetros P e λ , através dos métodos Yule-Walker, mínimos quadrados condicional e máxima verossimilhança condicional.

Tabela 4.6: Estimativas dos parâmetros do modelo.

Método	\hat{P}	$\hat{\lambda}$
YW	0.4526	1.1507
MQC	0.4367	1.1849
MVC	0.6416	0.7336

Repare na Tabela 4.6 que as estimativas utilizando os métodos de YW e MQC são muito próximas, como já era esperado. Na Seção 4.4 foi observado que o estimador de máxima verossimilhança condicional apresentou melhores propriedades para a estimação dos parâmetros do processo $\text{SINAR}(1)_s$. Diante desse fato o modelo proposto para a série do IQA será baseado no estimador de máxima verossimilhança condicional. Dessa forma, a equação estimada desse modelo é dada por

$$y_t = 0.6416 \circ y_{t-7} + \epsilon_t, \quad \{\epsilon_t\} \sim \mathcal{Po}(0.7336). \quad (4.15)$$

A Figura 4.6 apresenta o histograma e o box-plot dos resíduos. Nota-se pelo histograma e pelo box-plot dos resíduos, que os mesmos se distribuem de maneira assimétrica.

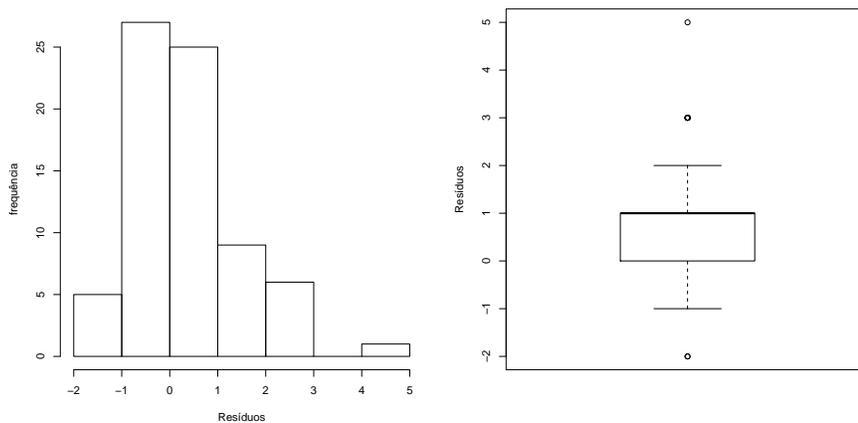


Figura 4.6: Box-plot dos resíduos.

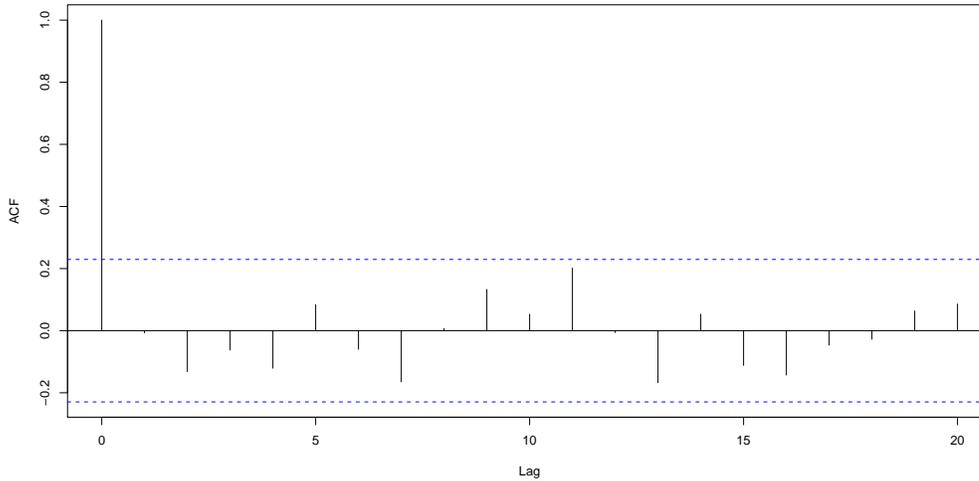


Figura 4.7: ACF dos resíduos.

A Figura 4.7 mostra a função de autocorrelação amostral dos resíduos e pode-se observar que os resíduos se encontram dentro dos limites de confiança. Entretanto, esses limites são obtidos através da suposição de normalidade, suposição que não é satisfeita no modelo considerado. Portanto, podemos apenas afirmar que as autocorrelações amostrais dos resíduos pertencem ao intervalo $(-0.2, 0.2)$.

A fim de realizar previsões com o modelo ajustado, será utilizado o estimador para a previsão definido em (4.14). As previsões dos índices de qualidade do ar para os dias 22/03/07, 24/03/07 e 26/03/07 são dadas na Tabela 4.7.

Tabela 4.7: Previsões para a série IQA com origem em $n = 80$, utilizando o modelo $\text{SINAR}(1)_7$ dado por (4.15).

h	y_{n+h}	\hat{y}_{n+h}	\hat{e}_{n+h}
1	2	2.0168	-0.0168
3	2	1.3752	0.6248
5	3	1.3752	1.6248

Podemos observar na Tabela 4.7 que as previsões encontradas não são valores inteiros como já discutido no Capítulo 4.

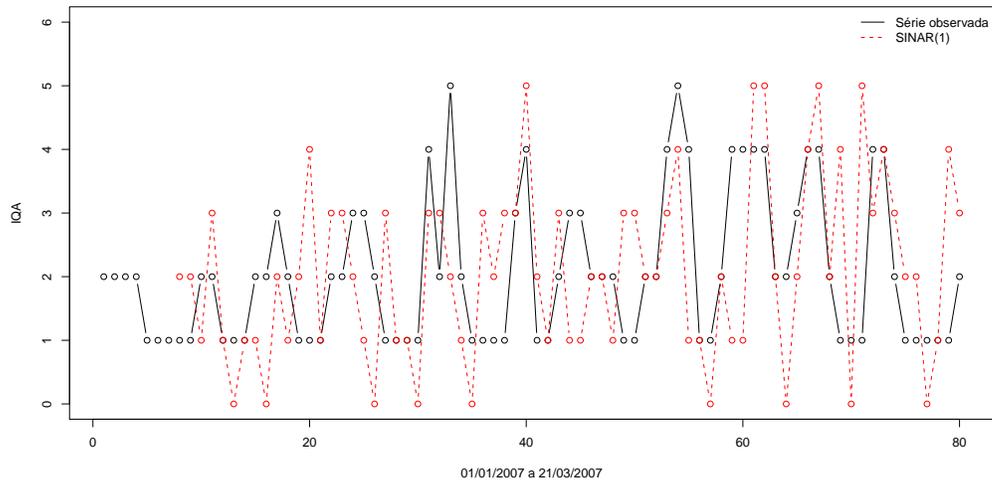


Figura 4.8: Série observada e série ajustada.

Repare na Figura 4.8 que perdemos as sete primeiras observações quando realizamos o ajuste do modelo proposto e, além disso, podemos ver que o modelo ajustado acompanha a média da série, mas não acompanha bem as observações. Portanto, outros parâmetros talvez devessem ser considerados no modelo para obter melhor ajuste, como já dito anteriormente.

CAPÍTULO 5

Considerações finais

As contribuições científicas deste trabalho concentram em dois pontos principais. O primeiro dedica-se ao estudo do processo INAR(1), introduzido por Al-Osh & Alzaid (1987), quando o processo apresenta raiz unitária. O segundo consiste em propor o modelo de valores inteiros com estrutura sazonal baseado no processo introduzido por Al-Osh & Alzaid (1987), denotado por $\text{SINAR}(1)_s$.

No Capítulo 2, foram apresentadas características e propriedades do processo INAR(1), além dos procedimentos existentes para estimação do processo e as propriedades assintóticas dos estimadores, a fim de obter uma base sólida para a construção do processo sazonal e para o estudo do processo na presença de raiz unitária.

No contexto da primeira contribuição deste trabalho, os resultados empíricos evidenciaram que a utilização do teste usual de Dickey-Fuller, em processos INAR(1), pode resultar em conclusões errôneas. Diante desse fato, para vários valores do percentil α , os pontos críticos da estatística de teste do teste de Dickey-Fuller, quando o teste é utilizado em processo INAR(1), foram calculados. Comparações entre a utilização do teste de Dickey-Fuller em processos com marginais contínuas e discretas também foram abordadas e se encontram no Capítulo 3.

No que tange aos processos de contagem com estrutura sazonal, foi proposto um modelo de valores inteiros com estrutura sazonal baseado no modelo apresentado por Al-Osh & Alzaid (1987). As principais propriedades do modelo proposto foram derivadas,

tais como os momentos, a função de autocovariância e a função de autocorrelação e esses resultados encontram-se no Capítulo 4. Ensaios de Monte Carlo foram realizados para comparar os vícios e os erros quadráticos médios de três estimadores para os parâmetros do modelo proposto. Os resultados das simulações de Monte Carlo realizadas mostraram que o estimador de máxima verossimilhança condicional apresenta melhores propriedades quando comparado com os estimadores de Yule-Walker e mínimos quadrados condicional. Dessa forma, para estimar os parâmetros (P, λ) do processo $\text{SINAR}(1)_s$, recomenda-se utilizar o método de máxima verossimilhança condicional. Além disso, através de um exemplo prático, nota-se que o modelo proposto se adequou para a modelagem de dados inteiros com comportamento sazonal.

Como possíveis trabalhos a serem desenvolvidos futuramente, destacamos: encontrar a distribuição assintótica da estatística de teste do teste de Dickey-Fuller quando o teste é utilizado em processos INAR e generalizar o processo autorregressivo de valores inteiros com estrutura sazonal para ordem P , ou seja, $\text{SINAR}(P)_s$.

Neste apêndice, obteremos as demonstrações das proposições, lemas e teoremas descritos ao longo deste trabalho.

Prova do Lema 2.1.

1. $0 \circ X = \sum_{j=1}^X z_j$, tal que $P(z_j = 1) = p = 0$. Logo, $0 \circ X = \sum_{j=1}^X z_j = \sum_{j=1}^X 0 = 0$.

2. $1 \circ X = \sum_{j=1}^X z_j$, tal que $P(z_j = 1) = 1$. Logo, $1 \circ X = \sum_{j=1}^X z_j = \sum_{j=1}^X 1 = X$.

3. Seja $\mathcal{G}_X(t) = \mathbb{E}(t^X)$ a função geradora de probabilidade de X . Se $Y = p \circ X = \sum_{j=1}^X z_j$, temos que

$$\mathcal{G}_Y(t) = \mathcal{G}_{p \circ X}(t) = \mathcal{G}_X(\mathcal{G}_z(t)) = \mathcal{G}_X(pt + 1 - p),$$

onde $\mathcal{G}_z(t)$ é a função geradora de probabilidade da Bernoulli(p). Fazendo $Y = q \circ X = \sum_{j=1}^X z_j$ e $W = p \circ Y = \sum_{j=1}^Y w_j$, temos que

$$\mathcal{G}_W(t) = \mathcal{G}_Y(\mathcal{G}_w(t)) = \mathcal{G}_X(\mathcal{G}_z(\mathcal{G}_w(t))) = \mathcal{G}_X(pqt + 1 - pq).$$

Portanto, $p \circ (q \circ X) \stackrel{d}{=} (pq) \circ X$.

4. $p \circ (X+Y) = \sum_{j=1}^{X+Y} z_j = \sum_{j=1}^X z_j + \sum_{j=X+1}^Y z_j = p \circ X + \sum_{j=1}^Y z_{j+X} \stackrel{d}{=} p \circ X + p \circ Y.$
5. $\mathbb{E}(p \circ X | X) = \mathbb{E}\left(\sum_{j=1}^X z_j | X\right) = \sum_{j=1}^X \mathbb{E}(z_j | X) = \sum_{j=1}^X \mathbb{E}(z_j) = pX.$
6. $\mathbb{E}(p \circ X) = \mathbb{E}(\mathbb{E}(p \circ X | X)) = \mathbb{E}(pX) = p\mathbb{E}(X).$
7. $\text{Var}(p \circ X | X) = \text{Var}\left(\sum_{j=1}^X z_j | X\right) = \sum_{j=1}^X \text{Var}(z_j | X) = \sum_{j=1}^X \text{Var}(z_j) = p(1-p)X.$
8. $\text{Var}(p \circ X) = \text{Var}(\mathbb{E}(p \circ X | X)) + \mathbb{E}(\text{Var}(p \circ X | X)) = p^2 \text{Var}(X) + p(1-p)\mathbb{E}(X).$
9.
$$\begin{aligned} \text{Cov}(p \circ X, q \circ Y) &= \mathbb{E}[(p \circ X)(q \circ Y)] - \mathbb{E}(p \circ X)\mathbb{E}(q \circ Y) \\ &= \mathbb{E}[\mathbb{E}[(p \circ X)(q \circ Y) | X, Y]] - \mathbb{E}(p \circ X)\mathbb{E}(q \circ Y) \\ &= \mathbb{E}\left[\mathbb{E}\left[\left(\sum_{i=1}^X z_{i,1}\right)\left(\sum_{j=1}^Y z_{j,2}\right) | X, Y\right]\right] - \mathbb{E}(p \circ X)\mathbb{E}(q \circ Y) \\ &= \mathbb{E}\left[\sum_{i=1}^X \sum_{j=1}^Y \mathbb{E}[z_{i,1}z_{j,2} | X, Y]\right] - pq\mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}\left[\sum_{i=1}^X \sum_{j=1}^Y \mathbb{E}[z_{i,1}z_{j,2}]\right] - pq\mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}[XYpq] - pq\mathbb{E}(X)\mathbb{E}(Y) \\ &= pq \text{Cov}(X, Y). \end{aligned}$$

Prova do Teorema 3.1.

Pelo Corolário 2.1 temos, assintoticamente, que

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, (p(1-p)^2/\lambda + (1-p^2))),$$

supondo que $\lim_{n \rightarrow \infty} \lim_{\lambda \rightarrow \infty} P(X_n \leq x) = \lim_{\lambda \rightarrow \infty} \lim_{n \rightarrow \infty} P(X_n \leq x)$, em que $X_n = \sqrt{n}(\hat{p} - p)$, temos que para $n, \lambda \rightarrow \infty$

$$\sqrt{n}(\hat{p} - p) \xrightarrow{d} \mathcal{N}(0, (1-p^2)).$$

Prova da Proposição 4.1.

Substituindo sucessivamente y_{t-s}, y_{t-2s}, \dots , obtemos

$$\begin{aligned}
 y_t &= P \circ y_{t-s} + \epsilon_t \\
 &= P \circ (P \circ y_{t-2s} + \epsilon_{t-s}) + \epsilon_t \\
 &\stackrel{d}{=} P^2 \circ y_{t-2s} + P \circ \epsilon_{t-s} + \epsilon_t \\
 &\stackrel{d}{=} P^2 \circ (P \circ y_{t-3s} + \epsilon_{t-2s}) + P \circ \epsilon_{t-s} + \epsilon_t \\
 &\stackrel{d}{=} P^3 \circ y_{t-3s} + P^2 \circ \epsilon_{t-2s} + P \circ \epsilon_{t-s} + \epsilon_t \\
 &\vdots \\
 &\stackrel{d}{=} P^h \circ y_{t-hs} + \sum_{j=0}^{h-1} P^j \circ \epsilon_{t-js}.
 \end{aligned}$$

Assim,

$$\begin{aligned}
 \gamma(k) &= \text{Cov}(y_{t-k}, y_t) = \text{Cov}(y_{t-k}, P^h \circ y_{t-hs}) + \text{Cov}(y_{t-k}, \sum_{j=0}^{h-1} P^j \circ \epsilon_{t-js}) \\
 &= \text{Cov}(y_{t-k}, P^h \circ y_{t-hs}) + \mathbb{E} \left(y_{t-k} \sum_{j=0}^{h-1} P^j \circ \epsilon_{t-js} \right) - \mathbb{E}(y_{t-k}) \mathbb{E} \left(\sum_{j=0}^{h-1} P^j \circ \epsilon_{t-js} \right) \\
 &= P^h \text{Cov}(y_{t-k}, y_{t-hs}) + \sum_{j=0}^{h-1} P^j \mathbb{E}(y_{t-k} \epsilon_{t-js}) - \sum_{j=0}^{h-1} P^j \mathbb{E}(y_{t-k} \epsilon_{t-js}) \\
 &= P^h \text{Cov}(y_{t-k}, y_{t-hs}) = \begin{cases} \gamma(0) P^{k/s}, & \text{se } k \text{ é múltiplo de } s, \\ 0, & \text{caso contrário,} \end{cases}
 \end{aligned}$$

Prova da Proposição 4.2.

Através de uma amostra y_1, \dots, y_n de um processo $\text{SINAR}(1)_s$ estacionário, a estimação dos parâmetros P e λ através do método de mínimos quadrados condicional consiste em minimizar a soma dos quadrados

$$Q_n(P, \lambda) = \sum_{t=s+1}^n (y_t - Py_{t-s} - \lambda)^2. \quad (5.1)$$

Derivando (5.1) em relação a P e λ e igualando a 0, temos que

$$\Psi_n(P, \lambda) = \begin{bmatrix} \sum_{t=s+1}^n (y_t - Py_{t-s} - \lambda)y_{t-s} \\ \sum_{t=s+1}^n (y_t - Py_{t-s} - \lambda) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Definindo $\Psi_n(P, \lambda) = \sum_{t=s+1}^n \Psi_t$, onde $\Psi_t = (\Psi_{Pt}, \Psi_{\lambda t}^\top)$ em que $\Psi_{Pt} = (y_t - Py_{t-s} - \lambda)y_{t-s}$ e $\Psi_{\lambda t} = (y_t - Py_{t-s} - \lambda)$. De acordo com o Teorema 2 de Freeland & MacCabe (2005) a distribuição assintótica de \hat{P} e $\hat{\lambda}$ é dada por $\sqrt{n}(\begin{smallmatrix} \hat{P}-P \\ \hat{\lambda}-\lambda \end{smallmatrix}) \xrightarrow{d} \mathcal{N}(0, V^{-1}WV^{-1})$, onde $W = \mathbb{E}(\Psi_t \Psi_t^\top)$ e $V = \mathbb{E}\left(\frac{\partial \Psi_t}{\partial \Theta_i}\right)$, $\Theta = (p, \lambda)^\top$, $i = 1, 2$.

Então,

$$\begin{aligned} \mathbb{E}\left(\frac{\partial \Psi_{Pt}}{\partial P}\right) &= \mathbb{E}(-y_{t-s}^2) = -\left(\frac{\lambda}{1-P} + \frac{\lambda^2}{(1-P)^2}\right), \\ \mathbb{E}\left(\frac{\partial \Psi_{Pt}}{\partial \lambda}\right) &= \mathbb{E}(-y_{t-s}) = -\frac{\lambda}{1-P}, \\ \mathbb{E}\left(\frac{\partial \Psi_{\lambda t}}{\partial P}\right) &= \mathbb{E}(-y_{t-s}) = -\frac{\lambda}{1-P}, \\ \mathbb{E}\left(\frac{\partial \Psi_{\lambda t}}{\partial \lambda}\right) &= -1, \end{aligned}$$

logo,

$$V = \begin{bmatrix} -\left(\frac{\lambda}{1-P} + \frac{\lambda^2}{(1-P)^2}\right) & -\frac{\lambda}{1-P} \\ -\frac{\lambda}{1-P} & -1 \end{bmatrix} \Rightarrow V^{-1} = \begin{bmatrix} \left(\frac{\lambda}{1-P} + \frac{\lambda^2}{(1-P)^2}\right) & -1 \\ -1 & 1 + \frac{\lambda}{1-P} \end{bmatrix}.$$

$$\begin{aligned} \mathbb{E}(\Psi_{Pt}^2) &= \mathbb{E}[\mathbb{E}((y_t - Py_{t-s} - \lambda)^2 | y_{t-s}) y_{t-s}^2] \\ &= \mathbb{E}[(P(1-P)y_{t-s} + \lambda)y_{t-s}^2] \\ &= P(1-P)\mathbb{E}(y_{t-s}^3) + \lambda\mathbb{E}(y_{t-s}^2) \\ &= P\lambda + \lambda^2 \left(\frac{1+3P}{1-P}\right) + \lambda^3 \left(\frac{1+P}{(1-P)^2}\right), \end{aligned}$$

$$\begin{aligned} \mathbb{E}(\Psi_{\lambda t}^2) &= \mathbb{E}[\mathbb{E}((y_t - Py_{t-s} - \lambda)^2 | y_{t-s})] \\ &= \mathbb{E}[P(1-P)y_{t-s} + \lambda] \\ &= \lambda(1+P), \end{aligned}$$

$$\begin{aligned}
\mathbb{E}(\Psi_{Pt}\Psi_{\lambda t}) &= \mathbb{E}[\mathbb{E}((y_t - Py_{t-s} - \lambda)^2 | y_{t-s})y_{t-s}] \\
&= \mathbb{E}[(P(1-P)y_{t-s} + \lambda)y_{t-s}] \\
&= P(1-P)\mathbb{E}(y_{t-s}^2) + \lambda\mathbb{E}(y_{t-s}) \\
&= P\lambda + \frac{1+P}{1-P}\lambda^2.
\end{aligned}$$

Assim,

$$W = \begin{bmatrix} P\lambda + \lambda^2 \left(\frac{1+3P}{1-P}\right) + \lambda^3 \left(\frac{1+P}{(1-P)^2}\right) & P\lambda + \frac{1+P}{1-P}\lambda^2 \\ P\lambda + \frac{1+P}{1-P}\lambda^2 & (1+P)\lambda \end{bmatrix},$$

e portanto

$$V^{-1}WV^{-1} = \begin{bmatrix} \frac{P(1-P)^2}{\lambda} + (1-P^2) & -(1+P)\lambda \\ -(1+P)\lambda & \lambda + \frac{1+P}{1-P}\lambda^2 \end{bmatrix}.$$

Prova da Proposição 4.3.

Substituindo sucessivamente $y_{n+h-s}, y_{n+h-2s}, \dots$, obtemos

$$\begin{aligned}
y_{n+h} &= P \circ y_{n+h-s} + \epsilon_{n+h} \\
&= P \circ (P \circ y_{n+h-2s} + \epsilon_{n+h-s}) + \epsilon_{n+h} \\
&\stackrel{d}{=} P^2 \circ y_{n+h-2s} + P \circ \epsilon_{n+h-s} + \epsilon_t \\
&\stackrel{d}{=} P^2 \circ (P \circ y_{n+h-3s} + \epsilon_{n+h-2s}) + P \circ \epsilon_{n+h-s} + \epsilon_{n+h} \\
&\stackrel{d}{=} P^3 \circ y_{n+h-3s} + P^2 \circ \epsilon_{n+h-2s} + P \circ \epsilon_{n+h-s} + \epsilon_{n+h} \\
&\vdots \\
&\stackrel{d}{=} P^q \circ y_{n-r} + \sum_{j=0}^{q-1} P^j \circ \epsilon_{n+h-js},
\end{aligned}$$

onde $h = qs - r, q = \lceil \frac{h}{s} \rceil, q \geq 1, r = 0, \dots, s-1$ e $\lceil \frac{h}{s} \rceil = \min\{n \in \mathbb{Z} \mid \lceil \frac{h}{s} \rceil \leq n\}$.

Então,

$$\begin{aligned}
1. \mathbb{E}(y_{n+h}|y_1, \dots, y_n) &= \mathbb{E} \left(P^q \circ y_{n-r} + \sum_{j=0}^{q-1} P^j \circ \epsilon_{n+h-j_s} \right) \\
&= P^q y_{n-r} + \lambda \sum_{j=0}^{q-1} P^j \\
&= P^q y_{n-r} + \frac{\lambda(1 - P^q)}{1 - P} \\
&= P^q \left(y_{n-r} - \frac{\lambda}{1 - P} \right) + \frac{\lambda}{1 - P}.
\end{aligned}$$

$$\begin{aligned}
2. \text{Var}(y_{n+h}|y_1, \dots, y_n) &= \text{Var} \left(P^q \circ y_{n-r} + \sum_{j=0}^{q-1} P^j \circ \epsilon_{n+h-j_s} \right) \\
&= \text{Var}(P^q \circ y_{n-r}) + \text{Var} \left(\sum_{j=0}^{q-1} P^j \circ \epsilon_{n+h-j_s} \right) \\
&= P^q(1 - P^q)y_{n-r} + \sum_{j=0}^{q-1} \text{Var}(P^j \circ \epsilon_{n+h-j_s}) \\
&= P^q(1 - P^q)y_{n-r} + \frac{\lambda(1 - P^q)}{1 - P}.
\end{aligned}$$

$$\begin{aligned}
3. \lim_{h \rightarrow \infty} \mathbb{E}(y_{n+h}|y_1, \dots, y_n) &= \lim_{h \rightarrow \infty} \left[P^q \left(y_{n-r} - \frac{\lambda}{1 - P} \right) + \frac{\lambda}{1 - P} \right] = \frac{\lambda}{1 - P}. \\
\lim_{h \rightarrow \infty} \text{Var}(y_{n+h}|y_1, \dots, y_n) &= \lim_{h \rightarrow \infty} \left[P^q(1 - P^q)y_{n-r} + \frac{\lambda(1 - P^q)}{1 - P} \right] = \frac{\lambda}{1 - P}.
\end{aligned}$$

Função escore

A distribuição condicional de y_t dado y_{t-s} é dada por

$$f(y_t|y_{t-s}) = \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{e^{-\lambda} \lambda^{(y_t)-i}}{((y_t) - i)! \binom{y_{t-s}}{i}} P^i (1 - P)^{(y_{t-s})-i}.$$

$$\begin{aligned}
\frac{\partial f(y_t|y_{t-s})}{\partial P} &= \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{e^{-\lambda} \lambda^{(y_t)-i}}{((y_t)-i)!} \binom{y_{t-s}}{i} P^i (1-P)^{(y_{t-s})-i} \left[\frac{i}{P(1-P)} - \frac{y_{t-s}}{1-P} \right] \\
&= \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{i}{P(1-P)} \binom{y_{t-s}}{i} P^i (1-P)^{(y_{t-s})-i} \frac{e^{-\lambda} \lambda^{(y_t)-i}}{((y_t)-i)!} \\
&\quad - \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{y_{t-s}}{1-P} \binom{y_{t-s}}{i} P^i (1-P)^{(y_{t-s})-i} \frac{e^{-\lambda} \lambda^{(y_t)-i}}{((y_t)-i)!} \\
&= \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{y_{t-s}}{(1-P)} \binom{y_{t-s}-1}{i-1} P^{i-1} (1-P)^{(y_{t-s})-i} \frac{e^{-\lambda} \lambda^{(y_t)-i}}{((y_t)-i)!} \\
&\quad - \frac{y_{t-s}}{1-P} f(y_t|y_{t-s}) \\
&= \frac{y_{t-s}}{(1-P)} \sum_{i=0}^{\min(y_t-1, y_{t-s}-1)} \binom{y_{t-s}-1}{i} P^i (1-P)^{(y_{t-s})-i-1} \frac{e^{-\lambda} \lambda^{(y_t)-i-1}}{((y_t)-i-1)!} \\
&\quad - \frac{y_{t-s}}{1-P} f(y_t|y_{t-s}) \\
&= \frac{y_{t-s}}{(1-P)} [f(y_t-1|y_{t-s}-1) - f(y_t|y_{t-s})]. \tag{5.2}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial f(y_t|y_{t-s})}{\partial \lambda} &= \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{(y_t-i) e^{-\lambda} \lambda^{(y_t)-i-1}}{((y_t)-i)!} \binom{y_{t-s}}{i} P^i (1-P)^{(y_{t-s})-i} \\
&\quad - \sum_{i=0}^{\min(y_t, y_{t-s})} \frac{e^{-\lambda} \lambda^{(y_t)-i}}{((y_t)-i)!} \binom{y_{t-s}}{i} P^i (1-P)^{(y_{t-s})-i} \\
&= \sum_{i=0}^{\min(y_t-1, y_{t-s})} \frac{e^{-\lambda} \lambda^{(y_t)-i-1}}{((y_t)-i-1)!} \binom{y_{t-s}}{i} P^i (1-P)^{(y_{t-s})-i} - f(y_t|y_{t-s}) \\
&= f(y_t-1|y_{t-s}) - f(y_t|y_{t-s}). \tag{5.3}
\end{aligned}$$

A função de log-verossimilhança condicional de y_t dado y_{t-s} é dada por

$$\ell(p, \lambda; y|y_0, \dots, y_{s-1}) = \log \left(\prod_{t=s}^n f(y_t|y_{t-s}) \right) = \sum_{t=s}^n \log(f(y_t|y_{t-s})).$$

De (5.2) e (5.3), temos que

$$\frac{\partial \ell(p, \lambda; y|y_0, \dots, y_{s-1})}{\partial P} = \sum_{t=s}^n \frac{y_{t-s}}{1-P} \frac{[f(y_t - 1|y_{t-s} - 1) - f(y_t|y_{t-s})]}{f(y_t|y_{t-s})}$$

$$\frac{\partial \ell(p, \lambda; y|y_0, \dots, y_{s-1})}{\partial \lambda} = \sum_{t=s}^n \frac{[f(y_t - 1|y_{t-s}) - f(y_t|y_{t-s})]}{f(y_t|y_{t-s})}.$$

Referências Bibliográficas

- [1] Al-Osh, M. A. & Alzaid, A. A. (1987). First-order integer valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, 8, 261–275.
- [2] Barsky, R. & Miron, J. (1989). The seasonal and the business cycle. *Journal of Political Economy*, 97, 503–534.
- [3] Box, G. E. P., Jenkins, G.M. & Reinsel, G. C. (1994). *Time series analysis: forecasting and control*. Prentice-Hall, New Jersey.
- [4] Brännäs, K. (1994). Estimation and testing in integer-valued AR(1) models. *Umeå Economic Studies*, 335. Umeå University, Sweden.
- [5] Brännäs, K., Hellström, J. & Nordström, J. (2002). A new approach to modelling and forecasting monthly guest nights in hotels. *International Journal of Forecasting*, 18, 19–30.
- [6] Brockwell, P. J. & Davis, R. A. (1991). *Time Series: Theory and Methods*, Springer-Verlag, New York.
- [7] Burns, A. F. & Mitchell, W. C. (1946). *Measuring business cycles*. National Bureau of Economic Research, New York.
- [8] Canova, F. (1992). An alternative approach to modeling and forecasting seasonal time series. *Journal of Business & Economic Statistics*, 10, 97–108.

- [9] Chang, Y. Park, J. Y. (2003). A sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis*, 24, 370–400.
- [10] Cribari-Neto, F. & Zarkos, S.G. (1999). R: yet another econometric programming environment. *Journal of Applied Econometrics*, 14, 319–329.
- [11] Dickey, D. & Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427–431.
- [12] Dickey, D. A., Hasza, D. P. & Fuller, W. A. (1984). Testing for unit roots in seasonal time series. *Journal of American Statistical Association*, 79, 355–367.
- [13] Du, J. & Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis*, 12, 129–142.
- [14] Franke, J. & Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In: Subba Rao, T. (Ed.), *Developments in Time Series Analysis*. Chapman-Hall, London, 310–330.
- [15] Freeland, K. (1998). *Statistical analysis of discrete time series with application to the analysis of workers compensation claims data*. PhD dissertation - University of British Columbia, Canada.
- [16] Freeland, K. & McCabe, B.P.M. (2004a). Analysis of low count time series data by Poisson autoregression. *Journal of Time Series Analysis*, 25, 701–722.
- [17] Freeland, K. & McCabe, B.P.M. (2004b). Forecasting discrete valued low count time series. *International Journal of Forecasting*, 20, 427–434.
- [18] Freeland, K. & McCabe, B.P.M. (2005). Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Statistics & Probability Letters*, 73, 147–153.
- [19] Fuller, W. A. (1976). *Introduction to the statistical time series*, Wiley.
- [20] Gallager, R. G. (2009). *Discrete stochastic processes*, Rice University.

- [21] Gauthier, G. & Latour, A. (1994). Convergence forte des estimateurs des paramètres d'un processus GENAR(p). *Annales des Sciences Mathématiques du Québec*, 18, 49–71.
- [22] Hamilton, J. M. (1994). *Time Series Analysis*, Princeton University Press.
- [23] Hansen, L. e Sargent, T. (1993). Seasonality and approximation errors in rational expectations models. *Journal of Econometrics*, 55, 21–55.
- [24] Hassler, U. (1994). (Mis)specification of long memory in seasonal time. *Journal of Time Series Analysis*, 15, 19–30.
- [25] Hellström, J. (2001). Unit root testing in integer-valued AR(1) models. *Economics Letters*, 70, 9–14.
- [26] Ihaka, R. & Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5, 299–314.
- [27] Jacobs, P. A. & Lewis, P. A. W. (1978a). Discrete time series generated by mixtures. I: correlation and runs properties. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 40, 94-105.
- [28] Jacobs, P. A. & Lewis, P. A. W. (1978b). Discrete time series generated by mixtures II: asymptotic properties. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, 40, 222-228.
- [29] Jung, R. C. & Tremayne, A. R. (2003). Testing for serial dependence in time series models of counts. *Journal of Time Series Analysis*, 24, 65-84.
- [30] Jung, R. C., Ronning, G. & Tremayne, A. R. (2005). Estimation in conditional first order autoregression with discrete support. *Statistical Papers*, 46, 195-224.
- [31] Klimko, L. & Nelson, P. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistic*, 6, 629-642.
- [32] Kwiatkowski, D., Phillips, P. C. B, Schmidt, P. & Shin, Y. (1992). Testing the null hypothesis of stationary against the alternative of a unit root. *Journal of Econometrics*, 54, 159–178.

- [33] Latour, A. (1997). The multivariate GINAR(p) process. *Advances in Applied Probability*, 29, 228-248.
- [34] Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *Journal of Time Series Analysis*, 19, 439-455.
- [35] Macaulay, F. R. (1938). *Some theoretical problems suggested by movements of interest rates, bond yields and stock prices in the United States since 1856*. National Bureau of Economic Research, New York.
- [36] Maddala, G. S. & Kim, I.-M. (1998). *Unit roots cointegration and structural change*, Cambridge.
- [37] Mckenzie, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Advances in Applied Probability*, 20, 822-835.
- [38] Mitchell, W. C. (1927). *The business cycles: The problem and its setting*. National Bureau of Economic Research, New York.
- [39] Montgomery, D. C. (2004). *Introdução ao controle estatístico de qualidade*, LTC.
- [40] Morettin, P. A. & Toloi, C. (2004). *Análise de Séries Temporais*, ABE - Projeto Fisher.
- [41] Phillips, P. C. B. (1987). Time series regression with a unit root. *Econometrica*, 55, 277-301.
- [42] Phillips, P. & Perron, P. (1988). Testing for unit root in time series regression. *Biometrika*, 75, 335-346.
- [43] Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association*, 85, 338-344.
- [44] Priestley, M. B. (1981). *Spectral Analysis and Time Series*, Academic press.
- [45] Said, S. & Dickey, D. (1984), Testing for unit root autoregressive-moving average models of unknown order. *Biometrika*, 71, 599-608.

- [46] Silva, M.E. & Oliveira, V. L. (2005). Difference equations for the higher order moments and cumulants of the INAR(p) model. *Technical Report CMA/9/00*, Universidade de Porto.
- [47] Sprott, D.A. (1983). Estimating the parameters of a convolution by maximum likelihood. *Journal of the American Statistical Association*, 78, 457-460.
- [48] Steutel, F.W. & Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, 7, 893-899.
- [49] Stoffer, D.S. (1987). Walsh-Fourier analysis of discrete-valued time series. *Journal of Time Series Analysis*, 8, 449-467.
- [50] Subba Rao, T. (1994). *Developments in Time Series Analysis*, Chapman-Hall.
- [51] Venables, W. N., Smith, D.M. & R Development Core Team. (2009). An introduction to R. Disponível em: <http://cran.r-project.org/doc/manuals/R-intro.pdf>. Arquivo obtido em 17 de agosto de 2010.
- [52] Zheng, H. & Basawa, I. V. (2008). First-order observation-driven integer-valued autoregressive processes. *Statistics and probability letters*, 78, 1-9.