

Universidade Federal do Espírito Santo
Centro de Ciências Exatas
Programa de Pós-Graduação em Matemática

TESTE DE RAIZ UNITÁRIA
(DICKEY-FULLER)
NO MODELO INAR(1).

BERNARDO IGNATOWSKI BARCELOS

Vitória ES
2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

BERNARDO IGNATOWSKI BARCELOS

TESTE DE RAIZ UNITÁRIA
(DICKEY-FULLER)
NO MODELO INAR(1).

Dissertação apresentada ao Programa de Pós-Graduação em Matemática da Universidade Federal do Espírito Santo, como parte dos requisitos para obtenção do título de Mestre em Matemática.

Orientador: Prof. Dr. VALDÉRIO A. REISEN

Vitória ES
2008

AGRADECIMENTOS

Agradeço primeiramente a Deus;

Agradeço minha família;

Ao meu orientador, o Prof. Valdério A. Reisen, por acreditar em mim e ter me orientado;

Aos Professores da banca presente, Prof. Glaura C. Franco, Prof. Regina M. de Aquino e o Prof. Ricardo S. Leite;

Ao pessoal da Matemática, em especial ao Douglas, Eleonésio, Paulo e Pedro, pela convivência nesses mais de 2 anos;

Ao pessoal da Estatística, em especial a Alyne e Nataly pela ajuda na fase final da dissertação e ao Kennedy pelo trabalho que fizemos juntos durante toda a dissertação;

Aos Professores do PPGMAT;

Ao Fundo de Apoio à Ciência e Tecnologia do Município de Vitória (FACITEC) pelo apoio financeiro;

A Isabel pela paciência de nos atender durante esses anos.

RESUMO

Esta dissertação tem como objetivo estudar o modelo Auto-Regressivo de Valores Inteiros de Ordem 1 (INAR(1)), proposto por Al-Osh e Alzaid (1987), no contexto de raiz unitária (RU). O popular teste de RU proposto por Dickey e Fuller (1979) é considerado no estudo para testar o processo INAR(1). Algumas propriedades do modelo também são obtidas. Através de ensaios de Monte Carlo, o poder do teste de Dickey-Fuller sob o processo INAR foi avaliado, e o processo AR(1) foi considerado para efeito de comparação. Os resultados empíricos evidenciaram que o desempenho do teste depende do tamanho da amostra e dos parâmetros do modelo INAR(1). Para pequenos tamanhos amostrais ($n < 200$), o uso do teste com a especificação correta do modelo resulta em um teste mais poderoso, quando comparado com o modelo AR(1). Entretanto, para grandes tamanhos amostrais ($n > 200$), o teste não mostrou diferença quanto à especificação correta ou não do modelo INAR(1), isto é, o poder do teste para esse modelo mostrou-se equivalente ao do modelo AR(1), o que é esperado do ponto de vista teórico. O estudo empírico motivou à aplicação da metodologia a uma série real.

ABSTRACT

This work has as main objective to investigate the study of the Integer-Valued Autoregressive model of order one (INAR(1)), proposed by Al-Osh and Alzaid (1987), in the context of unit root. The popular test of unit root proposed by Dickey and Fuller (1979) is considered in the study to test the INAR(1) process. Some properties of the model are also obtained. Through Monte Carlo experiments, the power of Dickey-Fuller test under the INAR process was evaluated, and the AR(1) process was considered for comparison purposes. The results have showed that the test performance depends on the size of the sample and on the parameters of the INAR(1) model. For small sample sizes ($n < 200$), the test with the correct model specification results in a more powerful test compared with the AR(1) model. However, for large sample sizes ($n > 200$), the test has showed no difference on the correct model specification, that is, the power of the test for the INAR model has given equivalent power of the AR(1) model. The empirical study has led to an application of the methodology to a real data set.

Sumário

Lista de Figuras	viii
1 Introdução	1
2 Conceitos Fundamentais	4
2.1 Espaços de Probabilidade	4
2.2 Processos Estocásticos	5
2.2.1 Funções de Autocovariância e de Autocorrelação	7
2.2.2 Exemplos de Processos Estocásticos	10
2.2.3 Função de Autocorrelação Parcial	12
2.2.4 Processo Média Móvel (MA)	13
3 Modelo Auto-Regressivo de Ordem 1 (AR(1))	18
3.1 Modelo AR(1)	18
3.2 Estimação de Mínimos Quadrados Ordinários	25
3.2.1 AR(1) sem Tendência	25
3.2.2 AR(1) com Tendência	26
4 Testes de Raízes Unitárias	28
4.1 O Teste de Dickey-Fuller (DF)	28
5 Modelo Auto-Regressivo de Valores Inteiros de Ordem 1 (INAR(1))	35
5.1 Distribuição de Poisson	35
5.2 Operador <i>thinning</i> binomial	39
5.3 Modelo INAR(1)	41
5.4 Estimador de Mínimos Quadrados Condicional	45
6 Resultados Empíricos	48
7 Aplicação	52
8 Conclusão	55
Referências Bibliográficas	58
Apêndice A	59
Apêndice B	63

Lista de Figuras

3.1	Série AR(1) simulada, $X_t = 0.5X_{t-1} + \epsilon_t$	22
3.2	FAC (à esquerda) e FACP (à direita) para a série simulada, $X_t = 0.5X_{t-1} + \epsilon_t$	22
3.3	Série AR(1) simulada, $X_t = X_{t-1} + \epsilon_t$	24
3.4	FAC (à esquerda) e FACP (à direita) para a série simulada, $X_t = X_{t-1} + \epsilon_t$	24
5.1	Série INAR(1) estacionária simulada $X_t = 0.5 \circ X_{t-1} + \epsilon_t$	44
5.2	FAC (à esquerda) e FACP (à direita) para a série INAR(1) estacionária simulada $X_t = 0.5 \circ X_{t-1} + \epsilon_t$	44
5.3	Série INAR(1) não-estacionária simulada, $X_t = X_{t-1} + \epsilon_t$	45
5.4	FAC (à esquerda) e FACP (à direita) para a série INAR(1) não-estacionária simulada $X_t = X_{t-1} + \epsilon_t$	45
6.1	Poder do teste de DF	51
7.1	Série do número anual de veículos produzidos no Brasil.	52
7.2	FAC (à esquerda) e FACP (à direita) da série do número anual de veículos produzidos no Brasil.	53
7.3	Série diferenciada do número anual de veículos produzidos no Brasil.	53
7.4	FAC (à esquerda) e FACP (à direita) da série diferenciada do número anual de veículos produzidos no Brasil.	54

Capítulo 1

Introdução

Os modelos usuais de séries temporais, tais como o processo estacionário Auto-Regressivo Média Móvel (ARMA) (Box e Jenkins, 1970), assumem que as marginais são contínuas, e em geral de distribuição Gaussiana. Essa suposição é inapropriada para modelar *séries temporais com marginais discretas*, usualmente definidas como *séries temporais de contagem*.

Esta pesquisa dedica-se ao estudo do processo Auto-Regressivo de Valor Inteiro de Ordem 1 (INAR(1)), introduzido por Al-Osh e Alzaid (1987, 1988), quando o mesmo apresenta raiz unitária. O principal objetivo é verificar o comportamento do teste Dickey-Fuller (DF) (1979) no modelo INAR(1). O teste DF tem sido um dos mais populares procedimentos para testar a presença de raiz unitária em modelos Auto-Regressivo de Ordem 1 (AR(1)) e, na literatura, existem inúmeros artigos e livros publicados com dedicação a essa importante vertente de investigação.

Diversos trabalhos sobre diferentes metodologias para séries temporais de contagem têm sido propostas, podem-se citar Wedderburn (1974), Liang e Zeger (1982), Jørgensen (1983), Zeger (1988), McCullagh e Nelder (1989), entre outros. Dentre essas, o estudo de séries temporais de valores inteiros (contagem) tem chamado a atenção por sua importância em modelar séries de diversas áreas do conhecimento. Em destaque, podem-se citar as áreas de medicina e de economia. Em ambas as áreas, séries temporais de contagem aparecem de forma natural. Como exemplo, na medicina, o número de pessoas internadas por causas epidemiológicas. Na economia é comum estudar o número de ações vendidas em um período, o número de

patentes registradas, o número de incidentes causados em navios, o número de acidentes aéreos, o número de carros roubados, o número de passageiros transportados por uma empresa entre muitos outros exemplos.

Os estudos iniciais para modelar séries temporais de valores inteiros são baseados no processo INAR(1). Inferências sobre o modelo foram considerados nos artigos de Brännäs (1994, 1995a, 1995b). Brännäs e Hall (2001) estenderam o modelo proposto por Al-Osh e Alzaid (1988) denominado como modelo Média Móvel de Valores Inteiros (INMA). Na mesma direção, Alzaid e Al-Osh (1990) propuseram o INAR de ordem p , isto é, INAR(p).

Silva e Oliveira (2004) investigaram momentos e cumulantes de grande ordem do modelo INAR(1). As autoras estenderam o estudo no contexto do domínio da frequência. Ferland, Latour e Oraichi (2005), com base no modelo INAR, estudaram o processo com erros GARCH. Jung, Kukuk e Liesenfeld (2006) investigaram, empiricamente, vários modelos para processo de contagem e analisaram como exemplo a série de números de atendimentos hospitalares por causas respiratórias. Outras informações sobre o INAR podem ser observadas em Du e Li (1991), Jung e Tremayne (2001), Zheng, Basawa e Datta (2006). As recentes produções bibliográficas, Subba Rao (1993), Kedem e Fokianos (2002) e Winkkelmann (2003) apresentam teoria, procedimentos e aplicações do processo de contagem.

O estudo de processos de contagem tem sido na atualidade uma das áreas de grandes destaques em diferentes campos de atuação. Em geral, as pesquisas dedicam-se à formulação e propriedades de modelos, estimação, testes e distribuições assintóticas dos estimadores dos modelos para diferentes distribuições marginais discretas. As metodologias bem definidas para séries temporais discretas com marginais contínuas estão sendo consideradas para series temporais de contagem. Portanto, os estudos nessa área necessitam ainda de muita investigação, tornando-a uma linha de pesquisa bastante interessante com grande impacto científico.

No que tange à raiz unitária em processo de contagem, a literatura ainda é bastante restrita. Hellström (2001) é a referência pioneira na área. O autor investigou a distribuição empírica do teste DF no modelo INAR(1). Esse artigo foi a motivação da presente dissertação com a extensão do estudo no sentido de verificar o poder do teste de DF em processos INAR(1) com marginal de Poisson de diferentes valores do parâmetro λ e do tamanho amostral. Em síntese, a investigação de Monte Carlo indicou que o teste DF apresentou bom desempenho em processos INAR(1). Esse estudo empírico suporta o uso do teste DF mesmo para séries temporais onde as distribuições marginais não seguem a família Gaussiana. Como motivação de uso da metodologia, a *série de veículos produzidos no Brasil*, período de 1957 a 2007, foi analisada.

Esta dissertação está organizada da seguinte forma: Capítulo 2 apresenta alguns conceitos de processos, funções de autocorrelação e outras definições em séries temporais. Capítulo 3 dedica-se ao modelo AR(1) e o Capítulo 4 sumariza o teste de raiz unitária. O processo de contagem INAR(1) é a motivação do Capítulo 5 onde são discutidas as propriedades desse modelo. Os Capítulos 6 e 7 apresentam os resultados numéricos e aplicação da metodologia proposta, respectivamente, e o Capítulo 8 descreve as conclusões.

Capítulo 2

Conceitos Fundamentais

Neste capítulo são introduzidos alguns conceitos como definições de espaço de probabilidade, processos estocásticos e funções de autocorrelação e autocorrelação parcial de séries temporais. Em adição, são discutidos diferentes propriedades de processos Médias Móveis finito (MA(M)) e infinito (MA(∞)). Maiores informações sobre as definições abordadas ver Fuller (1976).

2.1 Espaços de Probabilidade

Definição 2.1.1. *O conjunto Ω de todos os resultados possíveis de um experimento aleatório é chamado espaço amostral.*

Definição 2.1.2. *Seja \mathbf{A} um subconjunto do espaço amostral Ω e \mathcal{A} a coleção de todos os subconjuntos de \mathbf{A} . Diz-se que \mathbf{A} é um evento do espaço amostral e \mathcal{A} é chamado espaço de eventos.*

Definição 2.1.3. *Uma coleção não vazia \mathcal{A} de subconjuntos de Ω é dita ser uma σ -álgebra se satisfaz:*

1. Se $\mathbf{A} \in \mathcal{A}$ então $\mathbf{A}^c \in \mathcal{A}$, onde \mathbf{A}^c é o complementar de \mathbf{A} ;
2. Se $\mathbf{A}_1, \mathbf{A}_2, \dots \in \mathcal{A}$ é uma sequência enumerável de eventos em \mathcal{A} , então $\bigcup_{i=1}^{\infty} \mathbf{A}_i \in \mathcal{A}$;
3. $\emptyset \in \mathcal{A}$.

Definição 2.1.4. *Seja Ω um espaço amostral e \mathbf{A} um evento do espaço amostral. A função de probabilidade $P(\cdot)$ é uma função que satisfaz os axiomas:*

A1. $P(\mathbf{A}) \geq 0$, para todo $\mathbf{A} \in \mathcal{A}$;

A2. $P(\Omega) = 1$;

A3. Se $\mathbf{A}_1, \mathbf{A}_2, \dots \in \mathcal{A}$ é uma sequência enumerável de eventos em \mathcal{A} e $\mathbf{A}_i \cap \mathbf{A}_j = \emptyset$ para todo $i \neq j$, então $P\left(\bigcup_{i=1}^{\infty} \mathbf{A}_i\right) = \sum_{i=1}^{\infty} P(\mathbf{A}_i)$.

Definição 2.1.5. Um espaço de probabilidade é o trio (Ω, \mathcal{A}, P) , onde Ω é um conjunto não vazio, \mathcal{A} é uma σ -álgebra de eventos de Ω e P é uma função de probabilidade definida em \mathcal{A} .

Embora seja possível enumerar, teoricamente, todos os resultados de um experimento, na prática isso pode ser inviável. Para cada resultado ω do experimento será atribuído um número real $X(\omega)$.

Definição 2.1.6. Uma variável aleatória X é uma função de valor real definida em Ω tal que o conjunto $\{w : X(w) \leq x\} \subset \mathcal{A}$ é uma σ -álgebra de eventos, para todo $x \in \mathbb{R}$.

Definição 2.1.7. A função de distribuição acumulada de uma variável aleatória X , denotada por $F_X(\cdot)$, é definida como uma função com domínio em \mathbb{R} e contradomínio $[0, 1]$ que satisfaz $F_X(x) = P(\{\omega : X(\omega) \leq x\})$ para todo $x \in \mathbb{R}$.

2.2 Processos Estocásticos

Definição 2.2.1. Seja (Ω, \mathcal{A}, P) um espaço de probabilidade e T um conjunto de índices. Um processo estocástico é uma função real $X(t, \omega)$ definida em $T \times \Omega$ tal que, para cada t fixo, $X(t, \omega)$ é uma variável aleatória em (Ω, \mathcal{A}, P) .

A função $X(t, \omega)$ é frequentemente escrita como $X_t(\omega)$ ou X_t . Um processo estocástico pode ser considerado como uma família $\{X_t; t \in T\}$ de variáveis aleatórias. Se t assume valores em um intervalo contínuo real (finito ou infinito) tal que T é subconjunto de \mathbb{R} , X_t é definido como processo de parâmetro contínuo. Se t assume valores em um conjunto discreto, tipicamente, $t \in \mathbb{Z}$, então X_t é definido como processo de parâmetro discreto.

Para cada w fixo, $X(t, w)$ é uma função real de t , denominada realização do processo ou série temporal. Para cada t , X_t é uma variável aleatória associada a uma distribuição de probabilidade. Se X_t é uma variável discreta, as propriedades da variável serão especificadas por uma distribuição de probabilidade discreta. Nesse contexto, X_t é definida como *série temporal de contagem*. Se X_t é uma variável contínua, suas propriedades serão descritas por uma função de densidade de probabilidade.

Definição 2.2.2. *Sejam $X_{t_i}, i = 1, 2, \dots, n$ variáveis aleatórias com $t_1, t_2, \dots, t_n \in T$. A função de distribuição conjunta de $\{X_{t_1}, X_{t_2}, \dots, X_{t_n}\}$, é definida por*

$$F_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = P\{X_{t_1} \leq x_{t_1}, \dots, X_{t_n} \leq x_{t_n}\}. \quad (2.1)$$

Definição 2.2.3. *Um processo $\{X_t; t \in T\}$ é dito estritamente estacionário se*

$$F_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = F_{X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h}}(x_{t_1}, x_{t_2}, \dots, x_{t_n}),$$

ou seja, se a função de distribuição conjunta de $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ e $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$ são iguais para todo $h \in \mathbb{Z}$ e para todo t_1, t_2, \dots, t_n .

Se $\{X_t; t \in T\}$ é um processo estocástico estritamente estacionário com $E[|X_t|] < \infty$ para todo t , então o valor esperado de $\{X_t\}$ é constante, pois a função de distribuição conjunta é a mesma para todo t . Da mesma maneira, se $E[X_t^2] < \infty$, então a variância de $\{X_t\}$ é constante para todo t .

Definição 2.2.4. *O operador diferença Δ é definido como*

$$\Delta X_t = X_t - X_{t-1}.$$

A d -ésima diferença é dada por

$$\Delta^d X_t = \Delta^{d-1} X_t - \Delta^{d-1} X_{t-1}.$$

Definição 2.2.5. O operador de retardo, denotado por B , é definido como

$$B^k X_t = X_{t-k}.$$

Observe que

$$(1 - B)X_t = X_t - X_{t-1} = \Delta X_t, \text{ ou seja, } \Delta = 1 - B.$$

2.2.1 Funções de Autocovariância e de Autocorrelação

Definição 2.2.6. Seja $\{X_t; t \in T\}$ um processo estocástico. A média do processo é definida por $E[X_t] = \mu_t$, a variância por $\text{var}[X_t] = E[X_t - \mu_t]^2 = \sigma_t^2$ e a função de autocovariância por

$$\gamma(t_1, t_2) = \text{cov}(X_{t_1}, X_{t_2}) = E[X_{t_1} - E[X_{t_1}]]E[X_{t_2} - E[X_{t_2}]].$$

A função de autocovariância (FAC) pode ser escrita como

$$\begin{aligned} \gamma(t_1, t_2) &= E[X_{t_1} - E[X_{t_1}]]E[X_{t_2} - E[X_{t_2}]] \\ &= E[X_{t_1}X_{t_2} - X_{t_1}E[X_{t_2}] - X_{t_2}E[X_{t_1}] + E[X_{t_1}]E[X_{t_2}]] \\ &= E[X_{t_1}X_{t_2}] - E[X_{t_1}E[X_{t_2}]] - E[X_{t_2}E[X_{t_1}]] + E[X_{t_1}]E[X_{t_2}] \\ &= E[X_{t_1}X_{t_2}] - E[X_{t_2}]E[X_{t_1}] - E[X_{t_1}]E[X_{t_2}] + E[X_{t_1}]E[X_{t_2}] \\ &= E[X_{t_1}X_{t_2}] - E[X_{t_1}]E[X_{t_2}]. \end{aligned}$$

Um processo estocástico é completamente definido em um sentido probabilístico se a distribuição (2.1) é conhecida para qualquer conjunto finito de variáveis aleatórias $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$. Porém, na prática isso não acontece, e adota-se o processo fracamente estacionário, o qual é definido a seguir.

Definição 2.2.7. O processo $\{X_t; t \in T\}$ é dito fracamente estacionário se

- i. $E[X_t^2] < \infty, \forall t \in T$;
- ii. $E[X_t] = \mu$, onde μ é constante para todo $t \in T$;

iii. $\gamma(t_1, t_2) = \text{cov}\{X_{t_1}, X_{t_2}\}$ é uma função de $|t_1 - t_2|$, $\forall t_1, t_2 \in T$.

No caso do processo ser fracamente estacionário também pode-se escrever a função de autocovariância como $\gamma(h) = \text{cov}(X_{t+h}, X_t)$. Claramente,

$$\gamma(0) = \text{cov}(X_t, X_t) = E[X_t^2] - E[X_t]^2 = \text{var}[X_t].$$

Proposição 2.2.1. *Seja $\{X_t; t \in T\}$ um processo fracamente estacionário com média $E(X_t) = 0$ e função de autocovariância $\gamma(h) = E[X_t X_{t+h}]$. A função $\gamma(\cdot)$ satisfaz as propriedades:*

- i. $\gamma(0) \geq 0$;
- ii. γ é par, isto é, $\gamma(h) = \gamma(-h)$;
- iii. $|\gamma(h)| \leq \gamma(0)$.

Demonstração:

- i. $\gamma(0) = E[X_t X_t] = E[X_t^2] = \text{var}[X_t] \geq 0$;
- ii. $\gamma(h) = E[X_t X_{t+h}]$, $\forall t \in T$.
Tomando $t' = t + h$, tem-se

$$\gamma(h) = E[X_t X_{t+h}] = E[X_{t'-h} X_{t'}] = \gamma(-h).$$

iii. Dado que

$$\begin{aligned} 0 \leq E[(X_{t+h} + X_t)^2] &= E[X_{t+h}^2 + 2X_{t+h}X_t + X_t^2] \\ &= E[X_{t+h}^2] + 2E[X_{t+h}X_t] + E[X_t^2] \\ &= E[X_{t+h}X_{t+h}] + 2E[X_{t+h}X_t] + E[X_tX_t]. \end{aligned}$$

Então,

$$\gamma(0) + 2\gamma(h) + \gamma(0) \geq 0,$$

portanto,

$$\gamma(0) + \gamma(h) \geq 0. \tag{2.2}$$

Um raciocínio análogo para $E[(X_{t+h} - X_t)^2]$ mostra que

$$\gamma(0) - \gamma(h) \geq 0. \quad (2.3)$$

De (2.2) e (2.3) obtém-se $\gamma(0) \geq -\gamma(h)$ e $\gamma(0) \geq \gamma(h)$, ou seja, $\gamma(0) \geq |\gamma(h)|$. \square .

Definição 2.2.8. Seja $\{X_t; t \in T\}$ um processo estocástico fracamente estacionário.

A função de autocorrelação de $\{X_t\}$ é definida como

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}.$$

Pela Proposição 2.2.1, segue que $-1 \leq \rho(h) \leq 1$. Dado que $|\gamma(h)| \leq \gamma(0)$, então

$$|\rho(h)| = \frac{|\gamma(h)|}{|\gamma(0)|} \leq 1.$$

Considera-se como uma das principais características da função de autocorrelação o fato da mesma ser par, ou seja,

$$\rho(-h) = \frac{\gamma(-h)}{\gamma(0)} = \frac{\gamma(h)}{\gamma(0)} = \rho(h),$$

e

$$\rho(0) = \frac{\gamma(0)}{\gamma(0)} = 1.$$

Definição 2.2.9. Uma função $f : \mathbb{Z} \rightarrow \mathbb{R}$ é dita ser não-negativa definida se satisfaz

$$\sum_{j=1}^n \sum_{k=1}^n a_j a_k f(t_k - t_j) \geq 0,$$

para todo $(a_1, a_2, \dots, a_n) \in \mathbb{R}$, $(t_1, t_2, \dots, t_n) \in \mathbb{Z}$ e n inteiro positivo.

Teorema 2.2.1. A função de autocovariância de um processo fracamente estacionário $\{X_t\}$ é não-negativa definida.

Demonstração:

Sejam (t_1, t_2, \dots, t_n) um conjunto qualquer de números inteiros, (a_1, a_2, \dots, a_n) um

conjunto qualquer de números reais e $\gamma(t_k - t_j)$ a covariância entre X_{t_k} e X_{t_j} . Como a função Variância é positiva, segue que

$$\begin{aligned}
 0 \leq \text{var} \left[\sum_{j=1}^n a_j X_{t_j} \right] &= E \left[\left(\sum_{j=1}^n a_j X_{t_j} \right)^2 \right] \\
 &= E \left[\sum_{j=1}^n \sum_{k=1}^n a_j a_k X_{t_j} X_{t_k} \right] \\
 &= \sum_{j=1}^n \sum_{k=1}^n a_j a_k E[X_{t_j} X_{t_k}] \\
 &= \sum_{j=1}^n \sum_{k=1}^n a_j a_k \gamma(t_k - t_j).
 \end{aligned}$$

□

2.2.2 Exemplos de Processos Estocásticos

Ruído Branco

Um processo $\{\epsilon_t; t \in \mathbb{Z}\}$ é chamado ruído branco se as variáveis aleatórias ϵ_t são não-correlacionadas com $E[\epsilon_t] = \mu_\epsilon$ (usualmente, $\mu_\epsilon = 0$), $\text{var}[\epsilon_t] = \sigma^2$ e $\text{cov}\{\epsilon_t, \epsilon_s\} = 0$ para todo $t \neq s$. A função de autocovariância do processo é dada por:

- Para $h \neq 0$,

$$\gamma(h) = E[\epsilon_t \epsilon_{t+h}] - E[\epsilon_t]E[\epsilon_{t+h}] = \text{cov}\{\epsilon_t, \epsilon_{t+h}\} = 0, \text{ dado que } t \neq t+h.$$

- Para $h = 0$,

$$\gamma(0) = E[\epsilon_t \epsilon_t] - E[\epsilon_t]E[\epsilon_t] = E[\epsilon_t^2] - E[\epsilon_t]^2 = \text{var}[\epsilon_t] = \sigma^2.$$

Resumindo,

$$\gamma(h) = \begin{cases} \sigma^2 & \text{se } h = 0 \\ 0 & \text{se } h \neq 0. \end{cases}$$

Logo, a função de autocorrelação do processo é dada por:

$$\rho(h) = \begin{cases} 1 & \text{se } h = 0 \\ 0 & \text{se } h \neq 0. \end{cases}$$

Passeio Aleatório

Seja $\{\epsilon_t; t \geq 1\}$ sequência de variáveis aleatórias independentes e identicamente distribuídas com $E[\epsilon_t] = 0$ e $\text{var}[\epsilon_t] = \sigma^2$. O processo $\{X_t\}$ definido por $X_t = X_{t-1} + \epsilon_t$ é dito ser um passeio aleatório (*random walk*).

Assumindo $X_0 = 0$, tem-se $X_t = \epsilon_1 + \epsilon_2 + \dots + \epsilon_t$. A média e variância do processo $\{X_t\}$ são dados por

$$\begin{aligned} E[X_t] &= E[\epsilon_1 + \dots + \epsilon_t] = t\mu_\epsilon = 0, \\ \text{var}[X_t] &= \text{var}[\epsilon_1 + \dots + \epsilon_t] \\ &= \text{var}[\epsilon_1] + \dots + \text{var}[\epsilon_t] \\ &= t\sigma^2. \end{aligned}$$

A função de autocovariância do processo é dada por

$$\gamma(t_1, t_2) = \sigma^2 \cdot \min(t_1, t_2).$$

Para provar essa propriedade, assume-se primeiramente que $t_2 > t_1$, então

$$\begin{aligned}
\gamma(t_1, t_2) &= E[X_{t_1}X_{t_2}] - E[X_{t_1}]E[X_{t_2}] \\
&= E[(\epsilon_1 + \dots + \epsilon_{t_1})(\epsilon_1 + \dots + \epsilon_{t_1} + \dots + \epsilon_{t_2})] \\
&\quad - E[\epsilon_1 + \dots + \epsilon_{t_1}]E[\epsilon_1 + \dots + \epsilon_{t_2}] \\
&= E[(\epsilon_1 + \dots + \epsilon_{t_1})^2 + (\epsilon_1 + \dots + \epsilon_{t_1})(\epsilon_{t_1+1} + \dots + \epsilon_{t_2})] - t_1\mu t_2\mu \\
&= \text{var}[\epsilon_1 + \dots + \epsilon_{t_1}] + E[\epsilon_1 + \dots + \epsilon_{t_1}]^2 \\
&\quad + E[\epsilon_1 + \dots + \epsilon_{t_1}]E[\epsilon_{t_1+1} + \dots + \epsilon_{t_2}] - t_1 t_2 \mu^2 \\
&= t_1\sigma^2 + t_1^2\mu^2 + t_1\mu(t_2 - t_1)\mu - t_1 t_2 \mu^2 \\
&= t_1\sigma^2 + t_1^2\mu^2 + t_1 t_2 \mu^2 - t_1^2\mu^2 - t_1 t_2 \mu^2 \\
&= t_1\sigma^2.
\end{aligned}$$

Analogamente, para $t_1 > t_2$ tem-se $\gamma(t_1, t_2) = t_2\sigma^2$.

2.2.3 Função de Autocorrelação Parcial

Considere um processo fracamente estacionário $\{X_t, t \in \mathbb{Z}\}$. Com o objetivo de obter a correlação entre X_t e X_{t+k} após a remoção das dependências lineares das variáveis intermediárias $X_{t+1}, \dots, X_{t+k-1}$, define-se a *função de autocorrelação parcial* (FACP)

$$\phi_{kk} = \text{corr}(X_t, X_{t+k}/X_{t+1}, \dots, X_{t+k-1})$$

a qual fornece informações adicionais com respeito à estrutura de dependência do processo estacionário. Pode-se mostrar (Wei, 1990) que, para $k = 1, 2, \dots$

$$\phi_{11} = \rho_1, \quad \phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

e, em geral,

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & \rho_1 & 1 \end{vmatrix}}.$$

2.2.4 Processo Média Móvel (MA)

Definição 2.2.10. Um processo estocástico $\{X_t : t \in \mathbb{Z}\}$ é chamado de Média Móvel de ordem M , denotado por $MA(M)$, se pode ser representado como

$$X_t = \sum_{j=-M}^M \alpha_j \epsilon_{t-j}, \quad (2.4)$$

onde M é inteiro não-negativo, α_j são números reais, $\alpha_{-M} \neq 0$ e $\{\epsilon_t\}$ é um processo ruído branco.

Definição 2.2.11. Se não existir M finito tal que $\alpha_j = 0$ para todo $|j| > M$, então $\{X_t\}$ é um processo de média móvel infinita e pode ser representado como

$$X_t = \sum_{j=-\infty}^{\infty} \alpha_j \epsilon_{t-j}.$$

Definição 2.2.12. O processo definido por

$$X_t = \sum_{j=0}^M \alpha_j \epsilon_{t-j}, \quad (2.5)$$

onde $\alpha_0 \neq 0$ e $\alpha_M \neq 0$, é chamado de média móvel unilateral de ordem M .

Sendo $\{X_t\}$ um processo MA unilateral de ordem M , sua média e variância são

dadas por

$$\begin{aligned}
E[X_t] &= E \left[\sum_{j=0}^M \alpha_j \epsilon_{t-j} \right] \\
&= \sum_{j=0}^M \alpha_j E[\epsilon_{t-j}] \\
&= 0 \\
\text{var}[X_t] &= E[X_t^2] \\
&= E \left[\left(\sum_{j=0}^M \alpha_j \epsilon_{t-j} \right)^2 \right] \\
&= E \left[\sum_{i=0}^M \sum_{j=0}^M \alpha_i \alpha_j \epsilon_{t-i} \epsilon_{t-j} \right] \\
&= \sum_{i=0}^M \sum_{j=0}^M \alpha_i \alpha_j E[\epsilon_{t-i} \epsilon_{t-j}] \\
&= \sum_{j=0}^M \alpha_j^2 E[\epsilon_{t-j} \epsilon_{t-j}] \\
&= \sigma^2 \sum_{j=0}^M \alpha_j^2.
\end{aligned}$$

A função de autocovariância é:

$$\begin{aligned}
\gamma(h) &= E[X_t X_{t+h}] \\
&= E \left[\left(\sum_{j=0}^M \alpha_j \epsilon_{t-j} \right) \left(\sum_{i=0}^M \alpha_i \epsilon_{t+h-i} \right) \right] \\
&= E \left[\sum_{i=0}^M \sum_{j=0}^M \alpha_i \alpha_j \epsilon_{t-j} \epsilon_{t+h-i} \right] \\
&= \sum_{i=0}^M \sum_{j=0}^M \alpha_i \alpha_j E[\epsilon_{t-j} \epsilon_{t+h-i}].
\end{aligned}$$

- Se $|h| > M$, então $t-j \neq t+h-i$, para todo i e j satisfazendo $0 \leq i, j \leq M$, logo, $\gamma(h) = 0$.

o Se $0 \leq |h| \leq M$, então

$$\gamma(h) = \sum_{i=0}^M \sum_{j=0}^M \alpha_i \alpha_j E[\epsilon_{t-j} \epsilon_{t+h-i}] = \sum_{j=0}^{M-|h|} \alpha_j \alpha_{j+|h|} \sigma^2 = \sigma^2 \sum_{j=0}^{M-|h|} \alpha_j \alpha_{j+|h|}.$$

A partir da função de autocovariância obtem-se a função de autocorrelação

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} \frac{\sigma^2 \sum_{j=0}^{M-|h|} \alpha_j \alpha_{j+|h|}}{\sigma^2 \sum_{j=0}^M \alpha_j^2}, & \text{para } 0 \leq |h| \leq M \\ 0, & \text{para } |h| > M. \end{cases}$$

No início desta seção foi definido o processo estocástico de média móvel como uma combinação linear de uma sequência de variáveis aleatórias não-correlacionadas. É também comum falar de média móvel de um processo estocástico arbitrário $\{X_t\}$. Assim, sendo Y_t definido como

$$Y_t = \sum_{i=-M}^M \alpha_i X_{t+i}.$$

Y_t é uma média móvel do processo $\{X_t\}$.

Teorema 2.2.2. *Se $\{X_t : t \in \mathbb{Z}\}$ é um processo estocástico estacionário com média 0 e função de autocovariância $\gamma_X(h)$, então $Y_t = \sum_{i=-M}^M \alpha_i X_{t+i}$, onde M é finito e os α_i são números reais, é um processo estocástico estacionário com média 0 e função de autocovariância*

$$\gamma(h) = \sum_{i=-M}^M \sum_{j=-M}^M \alpha_i \alpha_j \gamma_X(h + j - i).$$

Demonstração:

Dado que $Y_t = \sum_{i=-M}^M \alpha_i X_{t+i}$, segue-se que

$$E[Y_t] = E \left[\sum_{i=-M}^M \alpha_i X_{t+i} \right] = \sum_{i=-M}^M \alpha_i E[X_{t+i}] = 0,$$

e

$$\begin{aligned}
\gamma_Y(h) &= E[Y_t Y_{t+h}] \\
&= E \left[\left(\sum_{i=-M}^M \alpha_i X_{t+i} \right) \left(\sum_{j=-M}^M \alpha_j X_{t+h+j} \right) \right] \\
&= E \left[\left(\sum_{i=-M}^M \sum_{j=-M}^M \alpha_i \alpha_j X_{t+i} X_{t+h+j} \right) \right] \\
&= \sum_{i=-M}^M \sum_{j=-M}^M \alpha_i \alpha_j E[X_{t+i} X_{t+h+j}].
\end{aligned}$$

Tomando $t' = t + i$ e substituindo na equação acima, obtém-se

$$\gamma_Y(h) = \sum_{i=-M}^M \sum_{j=-M}^M \alpha_i \alpha_j E[X_{t'} X_{t'-i+h+j}] = \sum_{i=-M}^M \sum_{j=-M}^M \alpha_i \alpha_j \gamma_X(h + j - i).$$

□

A seguir são apresentados importantes resultados sobre processos médias móveis infinitos. As provas são apresentadas no Apêndice B.

Teorema 2.2.3. *Seja $\{X_j\}$ uma sequência de variáveis aleatórias num espaço de probabilidade (Ω, A, P) . Se*

$$\sum_{j=-\infty}^{\infty} E\{|X_j|\} < \infty,$$

então $\sum_{j=-\infty}^{\infty} X_j$ terá uma convergência quase certa e

$$E \left\{ \sum_{j=-\infty}^{\infty} X_j \right\} = \sum_{j=-\infty}^{\infty} E\{X_j\}.$$

Teorema 2.2.4. *Sejam as sequências de números reais $\{a_j\}$ e $\{b_j\}$ absolutamente somáveis. Seja $\{X_t\}$ uma sequência de variáveis aleatórias tal que $E\{X_t^2\} \leq K$ para $t = 0, \pm 1, \pm 2, \dots$ e para algum K finito.*

Definindo

$$(Y_t, Z_t) = \left(\sum_{j=-\infty}^{\infty} a_j X_{t-j}, \sum_{k=-\infty}^{\infty} b_k X_{t-k} \right).$$

Então

$$E\{Y_t\} = \lim_{n \rightarrow \infty} \sum_{j=-n}^n a_j E\{X_{t-j}\},$$

e

$$E\{Y_t Z_t\} = \lim_{n \rightarrow \infty} \sum_{j=-n}^n \sum_{k=-n}^n a_j b_k E\{X_{t-j} X_{t-k}\}.$$

Corolário 1. *Sejam as seqüências de números reais $\{a_j\}$ e $\{b_j\}$ absolutamente somáveis. Seja $\{e_t\}$ uma seqüência de variáveis aleatórias não-correlacionadas $(0, \sigma^2)$ e defina para $t = 0, \pm 1, \pm 2, \dots$*

$$(Y_t, Z_t) = \left(\sum_{j=-\infty}^{\infty} a_j e_{t-j}, \sum_{k=-\infty}^{\infty} b_k e_{t-k} \right),$$

então $E\{Y_t\} = 0$ e $E\{Y_t Z_t\} = \sigma^2 \sum_{j=-\infty}^{\infty} a_j b_j$.

Capítulo 3

Modelo Auto-Regressivo de Ordem 1 (AR(1))

Este capítulo apresenta o processo Auto-Regressivo de ordem 1 (AR(1)), com representação MA infinita. Através dos teoremas do capítulo anterior obtém-se a variância, a função de autocovariância, condições de estacionariedade entre outros resultados do modelo. Também são apresentados os estimadores de mínimos quadrados dos parâmetros do modelo. A dedicação especial ao modelo AR(1) é justificada por ser o modelo de interesse na pesquisa proposta por esta dissertação.

3.1 Modelo AR(1)

Definição 3.1.1. *Um processo estocástico $\{X_t : t \in \mathbb{Z}\}$ é chamado Auto-Regressivo de ordem p (AR(p)) se pode ser escrito na forma*

$$X_t = \sum_{i=1}^p \phi_i X_{t-i} + \epsilon_t,$$

onde $\phi_0 \neq 0$, $\phi_p \neq 0$ e $\{\epsilon_t\}$ é um processo ruído branco.

Outra forma de escrever o processo é

$$\phi_p(B)X_t = \epsilon_t,$$

onde $\phi_p(B) = (1 - \sum_{i=1}^p \phi_i B^i)$ é chamado de *operador auto-regressivo*.

Um caso particular no processo AR(p) é o processo Auto-Regressivo de ordem 1 (AR(1)) dado por

$$X_t = \phi X_{t-1} + \epsilon_t. \quad (3.1)$$

Como

$$X_{t-1} = \phi X_{t-2} + \epsilon_{t-1}, \quad X_{t-2} = \phi X_{t-3} + \epsilon_{t-2}, \dots, \quad X_{t-N} = \phi X_{t-N+1} + \epsilon_{t-N},$$

por substituição em (3.1) tem-se

$$\begin{aligned} X_t &= \phi^2 X_{t-2} + \phi \epsilon_{t-1} + \epsilon_t \\ &= \phi^3 X_{t-3} + \phi^2 \epsilon_{t-2} + \phi \epsilon_{t-1} + \epsilon_t \\ &\vdots \\ &= \phi^N X_{t-N} + \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i} \end{aligned}$$

Portanto,

$$X_t = \phi^N X_{t-N} + \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i}. \quad (3.2)$$

Sob a hipótese de que $|\phi| < 1$ e $E[X_t^2] < \infty$, obtém-se

$$X_t = \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i} + \phi^N X_{t-N}.$$

Logo,

$$\left(X_t - \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i} \right)^2 = (\phi^N X_{t-N})^2,$$

aplicando valor esperado:

$$E \left[\left(X_t - \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i} \right)^2 \right] = E [\phi^{2N} X_{t-N}^2] = \phi^{2N} E [X_{t-N}^2],$$

com $N \rightarrow \infty$, tem-se

$$\lim_{N \rightarrow \infty} E \left[\left(X_t - \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i} \right)^2 \right] = \lim_{N \rightarrow \infty} \phi^{2N} E [X_{t-N}^2].$$

Portanto,

$$\lim_{N \rightarrow \infty} E \left[\left(X_t - \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i} \right)^2 \right] = 0,$$

pois $\phi^{2N} \rightarrow 0$ e $E[X_{t-N}^2] < k$, onde k é uma constante.

Dessa forma, se $\{X_t\}$ satisfaz (3.1) com $|\phi| < 1$, então pode-se expressá-lo como média móvel infinita de $\{\epsilon_t\}$

$$X_t = \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \phi^i \epsilon_{t-i} = \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}.$$

A média do processo AR(1) é dada por

$$E[X_t] = E \left[\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} \right] = \sum_{i=0}^{\infty} E[\phi^i \epsilon_{t-i}] = \sum_{i=0}^{\infty} \phi^i E[\epsilon_{t-i}] = 0,$$

onde o resultado final da igualdade é derivado da aplicação do Teorema 2.2.3 e do Corolário 1, pois ϕ^i absolutamente somável e $E[\epsilon_{t-i}] = 0$. A função de autocovariância do processo é dada por

$$\begin{aligned} \gamma(h) &= E[X_t X_{t+h}] \\ &= E \left[\left(\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} \right) \left(\sum_{j=0}^{\infty} \phi^j \epsilon_{t+h-j} \right) \right] \\ &= \sigma^2 \sum_{i=0}^{\infty} \phi^i \phi^{i+h} \\ &= \sigma^2 \phi^h \sum_{i=0}^{\infty} (\phi^2)^i \\ &= \frac{\phi^h}{1 - \phi^2} \sigma^2, \end{aligned}$$

onde novamente faz-se uso do Corolário 1. A partir de (3.2) é possível calcular a função de autocovariância como se segue:

$$X_t = \phi^h X_{t-h} + \sum_{i=0}^{h-1} \phi^i \epsilon_{t-i}, \quad h = 1, 2, 3, \dots \quad (3.3)$$

Como X_{t-h} é função de ϵ_t com índices menores ou iguais a $t-h$ então X_{t-h} é não-correlacionada com $\epsilon_{t-h+1}, \epsilon_{t-h+2}, \dots$. Multiplicando ambos os lados da equação (3.3) por X_{t-h} e tomando o valor esperado tem-se:

$$\begin{aligned} \gamma(h) &= E[X_t X_{t-h}] \\ &= E[\phi^h X_{t-h}^2] + E\left[X_{t-h} \sum_{i=0}^{h-1} \phi^i \epsilon_{t-i}\right] \\ &= \phi^h E[X_{t-h} X_{t-h}] + E\left[\sum_{i=0}^{h-1} \phi^i X_{t-h} \epsilon_{t-i}\right] \\ &= \phi^h \gamma(0) + \sum_{i=0}^{h-1} \phi^i E[X_{t-h} \epsilon_{t-i}] \\ &= \phi^h \gamma(0). \end{aligned}$$

Sendo $\{X_t\}$ um processo AR(1) com $|\phi| < 1$, tem-se $E[X_t] = 0$, $\text{var}[X_t] = \frac{\sigma^2}{1-\phi^2} < \infty$, $\gamma(h) = \phi^h \gamma(0)$ e $\rho(h) = \phi^h$. Logo, o processo $\{X_t\}$ é fracamente estacionário quando $|\phi| < 1$. Como exemplo de um processo AR(1) estacionário simulou-se um processo com parâmetro $\phi = 0.5$ e tamanho de amostra $n = 250$. O processo simulado e as funções de autocorrelação estimadas estão nas Figuras 3.1 e 3.2, respectivamente.

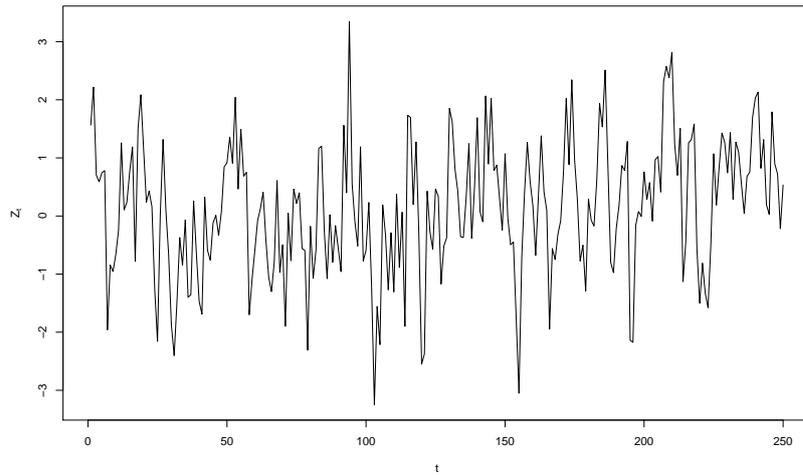


Figura 3.1: Série AR(1) simulada, $X_t = 0.5X_{t-1} + \epsilon_t$.

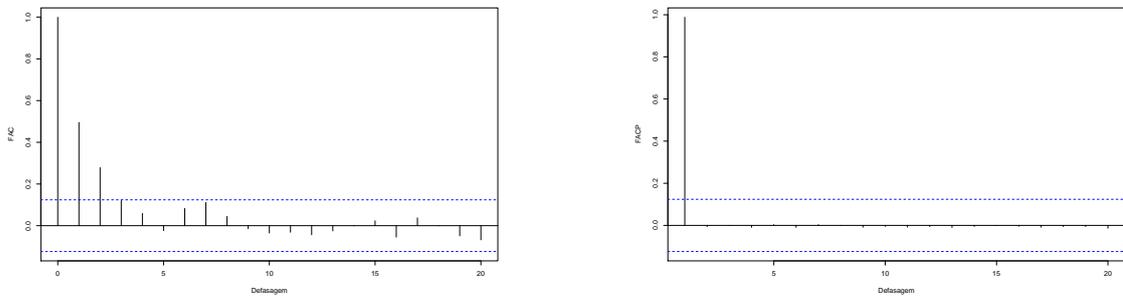


Figura 3.2: FAC (à esquerda) e FACP (à direita) para a série simulada, $X_t = 0.5X_{t-1} + \epsilon_t$.

Para $0 < \phi < 1$ a função de autocorrelação $\rho(h) = \phi^h$ decai exponencialmente, como observado na figura 3.2, e para $-1 < \phi < 0$ a função é alternadamente positiva e negativa com o valor absoluto decaindo exponencialmente.

Seja o processo AR(1) dado por $X_t = \phi X_{t-1} + \epsilon_t$ com $|\phi| \geq 1$, então

$$\begin{aligned}
 X_1 &= \epsilon_1 \\
 X_2 &= \phi\epsilon_1 + \epsilon_2 \\
 X_3 &= \phi^2\epsilon_1 + \phi\epsilon_2 + \epsilon_3 \\
 &\vdots \\
 X_N &= \phi^{N-1}\epsilon_1 + \phi^{N-2}\epsilon_2 + \dots + \phi^0\epsilon_N \\
 &= \sum_{i=0}^{N-1} \phi^i \epsilon_{N-i} \\
 &\vdots
 \end{aligned}$$

Dessa forma,

$$X_N = \sum_{i=0}^{\infty} \phi^i \epsilon_{N-i},$$

e com aplicação do Teorema 2.2.3, conclui-se que a variância é dada por

$$\text{var}[X_N] = \text{var} \left[\sum_{i=0}^{\infty} \phi^i \epsilon_{N-i} \right] = \sum_{i=0}^{\infty} \phi^{2i} \text{var}[\epsilon_{N-i}] = \sigma^2 \sum_{i=0}^{\infty} \phi^{2i}.$$

Através desse resultado, verifica-se que $\text{var}(X_N) \rightarrow \infty$ pois $|\phi| \geq 1$. E, assim, conclui-se que o processo AR(1) é não estacionário quando $|\phi| \geq 1$. Essa particularidade do modelo AR(1) é de especial interesse em diversas áreas de estudos, mais especificamente na Economia. As figuras a seguir apresentam um processo AR(1) simulado com $\phi = 1$ e suas funções de autocorrelação.

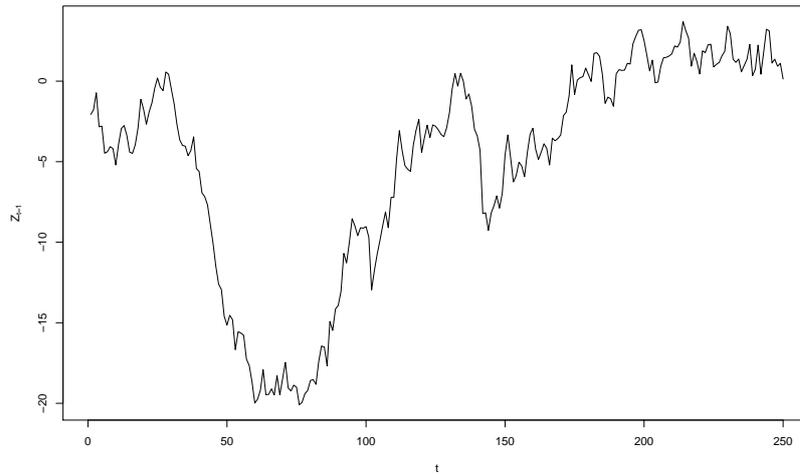


Figura 3.3: Série AR(1) simulada, $X_t = X_{t-1} + \epsilon_t$

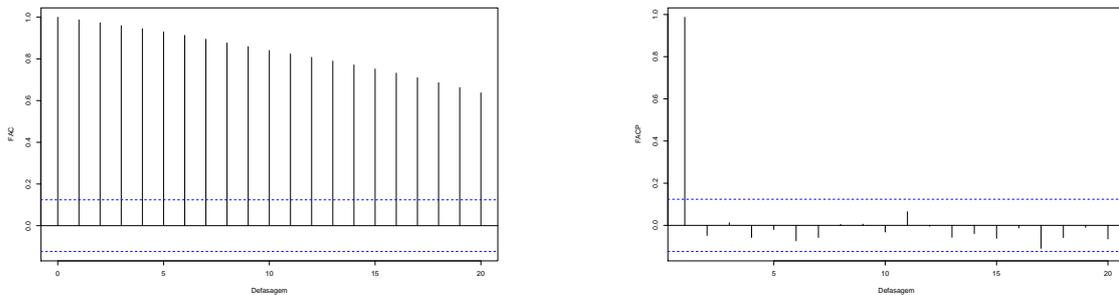


Figura 3.4: FAC (à esquerda) e FACP (à direita) para a série simulada, $X_t = X_{t-1} + \epsilon_t$.

Observa-se na Figura 3.4 que a FAC do modelo AR(1) não-estacionário apresenta um decaimento lento, diferente do AR(1) estacionário apresentado na Figura 3.2.

Além da verificação gráfica, um outro procedimento utilizado para verificar se uma série temporal é estacionária é determinar se a mesma possui raiz unitária. Para realizar o procedimento, diferentes tipos de testes podem ser aplicados. Os mais utilizados foram propostos por Dickey e Fuller (1979), Said e Dickey (1984) e Phillips e Perron (1988), e são denominados *Testes de Raiz Unitária*. Para o presente trabalho será considerado o teste apresentado por Dickey e Fuller, denominado *Teste de Dickey-Fuller* (DF), o qual é descrito no Capítulo 4.

3.2 Estimação de Mínimos Quadrados Ordinários

Esta seção dedica-se à apresentação da estimação do parâmetro ϕ do modelo AR(1). A estimação do parâmetro considera o método de mínimos quadrados ordinários aplicado ao modelo AR(1) na presença e ausência do termo de tendência determinística (intercepto ou constante), denotado por c .

3.2.1 AR(1) sem Tendência

Seja $\{X_t : t \in \mathbb{Z}\}$ definido como em (3.1), com ϵ_t sendo uma variável aleatória com $E[\epsilon_t] = 0$ e $\text{var}[\epsilon_t] = \sigma^2$, então

$$E[X_t|X_{t-1}] = E[\phi X_{t-1} + \epsilon_t|X_{t-1}] = E[\phi X_{t-1}|X_{t-1}] + E[\epsilon_t|X_{t-1}] = \phi X_{t-1}.$$

Sejam X_1, X_2, \dots, X_n , n observações do processo $\{X_t\}$, o estimador de mínimos quadrados de ϕ é o valor que minimiza a expressão

$$L = \sum_{t=2}^n (X_t - \phi X_{t-1})^2,$$

derivando L em relação a ϕ

$$\partial L / \partial \phi = -2 \sum_{t=2}^n (X_t - \hat{\phi} X_{t-1}) X_{t-1} = 0, \quad (3.4)$$

resolvendo (3.4), obtém-se o estimador de mínimos quadrados de ϕ :

$$\sum_{t=2}^n X_t X_{t-1} - \hat{\phi} \sum_{t=2}^n X_{t-1}^2 = 0,$$

então

$$\hat{\phi} = \frac{\sum_{t=2}^n X_t X_{t-1}}{\sum_{t=2}^n X_{t-1}^2}.$$

De acordo com a teoria de regressão em séries temporais, as propriedades estatísticas para $\hat{\phi}$ podem ser derivadas para o caso onde $|\phi| < 1$ (estacionários) e $\phi = 1$ (raiz unitária).

3.2.2 AR(1) com Tendência

Considerando agora o modelo AR(1) com tendência, tem-se

$$X_t = \phi X_{t-1} + \epsilon_t + c,$$

onde $\{\epsilon_t\}$ é um processo aleatório com média zero e variância σ^2 . Neste caso, tem-se que o valor esperado de X_t dado X_{t-1} é

$$E[X_t|X_{t-1}] = E[\phi X_{t-1}|X_{t-1}] + E[\epsilon_t|X_{t-1}] + E[c|X_{t-1}] = \phi X_{t-1} + c.$$

Sejam X_1, X_2, \dots, X_n , n observações do processo $\{X_t\}$, os estimadores de mínimos quadrados de ϕ e c são os valores que minimizam

$$L = \sum_{t=1}^n [X_t - (\phi X_{t-1} + c)]^2.$$

Derivando L em relação a ϕ e em relação a c obtém-se

$$\partial L / \partial \phi = -2 \sum_{t=1}^n (X_t - \hat{\phi} X_{t-1} - \hat{c}) X_{t-1} = 0, \quad (3.5)$$

$$\partial L / \partial c = -2 \sum_{t=1}^n (X_t - \hat{\phi} X_{t-1} - \hat{c}) = 0. \quad (3.6)$$

Resolvendo (3.5) e (3.6) resulta

$$\sum_{t=1}^n X_t - \hat{\phi} \sum_{t=1}^n X_{t-1} - n\hat{c} = 0,$$

daí

$$n\hat{c} = \sum_{t=1}^n X_t - \hat{\phi} \sum_{t=1}^n X_{t-1},$$

$$\hat{c} = \frac{1}{n} \left(\sum_{t=1}^n X_t - \hat{\phi} \sum_{t=1}^n X_{t-1} \right)$$

é o estimador de mínimos quadrados de c , e

$$\sum_{t=1}^n X_t X_{t-1} - \hat{\phi} \sum_{t=1}^n X_{t-1} X_{t-1} - \hat{c} \sum_{t=1}^n X_{t-1} = 0,$$

daí

$$\hat{\phi} \sum_{t=1}^n X_{t-1} X_{t-1} = \sum_{t=1}^n X_t X_{t-1} - \hat{c} \sum_{t=1}^n X_{t-1},$$

então

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

Portanto,

$$\hat{\phi} \left[\sum_{t=1}^n X_{t-1} X_{t-1} - \frac{1}{n} \left(\sum_{t=1}^n X_{t-1} \right)^2 \right] = \sum_{t=1}^n X_t X_{t-1} - \frac{1}{n} \sum_{t=1}^n X_t \sum_{t=1}^n X_{t-1}.$$

Obtém-se que:

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

é o estimador de mínimos quadrados de ϕ .

Pode-se observar que o termo de tendência causa mudança na estimativa de ϕ .

Portanto, é importante sempre observar se o processo contém ou não o termo.

Como descrito na sub-seção anterior, as propriedades de $\hat{\phi}$ também são derivadas para casos onde $|\phi| < 1$ e $\phi = 1$ (raiz unitária). Assim, com base no objetivo deste trabalho, as propriedades de estimação e teste de raiz unitária serão apresentados no próximo capítulo.

Capítulo 4

Testes de Raízes Unitárias

Estudos em testes de raízes unitárias (RU) têm sido uma área de pesquisa em evidência nas áreas de séries temporais e econometria desde a década de 70. Especial interesse tem sido dado à construção de testes para a verificação da existência de raízes unitárias no processo Auto-Regressivo. Este capítulo descreve de forma sintetizada o teste de RU proposto por Dickey-Fuller (DF) o qual é explorado, de forma empírica, no contexto do estudo proposto nesta dissertação.

4.1 O Teste de Dickey-Fuller (DF)

Para testar se uma série possui raiz unitária ou não é necessário aplicar testes de RU. Os testes RU são de grande importância, pois a maioria das séries econômicas possuem essa característica.

O teste de raiz unitária mais popular foi proposto por Dickey (1976) e Dickey e Fuller (1979), aqui referido como teste de Dickey-Fuller (DF). O teste se baseia na hipótese nula (H_0) de que o processo $\{X_t\}$ é não-estacionário (possui raiz unitária), contra a hipótese alternativa (H_1) de estacionariedade.

Como foi visto no capítulo anterior, o modelo AR(1) possui raiz unitária quando $\phi = 1$. Assim, pode-se definir a hipótese a ser testada como $H_0 : \phi = 1$ contra $H_1 : \phi < 1$. A estatística de teste sugerida por Dickey e Fuller, para o processo AR(1)

sem intercepto, é dada por

$$\tau = \frac{(\hat{\phi} - 1)}{e.p.(\hat{\phi})} = \frac{(\hat{\phi} - 1)}{\left[\hat{\sigma}_a^2 (\sum_{t=1}^n X_{t-1}^2)^{-1} \right]^{1/2}}$$

onde $\hat{\phi}$ é o valor estimado de ϕ através do método de Mínimos Quadrados Ordinários (MQO) e $\hat{\sigma}_a^2 = (n - 1)^{-1} \sum_{t=1}^n (X_t - \hat{\phi}X_{t-1})^2$.

Para o caso do modelo AR(1) com intercepto, isto é,

$$X_t = \phi X_{t-1} + \epsilon_t + c,$$

onde $\{\epsilon_t\}$ é um ruído branco com média 0 e variância σ^2 , a estatística τ é da forma

$$\tau = \frac{(\hat{\phi} - 1)}{e.p.(\hat{\phi})} = \frac{(\hat{\phi} - 1)}{\left[\hat{\sigma}_a^2 \left(\sum_{t=1}^n Z_{t-1}^2 - \frac{1}{n} (\sum_{t=1}^n Z_{t-1})^2 \right)^{-1} \right]^{1/2}},$$

onde $\hat{\phi}$ é também obtido através do método MQO e $\hat{\sigma}_a^2 = (n - 2)^{-1} \sum_{t=1}^n (X_t - \hat{c} - \hat{\phi}X_{t-1})^2$.

Sob a hipótese nula, o teste DF utiliza a distribuição τ dada por Fuller (1976).

As seguintes definições e resultados são apresentados como base para a obtenção da teoria assintótica da estatística τ .

Definição 4.1.1. *Um processo contínuo $\{W(t), 0 \leq t \leq 1\}$ é dito ser um Movimento Browniano se satisfaz:*

1. $W(0) = 0$;
2. $E[W(t)] = 0$;
3. $W(t)$ segue uma distribuição normal para cada t ;
4. Para quaisquer $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq 1$ as variáveis aleatórias $[W(t_2) - W(t_1)], [W(t_3) - W(t_2)], \dots, [W(t_n) - W(t_{n-1})]$ são independentes.

Se $\forall t, W(t) \sim N(0, t)$, então $W(t)$ é chamado *Movimento Browniano Padrão* (MBP).

Sejam a_1, a_2, \dots, a_n uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média 0 e variância σ^2 . Defina-se a função

$$F_n(x) = \begin{cases} 0 & \text{para } 0 \leq x < 1/n \\ a_1/\sqrt{n} & \text{para } 1/n \leq x < 2/n \\ (a_1 + a_2)/\sqrt{n} & \text{para } 2/n \leq x < 3/n \\ \vdots & \vdots \\ (a_1 + a_2 + \dots + a_n)/\sqrt{n} & \text{para } x = 1 \end{cases}$$

que pode ser escrita na forma

$$F_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[xn]} a_t,$$

onde $x \in [0, 1]$ e $[xn]$ é o maior inteiro menor ou igual a xn . Então,

$$F_n(x) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[xn]} a_t = \frac{\sqrt{[xn]}}{\sqrt{n}} \frac{1}{\sqrt{[xn]}} \sum_{t=1}^{[xn]} a_t.$$

Dado que $\left[\frac{\sqrt{[xn]}}{\sqrt{n}} \right] \rightarrow \sqrt{x}$ quando $n \rightarrow \infty$ e $\sum_{t=1}^{[xn]} a_t / \sqrt{[xn]} \xrightarrow{D} N(0, \sigma_a^2)$ pelo Teorema Central do Limite (TCL), segue que $F_n(x)$ converge em distribuição para $\sqrt{x}N(0, \sigma_a^2) = N(0, x\sigma_a^2)$ quando $n \rightarrow \infty$ ($F_n(x) \xrightarrow{D} N(0, x\sigma_a^2)$).

O limite da sequência de variáveis aleatórias $F_n(x)/\sigma_a$, $n=1,2,\dots$, pode ser descrito por um movimento Browniano, isto é,

$$\frac{F_n(x)}{\sigma_a} \xrightarrow{D} W(x),$$

que implica,

$$F_n(x) \xrightarrow{D} \sigma_a W(x),$$

onde $W(x)$ no tempo $t = x$ segue uma distribuição $N(0, x)$. Ou seja,

$$F_n(1) = \sum_{t=1}^n \frac{a_t}{\sqrt{n}} \xrightarrow{D} \sigma_a W(1), \quad (4.1)$$

onde $W(1)$ segue uma distribuição $N(0, 1)$.

Seja $Z_t = a_1 + a_2 + \dots + a_t$ e $Z_0 = 0$. Reescrevendo $F_n(x)$ como

$$F_n(x) = \begin{cases} 0 & \text{para } 0 \leq x \leq 1/n \\ Z_1/\sqrt{n} & \text{para } 1/n \leq x \leq 2/n \\ Z_2/\sqrt{n} & \text{para } 2/n \leq x \leq 3/n \\ \vdots & \vdots \\ Z_n/\sqrt{n} & \text{para } x = 1. \end{cases}$$

De (4.1) tem-se:

$$\frac{Z_n}{\sqrt{n}} \xrightarrow{D} \sigma_a W(1), \quad (4.2)$$

então,

$$\frac{Z_n^2}{n} \xrightarrow{D} \sigma_a^2 [W(1)^2]. \quad (4.3)$$

Tem-se também que

$$\int_0^1 F_n(x) dx = \frac{1}{n} \frac{Z_1}{\sqrt{n}} + \frac{1}{n} \frac{Z_2}{\sqrt{n}} + \dots + \frac{1}{n} \frac{Z_{n-1}}{\sqrt{n}} = n^{-3/2} \sum_{t=1}^n Z_{t-1},$$

pois $\int_0^1 F_n(x) dx$ é a soma da área sob o gráfico de $F_n(x)$.

De (4.2) tem-se:

$$n^{-3/2} \sum_{t=1}^n Z_{t-1} = \int_0^1 F_n(x) dx \xrightarrow{D} \sigma_a \int_0^1 W(x) dx. \quad (4.4)$$

Analogamente,

$$\int_0^1 [F_n(x)]^2 dx = \frac{1}{n} \left(\frac{Z_1}{\sqrt{n}} \right)^2 + \frac{1}{n} \left(\frac{Z_2}{\sqrt{n}} \right)^2 + \dots + \frac{1}{n} \left(\frac{Z_{n-1}}{\sqrt{n}} \right)^2 = n^{-2} \sum_{t=1}^n Z_{t-1}^2$$

e

$$n^{-2} \sum_{t=1}^n Z_{t-1}^2 = \int_0^1 [F_n(x)]^2 dx \xrightarrow{D} \sigma_a^2 \int_0^1 [W(x)]^2 dx. \quad (4.5)$$

Dado que

$$\begin{aligned} Z_t^2 &= (Z_{t-1} + a_t)^2 = Z_{t-1}^2 + 2Z_{t-1}a_t + a_t^2, \\ Z_{t-1}a_t &= \frac{1}{2}[Z_t^2 - Z_{t-1}^2 - a_t^2], \end{aligned}$$

então,

$$\sum_{t=1}^n Z_{t-1}a_t = \frac{1}{2} \left[\sum_{t=1}^n (Z_t^2 - Z_{t-1}^2 - a_t^2) \right] = \frac{1}{2}[Z_n^2 - Z_0^2] - \frac{1}{2} \sum_{t=1}^n a_t^2 = \frac{1}{2}Z_n^2 - \frac{1}{2} \sum_{t=1}^n a_t^2.$$

Portanto

$$n^{-1} \sum_{t=1}^n Z_{t-1}a_t = \frac{1}{2} \left[\frac{Z_n^2}{n} \right] - \frac{1}{2} \left[\frac{\sum_{t=1}^n a_t^2}{n} \right] \xrightarrow{D} \frac{1}{2} \sigma_a^2 [W(1)^2] - \frac{1}{2} \sigma_a^2 = \frac{1}{2} \sigma_a^2 \{ [W(1)]^2 - 1 \}, \quad (4.6)$$

pois $\frac{Z_n^2}{n} \xrightarrow{D} \sigma_a^2 [W(1)^2]$ e $\left[\frac{\sum_{t=1}^n a_t^2}{n} \right] \xrightarrow{P} \sigma_a^2$.

Teorema 4.1.1. *Considere o modelo AR(1) dado por $X_t = \phi X_{t-1} + a_t$, onde $\{a_t\}$ é ruído branco de média 0 e variância σ^2 . Então*

$$\hat{\tau} \xrightarrow{D} \frac{\frac{1}{2} \{ [W(1)]^2 - 1 \}}{\left\{ \int_0^1 [W(x)]^2 dx \right\}^{1/2}}, \quad (4.7)$$

onde $W(x)$ é o MBP, ou seja, para cada t , $W(x) \sim N(0, t)$.

Demonstração:

O estimador de mínimos quadrados de ϕ é dado por

$$\hat{\phi} = \frac{\sum_{t=1}^n X_{t-1} X_t}{\sum_{t=1}^n X_{t-1}^2}.$$

Note que

$$\begin{aligned} \hat{\phi} &= \frac{\sum_{t=1}^n X_{t-1} X_t}{\sum_{t=1}^n X_{t-1}^2} = \frac{\sum_{t=1}^n X_{t-1} (X_{t-1} + a_t)}{\sum_{t=1}^n X_{t-1}^2} \\ &= \frac{\sum_{t=1}^n X_{t-1}^2}{\sum_{t=1}^n X_{t-1}^2} + \frac{\sum_{t=1}^n X_{t-1} a_t}{\sum_{t=1}^n X_{t-1}^2} \\ &= 1 + \frac{\sum_{t=1}^n X_{t-1} a_t}{\sum_{t=1}^n X_{t-1}^2}. \end{aligned}$$

Seja τ a estatística dada por:

$$\hat{\tau} = \frac{\hat{\phi} - 1}{S_{\hat{\phi}}} = \frac{\hat{\phi} - 1}{\left[\hat{\sigma}_a^2 \left(\sum_{t=1}^n X_{t-1}^2 \right)^{-1} \right]^{1/2}},$$

onde $\hat{\sigma}_a^2 = \frac{\sum_{t=1}^n (X_t - \hat{\phi} X_{t-1})^2}{(n-1)}$.

Logo,

$$\begin{aligned} \hat{\tau} &= \frac{(\hat{\phi} - 1) \left[\sum_{t=1}^n X_{t-1}^2 \right]^{1/2}}{[\hat{\sigma}_a^2]^{1/2}} \\ &= \frac{\sum_{t=1}^n X_{t-1} a_t \left[\sum_{t=1}^n X_{t-1}^2 \right]^{1/2}}{\sum_{t=1}^n X_{t-1}^2 [\hat{\sigma}_a^2]^{1/2}} \\ &= \frac{n^{-1} \sum_{t=1}^n X_{t-1} a_t}{n^{-1} \left[\sum_{t=1}^n X_{t-1}^2 \right]^{1/2} [\hat{\sigma}_a^2]^{1/2}} \\ &= \frac{n^{-1} \sum_{t=1}^n X_{t-1} a_t}{\left[n^{-2} \sum_{t=1}^n X_{t-1}^2 \right]^{1/2} [\hat{\sigma}_a^2]^{1/2}} \xrightarrow{D} \frac{\frac{1}{2} \sigma_a^2 \{ [W(1)]^2 - 1 \}}{\left\{ \sigma_a^2 \int_0^1 [W(x)]^2 dx \right\}^{1/2} [\sigma_a^2]^{1/2}} = \frac{\frac{1}{2} \{ [W(1)]^2 - 1 \}}{\left\{ \int_0^1 [W(x)]^2 dx \right\}^{1/2}}. \end{aligned}$$

onde o resultado acima segue de (4.5) e (4.6) □

O teste usando (4.7) é chamado *teste de Dickey-Fuller*. A distribuição dessa estatística está tabelada em Fuller (1976).

Teorema 4.1.2. *Seja o modelo AR(1) com intercepto dado por $X_t = \phi X_{t-1} + a_t + c$, onde $\{a_t\}$ é ruído branco de média 0 e variância σ^2 . Então,*

$$\hat{\tau}_\mu \xrightarrow{D} \frac{\frac{1}{2} \{[W(1)]^2 - 1\} - W(1) \int_0^1 W(x) dx}{\left\{ \int_0^1 [W(x)]^2 dx - \left[\int_0^1 W(x) dx \right]^2 \right\}^{1/2}}$$

Esta prova foi omitida neste trabalho. Para maiores detalhes, Wei (2006). Neste caso a notação padrão utilizada para τ é τ_μ . A distribuição de τ_μ afasta-se mais da normal do que τ .

Outros testes de raízes unitárias, como Dickey-Fuller Aumentado (ADF) (Said e Dickey, 1984) e Phillips-Perron (PP) (Perron, 1988) também são bastante utilizados, entretanto não serão aqui descritos pois não fazem parte da proposta de pesquisa do presente trabalho.

Capítulo 5

Modelo Auto-Regressivo de Valores Inteiros de Ordem 1 (INAR(1))

Recentemente tem havido um crescente interesse no estudo de séries temporais de valor inteiro não-negativo e, em especial, em séries de contagem. Como exemplos de séries de contagem pode-se citar: número de chamadas telefônicas chegadas a uma central em um período de uma hora, número mensal de passageiros de uma determinada companhia de aviação, o número diário de hóspedes em um hotel, o número de internações em um hospital por causas respiratórias, entre outras. Alguns modelos ARMA de contagem com erros de Poisson foram propostos por McKenzie (1985, 1988). Como destaque para o modelo Auto-Regressivo de Valores Inteiros de ordem 1 (INAR(1)), apresentado por Al-Osh e Alzaid (1987). O modelo INAR baseia-se no operador *thinning* binomial proposto por Steutel e Van Harn (1979). Du e Li (1991) generalizaram o modelo INAR para ordem p , isto é, INAR(p). O operador *thinning* é definido na Seção 5.2, a definição e as propriedades do modelo INAR(1) são dadas nas Seções 5.3 e 5.4, respectivamente.

5.1 Distribuição de Poisson

Dentre as distribuições discretas abordadas em séries temporais, a de Poisson é a mais observada na teoria e na prática. Assim, antes de introduzir a teoria do processo INAR é apresentada uma revisão sobre a distribuição de Poisson. Maiores informações sobre distribuições de probabilidade pode-se citar James (1981) e Meyer

(2000). Uma variável aleatória \mathbf{X} possui distribuição Poisson com parâmetro λ se a distribuição de probabilidade de \mathbf{X} é dada por

$$\begin{aligned} P(\mathbf{X} = x) &= \frac{e^{-\lambda} \lambda^x}{x!}; & x = 0, 1, 2, \dots \\ &= 0 & ; \quad c.c. \end{aligned}$$

A distribuição de Poisson pode ser desenvolvida a partir da distribuição Binomial como se segue. Seja \mathbf{Y} uma variável aleatória com distribuição Binomial com parâmetros n (número de ensaios) e p (probabilidade de sucessos), onde sua distribuição de probabilidade é dada por

$$P(\mathbf{Y} = y) = \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y}; \quad y = 0, 1, 2, \dots$$

Tomando $np = b$ tem-se $p = b/n$ e $1-p = 1 - b/n = (n-b)/n$. Daí:

$$\begin{aligned} P(\mathbf{Y} = y) &= \frac{n(n-1)(n-2)\dots(n-y+1)}{y!} \left[\frac{b}{n}\right]^y \left[\frac{n-b}{n}\right]^{n-y} \\ &= \frac{b^y}{y!} \left[1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{y-1}{n}\right)\right] \left[1 - \frac{b}{n}\right]^n \left[1 - \frac{b}{n}\right]^{-y}. \end{aligned}$$

Fazendo $n \rightarrow \infty$ e $p \rightarrow 0$ de modo que $np = b$ permaneça fixo, tem-se que $\left(1 - \frac{1}{n}\right) \rightarrow 1$, $\left(1 - \frac{2}{n}\right) \rightarrow 1$, ..., $\left(1 - \frac{y-1}{n}\right) \rightarrow 1$, $\left(1 - \frac{b}{n}\right)^{-y} \rightarrow 1$. Dado que $\left(1 - \frac{b}{n}\right)^n \rightarrow e^{-b}$ quando $n \rightarrow \infty$ então

$$P(Y = y) = \frac{b^y e^{-b}}{y!},$$

que é a distribuição de Poisson com parâmetro b , isto é, $Y \sim Po(b)$.

Para verificar que a expressão acima é realmente uma distribuição de probabilidade, esta deve satisfazer $\sum_{x=0}^{\infty} P(X = x) = 1$.

De fato, como

$$\begin{aligned}
 P(X = 0) &= e^{-\lambda}; \\
 P(X = 1) &= \lambda e^{-\lambda}; \\
 P(X = 2) &= \frac{\lambda^2 e^{-\lambda}}{2!}; \\
 &\vdots \\
 P(X = n) &= \frac{\lambda^n e^{-\lambda}}{n!}; \\
 &\vdots
 \end{aligned}$$

então

$$\sum_{x=0}^{\infty} P(X = x) = e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2 e^{-\lambda}}{2!} + \dots = e^{-\lambda} \left[1 + \lambda + \frac{\lambda^2}{2!} + \dots \right] = e^{-\lambda} e^{\lambda} = 1.$$

A função geratriz de momentos da distribuição de Poisson é dada por

$$\begin{aligned}
 M_x(t) &= E[e^{tx}] \\
 &= \sum_{x_i=0}^{\infty} e^{tx_i} p(x = x_i) \\
 &= \sum_{x_i=0}^{\infty} e^{tx_i} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \\
 &= e^{-\lambda} + e^t \frac{e^{-\lambda} \lambda}{1!} + e^{2t} \frac{e^{-\lambda} \lambda^2}{2!} + e^{3t} \frac{e^{-\lambda} \lambda^3}{3!} + \dots \\
 &= e^{-\lambda} \left[1 + e^t \lambda + e^{2t} \frac{\lambda^2}{2!} + e^{3t} \frac{\lambda^3}{3!} + \dots \right] \\
 &= e^{-\lambda} \left[1 + (e^t \lambda)^1 + \frac{(e^t \lambda)^2}{2!} + \frac{(e^t \lambda)^3}{3!} + \dots \right] \\
 &= e^{-\lambda} e^{\lambda e^t} \\
 &= e^{\lambda e^t - \lambda} \\
 &= e^{\lambda(e^t - 1)}.
 \end{aligned}$$

Com a função geratriz de momentos, pode-se facilmente calcular os momentos da

distribuição da variável X em torno da origem. Estes são dados por:

$$E[X] = \frac{d}{dt} M_x(t) |_{t=0} = e^{\lambda(e^t-1)} \lambda e^t |_{t=0} = 1 \cdot \lambda \cdot 1 = \lambda.$$

$$E[X^2] = \frac{d^2}{dt^2} M_x(t) |_{t=0} = e^{\lambda(e^t-1)} (\lambda e^t)^2 + e^{\lambda(e^t-1)} \lambda e^t |_{t=0} = 1\lambda^2 + 1\lambda = \lambda + \lambda^2.$$

Daí

$$\text{var}[X] = E[X^2] - (E[X])^2 = \lambda + \lambda^2 - \lambda^2 = \lambda.$$

$$\begin{aligned} E[X^3] &= \frac{d^3}{dt^3} M_x(t) |_{t=0} = e^{\lambda(e^t-1)} (\lambda e^t)^3 + e^{\lambda(e^t-1)} 2\lambda e^t \lambda e^t + e^{\lambda(e^t-1)} (\lambda e^t)^2 + e^{\lambda(e^t-1)} \lambda e^t |_{t=0} \\ &= \lambda + 3\lambda^2 + \lambda^3. \end{aligned}$$

Teorema 5.1.1. *Sejam X_1, X_2, \dots, X_n variáveis aleatórias independentes com distribuição de Poisson de parâmetros $\lambda_1, \lambda_2, \dots, \lambda_n$, respectivamente. Seja $Z = X_1 + X_2 + \dots + X_n$, então Z tem distribuição de Poisson com parâmetro $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_n$.*

Demonstração:

A função geratriz de momentos de X_i é dada por

$$M_{X_i}(t) = e^{\lambda_i(e^t-1)}$$

Como

$$\begin{aligned} M_Z(t) &= E[e^{tZ}] = E[e^{t(X_1+\dots+X_n)}] = E[e^{tX_1} \cdot e^{tX_2} \cdot \dots \cdot e^{tX_n}] \\ &= E[e^{tX_1}] \cdot \dots \cdot E[e^{tX_n}] = M_{X_1}(t) \cdot M_{X_2}(t) \cdot \dots \cdot M_{X_n}(t), \end{aligned}$$

então,

$$M_Z(t) = e^{\lambda_1(e^t-1)} \cdot \dots \cdot e^{\lambda_n(e^t-1)} = e^{(\lambda_1+\dots+\lambda_n)(e^t-1)},$$

que é a função geratriz de momentos de uma variável aleatória com distribuição de Poisson de parâmetro $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_n$. \square

5.2 Operador *thinning* binomial

Definição 5.2.1. *Seja X variável aleatória de valor inteiro não-negativo e seja α número real tal que $\alpha \in [0,1]$. O operador *thinning*, denotado por \circ , é definido como*

$$\alpha \circ X = \sum_{j=1}^X Y_j$$

onde Y_j é uma sequência de variáveis aleatórias de valor inteiro não-negativo independentes e identicamente distribuídas (i.i.d.), chamada série de contagem, independente de X e tal que $E[Y_j] = \alpha$, $\text{var}[Y_j] = \sigma^2$, $E[Y_j^3] = \gamma$ e $E[Y_j^4] = \kappa$, isto é, momentos finitos.

A definição 5.2.1 refere-se à versão geral do operador *thinning* binomial, proposto por Steutel e Van Harn (1979), onde Y_j é uma sequência de variáveis aleatórias de valor inteiro não-negativo i.i.d., independente de X tal que $P(Y_j = 1) = 1 - P(Y_j = 0) = \alpha$.

As propriedades do operador *thinning* foram estudadas por Gauthier e Latour (1994), e Silva e Oliveira (2004, 2005), sendo apresentadas na proposição abaixo.

Proposição 5.2.1. *(Propriedades do operador *thinning*) Seja $X_i (i = 1, \dots, 4)$, uma variável aleatória identicamente distribuída e $\alpha_i (i = 1, \dots, 4)$ uma constante real não-negativa. Considere a série de contagem, $Y_{i,j} (j = 1, \dots, X_i)$, com $\alpha_i \circ X_i (i = 1, \dots, 4)$, sendo mutuamente independente e identicamente distribuída, independente de X_i e tal que $E[Y_{j,i}] = \alpha_i$ e $\text{var}[Y_{j,i}] = \sigma_i^2$, $E[Y_{j,i}^3] = \gamma_i$ e $E[Y_{j,i}^4] = \kappa_i$. Portanto,*

- (i) $0 \circ X_1 = 0$,
- (ii) $1 \circ X_1 = X_1$,
- (iii) $\alpha_1 \circ (\alpha_2 \circ X_2) = (\alpha_1 \alpha_2) \circ X_2$,
- (iv) $\alpha \circ (X_1 + X_2) = \alpha \circ X_1 + \alpha \circ X_2$
- (v) $E[\alpha_1 \circ X_1] = \alpha_1 E[X_1]$,
- (vi) $E[(\alpha_1 \circ X_1) X_2] = \alpha_1 E[X_1 X_2]$,

- (vii) $E[(\alpha_1 \circ X_1)X_2X_3] = \alpha_1 E[X_1X_2X_3]$,
- (viii) $E[(\alpha_1 \circ X_1)^2] = \alpha_1^2 E[X_1^2] + \sigma_1^2 E[X_1]$,
- (ix) $E[(\alpha_1 \circ X_1)^2X_2] = \alpha_1^2 E[X_1^2X_2] + \sigma_1^2 E[X_1X_2]$,
- (x) $E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)] = \alpha_1\alpha_2 E[X_1X_2]$,
- (xi) $E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)X_3] = \alpha_1\alpha_2 E[X_1X_2X_3]$,
- (xii) $E[(\alpha_1 \circ X_1)^3] = \alpha_1^3 E[X_1^3] + 3\alpha_1\sigma_1^2 E[X_1^2] + (\gamma_1 - 3\alpha_1\sigma_1^2 - \alpha_1^3)E[X_1]$,
- (xiii) $E[(\alpha_1 \circ X_1)^2(\alpha_2 \circ X_2)] = \alpha_1^2\alpha_2 E[X_1^2X_2] + \sigma_1^2\alpha_2 E[X_1X_2]$,
- (xiv) $E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)(\alpha_3 \circ X_3)] = \alpha_1\alpha_2\alpha_3 E[X_1X_2X_3]$,
- (xv) $E[(\alpha_1 \circ X_1)^4] = \alpha_1^4 E[X_1^4] + 6\alpha_1^2\sigma_1^2 E[X_1^3] + (3\sigma_1^4 - 12\sigma_1^2\alpha_1^2 + 4\alpha_1\gamma_1 - 4\alpha_1^4)E[X_1^2] + (\kappa_1 - 3\sigma_1^4 + 6\sigma_1^2 + \alpha_1^2 - 4\alpha_1\gamma_1 + 3\alpha_1^4)E[X_1]$,
- (xvi) $E[(\alpha_1 \circ X_1)^3X_2] = \alpha_1^3 E[X_1^3X_2] + 3\alpha_1\sigma_1^2 E[X_1^2X_2] + (\gamma_1 - 3\alpha_1\sigma_1^2 - \alpha_1^3)E[X_1X_2]$,
- (xvii) $E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)] = \sigma_1^2\sigma_2^2 E[X_1X_2] + \alpha_1^2\alpha_2^2 E[X_1^2X_2^2] + \alpha_1^2\sigma_2^2 E[X_1^2X_2] + \alpha_2^2\sigma_1^2 E[X_1X_2^2]$,
- (xviii) $E[(\alpha_1 \circ X_1)^2X_2X_3] = \alpha_1^2 E[X_1^2X_2X_3] + \sigma_1^2 E[X_1X_2X_3]$,
- (xix) $E[(\alpha_1 \circ X_1)X_2X_3X_4] = \alpha_1 E[X_2X_3X_4]$,
- (xx) $E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)^3] = \alpha_1\alpha_2^3 E[X_1X_2^3] + 3\alpha_1\alpha_2\sigma_2^2 E[X_1X_2^2] + (\alpha_1\gamma_2 - 3\alpha_1\alpha_2\sigma_2^2 - \alpha_1\alpha_2^3)E[X_1X_2]$,
- (xxi) $E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)(\alpha_3 \circ X_3)^2] = \alpha_1\alpha_2\alpha_3^2 E[X_1X_2X_3] + \alpha_1\alpha_2\sigma_3^2 E[X_1X_2X_3] - \alpha_1\alpha_2\alpha_3^2 E[X_1X_2X_3] + \alpha_1\alpha_2\alpha_3^2 E[X_1X_2X_3^2]$,
- (xxii) $E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)(\alpha_3 \circ X_3)(\alpha_4 \circ X_4)] = \alpha_1\alpha_2\alpha_3\alpha_4 E[X_1X_2X_3X_4]$.

As provas das propriedades do operador thinning são apresentadas no Apêndice C.

5.3 Modelo INAR(1)

Definição 5.3.1. Um processo estocástico $\{X_t : t \in \mathbb{Z}\}$ é chamado *Auto-Regressivo com valores inteiros de ordem p (INAR(p))* se pode ser escrito na forma

$$X_t = \sum_{i=1}^p \alpha_i \circ X_{t-1} + \epsilon_t, \quad (5.1)$$

onde α_i é um número real tal que $\alpha_1 \neq 0$ e $\alpha_p \neq 0$, $\alpha_i \in [0,1]$, $\{\epsilon_t\}$ é um processo aleatório não correlacionado de valores inteiros não-negativos com média μ e variância σ^2 e " \circ " representa o operador *thinning binomial*.

Quando o operador *thinning* binomial é aplicado em 5.1, a realização de X_t tem duas componentes aleatórias:

- Os elementos sobreviventes do processo no tempo $t - 1$, cada um com probabilidade de sobrevivência α_i , denotado por $\alpha_i \circ X_{t-1}$;
- Os elementos que entram no processo no intervalo $(t - 1, t]$, chamados de elementos de entrada, ϵ_t .

O modelo Auto-Regressivo de Valores Inteiros não-negativos de ordem 1 (INAR(1)), introduzido por Al-Osh e Alzaid (1987), é expresso na forma mais simples de (5.1), isto é, como

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad (5.2)$$

onde $\alpha \in [0,1]$ e $\{\epsilon_t\}$ é uma sequência não correlacionada de variáveis aleatórias com valores inteiros não-negativos de média μ e variância σ^2 finita.

A distribuição marginal do modelo (5.2) pode ser expressa em termos da sequência

$\{\epsilon_t\}$ como

$$\begin{aligned}X_1 &= \epsilon_1 \\X_2 &= \alpha \circ \epsilon_1 + \epsilon_2 \\X_3 &= \alpha \circ (\alpha \circ \epsilon_1 + \epsilon_2) + \epsilon_3 \\&= \alpha \circ (\alpha \circ \epsilon_1) + \alpha \circ \epsilon_2 + \epsilon_3 \\&= \alpha^2 \circ \epsilon_1 + \alpha \circ \epsilon_2 + \epsilon_3 \\&\vdots \\X_n &\stackrel{d}{=} \sum_{j=0}^{\infty} \alpha^j \circ \epsilon_{n-j}.\end{aligned}$$

A média e a variância do processo X_t como definido em (5.2) são dadas por

$$\begin{aligned}E[X_t] &= E[\alpha \circ X_{t-1} + \epsilon_t] \\&= \alpha E[X_{t-1}] + \mu \\&= \alpha E[\alpha \circ X_{t-2} + \mu] + \mu \\&= \alpha^2 E[X_{t-2}] + \alpha\mu + \mu \\&= \alpha^t E[X_0] + \mu \sum_{j=0}^{t-1} \alpha^j\end{aligned}$$

e

$$\begin{aligned}
\text{var}[X_t] &= \text{var}[\alpha \circ X_{t-1} + \epsilon_t] \\
&= \text{var}[\alpha \circ X_{t-1}] + \text{var}[\epsilon_t] \\
&= E[(\alpha \circ X_{t-1})^2] - E[\alpha \circ X_{t-1}]^2 + E[\epsilon_t^2] - E[\epsilon_t]^2 \\
&= \alpha^2 E[X_{t-1}^2] + \text{var}[X_{t-1}]E[X_{t-1}] - \alpha^2 E[X_{t-1}]^2 + \sigma^2 + \mu^2 - \mu^2 \\
&= \alpha^2 [E[X_{t-1}^2] - E[X_{t-1}]^2] + \alpha(1 - \alpha)E[X_{t-1}] + \sigma^2 \\
&= \alpha^2 \text{var}[X_{t-1}] + \alpha(1 - \alpha)E[X_{t-1}] + \sigma^2 \\
&= \alpha^2 [\alpha^2 \text{var}[X_{t-2} + \alpha(1 - \alpha)E[X_{t-2}] + \sigma^2] + \alpha(1 - \alpha)E[X_{t-1}] + \sigma^2 \\
&= \alpha^4 \text{var}[X_{t-2}] + \alpha(1 - \alpha) [\alpha^3 E[X_{t-2}] + \alpha E[X_{t-1}]] + (\alpha^2 + 1)\sigma^2 \\
&\vdots \\
&= \alpha^{2t} \text{var}[X_0] + (1 - \alpha) \sum_{j=1}^t \alpha^{2j-1} E[X_{t-j}] + \sigma^2 \sum_{j=1}^t \alpha^{2(j-1)}.
\end{aligned}$$

A função de autocovariância é dada por

$$\begin{aligned}
\gamma(k) &= \text{Cov}(X_{t-k}, X_t) \\
&= \text{Cov}(X_{t-k}, \alpha^k \circ X_{t-k}) + \text{cov}\left(X_{t-k}, \sum_{j=0}^{k-1} \alpha^j \circ \epsilon_{t-j}\right) \\
&= E[X_{t-k}(\alpha^k \circ X_{t-k})] - E[X_{t-k}]E[\alpha^k \circ X_{t-k}] + E\left[X_{t-k} \sum_{j=0}^{k-1} \alpha^j \circ \epsilon_{t-j}\right] - \\
&\quad - E[X_{t-k}]E\left[\sum_{j=0}^{k-1} \alpha^j \circ \epsilon_{t-j}\right] \\
&= \alpha^k E[X_{t-k}^2] - \alpha^k E[X_{t-k}]^2 + \sum_{j=0}^{k-1} \alpha^j E[X_{t-k} \epsilon_{t-j}] - \sum_{j=0}^{k-1} \alpha^j E[X_{t-k}]E[\epsilon_{t-j}] \\
&= \alpha^k \text{var}[X_{t-k}] + \sum_{j=0}^{k-1} \alpha^j \text{Cov}(X_{t-k}, \epsilon_{t-j}), \\
&= \alpha^k \gamma(0)
\end{aligned}$$

e a função de autocorrelação por

$$\rho(k) = \gamma(k)/\gamma(0) = \alpha^k.$$

Logo, o processo INAR(1) dado por $X_t = \alpha \circ X_{t-1} + \epsilon_t$ é estacionário quando $|\alpha| < 1$ e não estacionário quando $|\alpha| = 1$. O comportamento de um processo INAR(1) estacionário é apresentado, empiricamente, nas Figuras 5.1 e 5.2.

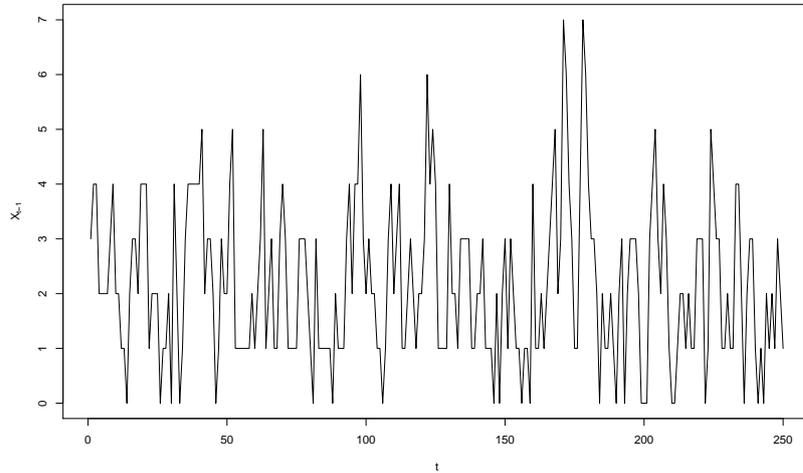


Figura 5.1: Série INAR(1) estacionária simulada $X_t = 0.5 \circ X_{t-1} + \epsilon_t$.

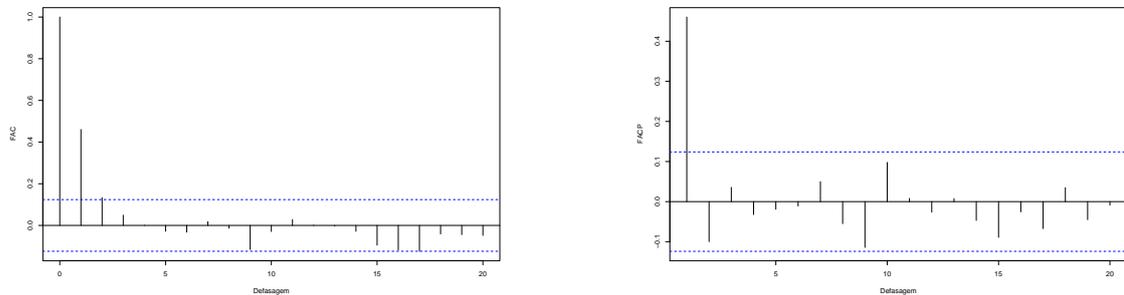


Figura 5.2: FAC (à esquerda) e FACP (à direita) para a série INAR(1) estacionária simulada $X_t = 0.5 \circ X_{t-1} + \epsilon_t$.

Portanto, a função de autocorrelação de um processo INAR(1) estacionário, assim como no processo AR(1), decai exponencialmente como pode ser observado na Figura 5.2. Como ilustração de um processo INAR(1) não-estacionário simulou-se o mesmo considerando $\alpha = 1$. As Figuras 5.3 e 5.4 apresentam o gráfico do processo e as funções de autocorrelação, onde observa-se que a FAC (Figura 5.4) apresenta um decaimento lento, assim como no processo AR(1). Entretanto, teoricamente, a FAC do processo INAR(1) só apresenta valores positivos.

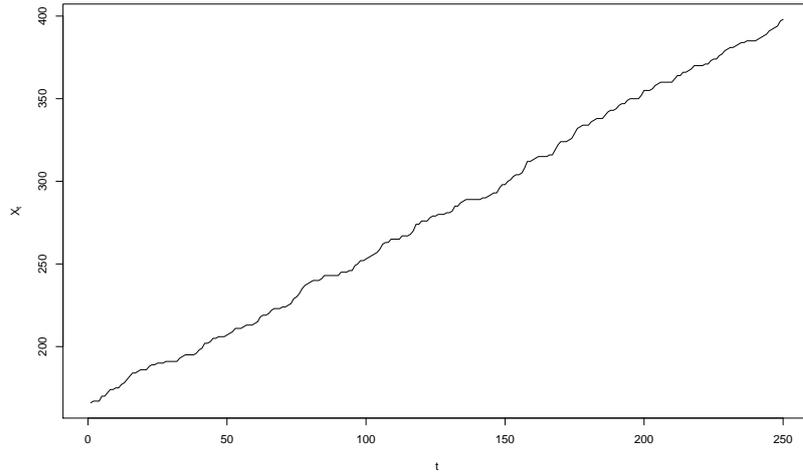


Figura 5.3: Série INAR(1) não-estacionária simulada, $X_t = X_{t-1} + \epsilon_t$.

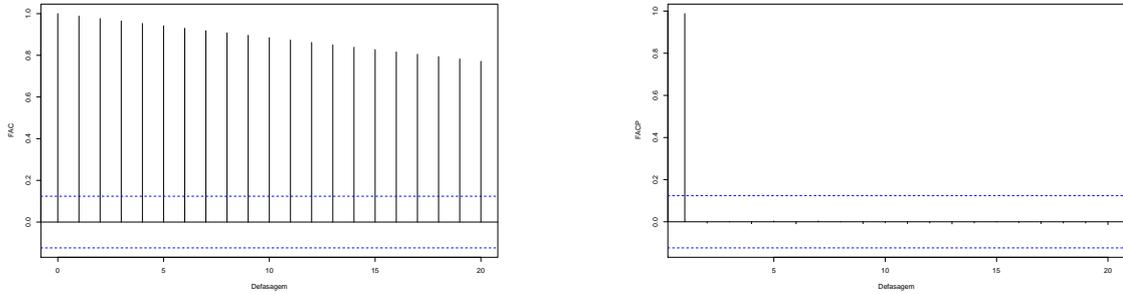


Figura 5.4: FAC (à esquerda) e FACP (à direita) para a série INAR(1) não-estacionária simulada $X_t = X_{t-1} + \epsilon_t$.

5.4 Estimador de Mínimos Quadrados Condicional

Seja $\{X_t, t \in \mathbb{Z}\}$ definido como em (5.2) e suponha-se que $\{\epsilon_t\}$ tenha distribuição de Poisson com parâmetro λ , então, tem-se

$$\begin{aligned}
 E[X_t | X_{t-1}] &= E[\alpha \circ X_{t-1} + \epsilon_t | X_{t-1}] \\
 &= E[\alpha \circ X_{t-1}] + E[\epsilon_t | X_{t-1}] \\
 &= E[Y_1 + \dots + Y_{X_{t-1}} | X_{t-1}] + E[\epsilon_t] \\
 &= E[Y_1] + \dots + E[Y_{X_{t-1}}] + \lambda \\
 &= \alpha + \dots + \alpha + \lambda \\
 &= \alpha X_{t-1} + \lambda.
 \end{aligned}$$

Sejam X_1, X_2, \dots, X_n , n observações do processo $\{X_t\}$, o estimador de mínimos quadrados condicional de α e λ são os valores que minimizam

$$L = \sum_{t=1}^n [X_t - (\alpha X_{t-1} + \lambda)]^2.$$

Derivando L em relação a α e em relação a λ obtem-se

$$\partial L / \partial \alpha = -2 \sum_{t=1}^n (X_t - \hat{\alpha} X_{t-1} - \hat{\lambda}) X_{t-1} = 0, \quad (5.3)$$

$$\partial L / \partial \lambda = -2 \sum_{t=1}^n (X_t - \hat{\alpha} X_{t-1} - \hat{\lambda}) = 0. \quad (5.4)$$

Resolvendo (5.3) e (5.4) resulta que

$$\sum_{t=1}^n X_t - \hat{\alpha} \sum_{t=1}^n X_{t-1} - n\hat{\lambda} = 0.$$

Tem-se, então, que

$$n\hat{\lambda} = \sum_{t=1}^n X_t - \hat{\alpha} \sum_{t=1}^n X_{t-1}.$$

Portanto,

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

é o estimador de mínimos quadrados de λ , e

$$\sum_{t=1}^n X_t X_{t-1} - \hat{\alpha} \sum_{t=1}^n X_{t-1} X_{t-1} - \hat{\lambda} \sum_{t=1}^n X_{t-1} = 0.$$

Logo,

$$\hat{\alpha} \sum_{t=1}^n X_{t-1} X_{t-1} = \sum_{t=1}^n X_t X_{t-1} - \hat{\lambda} \sum_{t=1}^n X_{t-1}.$$

Então

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

Portanto

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

Obtem-se que

$$\hat{\alpha} = \frac{\sum_{t=1}^n X_tX_{t-1} - \frac{1}{n} \sum_{t=1}^n X_t \sum_{t=1}^n X_{t-1}}{\sum_{t=1}^n X_{t-1}^2 - \frac{1}{n} \left(\sum_{t=1}^n X_{t-1} \right)^2},$$

é o estimador de mínimos quadrados de α .

Capítulo 6

Resultados Empíricos

Este capítulo dedica-se à investigação empírica do teste de DF com objetivo de verificar o poder 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 os contextos de diferentes tamanhos amostrais e valores dos parâmetros do modelo INAR(1), incluindo o valor da variância do processo, isto é, o valor esperado da distribuição de Poisson. O modelo AR(1) também é considerado para efeito de comparação.

Para o estudo empírico proposto, amostras foram simuladas dos modelos AR(1) e INAR(1), isto é, processos $Y_t = \alpha Y_{t-1} + \mu_t + \lambda$, $X_t = \alpha \circ X_{t-1} + \epsilon_t$, respectivamente. Nos modelos, $\{\mu_t\}$ é um processo independente e identicamente distribuído (i.i.d.) $N(0, \lambda)$ e o processo $\{\epsilon_t\}$ é i.i.d Poisson com parâmetro λ . Considerar a mesma variância em ambos os processos permite também o estudo do efeito desse parâmetro no desempenho do teste. Os valores adotados para os parâmetros foram $\alpha = 0.8, 0.9, 0.95, 1$ e $\lambda = 0.05, 0.1, 0.5, 3$, com amostras de tamanho $n = 100, 200$ e 500 . Em cada caso foram realizadas 1000 replicações e o nível de significância usado nos testes foi fixado em 5%. Os procedimentos de simulações dos modelos foram obtidos pelo programa R.

Os resultados na tabela a seguir referem-se a fração de rejeição do teste sob a hipótese nula $H_0 : \alpha = 1$ contra a alternativa $H_1 : \alpha < 1$.

Tabela 6.1: Poder do teste de DF nos modelos AR(1) e INAR(1)

1a: $\alpha = 1.0$				1b: $\alpha = 0.95$			
n	λ	AR	INAR	n	λ	AR	INAR
100	0.05	0.060	0.044	100	0.05	0.089	0.539
	0.1	0.052	0.055		0.1	0.071	0.259
	0.5	0.042	0.047		0.5	0.083	0.094
	3	0.055	0.038		3	0.085	0.075
200	0.05	0.045	0.038	200	0.05	0.193	0.839
	0.1	0.048	0.039		0.1	0.199	0.544
	0.5	0.039	0.035		0.5	0.176	0.180
	3	0.055	0.050		3	0.199	0.191
500	0.05	0.055	0.051	500	0.05	0.833	1.000
	0.1	0.043	0.041		0.1	0.835	0.987
	0.5	0.041	0.047		0.5	0.837	0.830
	3	0.049	0.064		3	0.845	0.828

1c: $\alpha = 0.90$				1d: $\alpha = 0.80$			
n	λ	AR	INAR	n	λ	AR	INAR
100	0.05	0.189	0.951	100	0.05	0.665	1.000
	0.1	0.196	0.842		0.1	0.628	0.996
	0.5	0.193	0.199		0.5	0.677	0.858
	3	0.185	0.196		3	0.645	0.645
200	0.05	0.661	1.000	200	0.05	1.000	1.000
	0.1	0.648	0.999		0.1	0.998	1.000
	0.5	0.623	0.704		0.5	0.998	0.999
	3	0.630	0.620		3	0.999	0.998
500	0.05	1.000	1.000	500	0.05	1.000	1.000
	0.1	1.000	1.000		0.1	1.000	1.000
	0.5	1.000	1.000		0.5	1.000	1.000
	3	1.000	1.000		3	1.000	1.000

A Tabela 6.1a mostra o nível de significância empírico do teste. Observa-se que em ambos os modelos o teste indica valores em torno do nível nominal de 5%. Valores diferentes da variância λ não indicam nenhum comportamento diferenciado. Isso evidencia que tanto para o modelo AR(1) quanto para o INAR(1), o teste apresenta bom desempenho, isto é, o teste pode ser usado em ambas situações de processos.

Com a variação do valor do coeficiente α , percebe-se claramente o efeito do valor da variância nos testes (Tabelas 6.1b-6.1d). Para o processo com variância pequena, o teste sob o modelo INAR(1) apresenta significativa rejeição da hipótese nula, espe-

cialmente para pequenos tamanhos amostrais. Entretanto, o teste comporta-se de forma similar em ambos processos quando considera-se tamanhos amostrais maiores. Essa evidência empírica mostra que para tamanhos pequenos de séries temporais, o teste tem que ser executado com a consideração do modelo especificado corretamente. Por exemplo, no caso onde $n = 100$ e $\lambda < 3$, em geral, o poder do modelo AR(1) é bem inferior ao poder do modelo INAR(1). Entretanto, para tamanhos $n > 200$, em geral, o teste DF pode ser aplicado em processo de contagem como forma aproximada do processo AR(1). Os resultados mostrados na tabela 6.1 podem ser facilmente visualizados através das curvas empíricas do poder do teste nos gráficos apresentados na Figura 6.1.

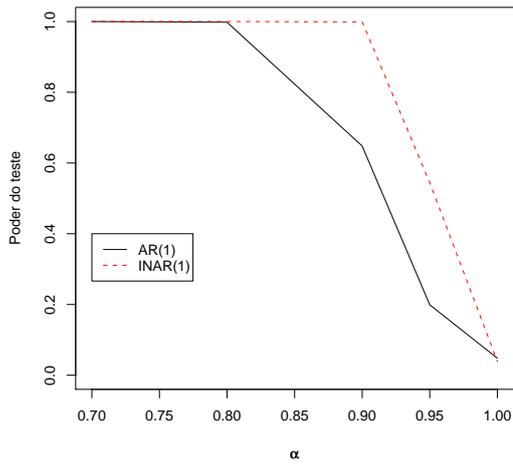
Os resultados empíricos observados comprovam a teoria assintótica da equivalência dos momentos para ambos os modelos. De acordo com a seção 2.2.2, se Y_t é um passeio aleatório, então $E[Y_t] = t\mu = 0$ (para $\mu = 0$) e $\text{var}[Y_t] = t\sigma^2$ onde μ e σ^2 são, respectivamente, a média e a variância do ruído branco. Para o processo INAR(1), de parâmetro α , a seção 5.3 mostra que

$$E[X_t] = \alpha^t E[X_0] + \mu \sum_{j=0}^{t-1} \alpha^j$$

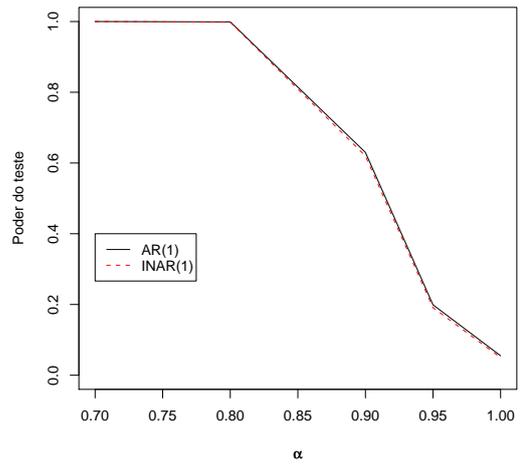
e

$$\text{var}[X_t] = \alpha^{2t} \text{var}[X_0] + (1 - \alpha) \sum_{j=1}^t \alpha^{2j-1} E[X_{t-j}] + \sigma^2 \sum_{j=1}^t \alpha^{2(j-1)}.$$

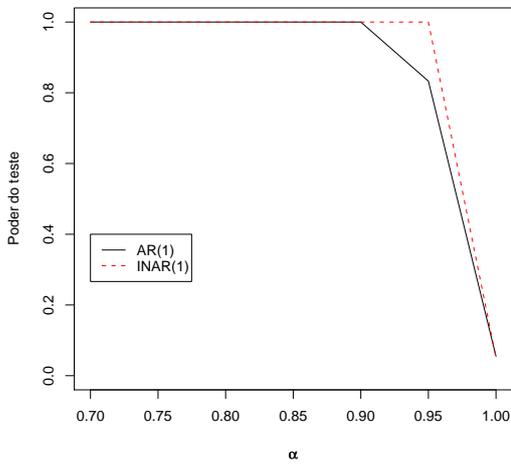
Note que para grandes valores de t e $\alpha = 1$, $E[X_t] \approx E[Y_t]$ e $\text{var}[X_t] \approx \text{var}[Y_t]$. Esses resultados teóricos explicam os empíricos apresentados nesta seção, isto é, para grandes tamanhos amostrais o teste DF comporta-se de maneira semelhante para ambos os modelos.



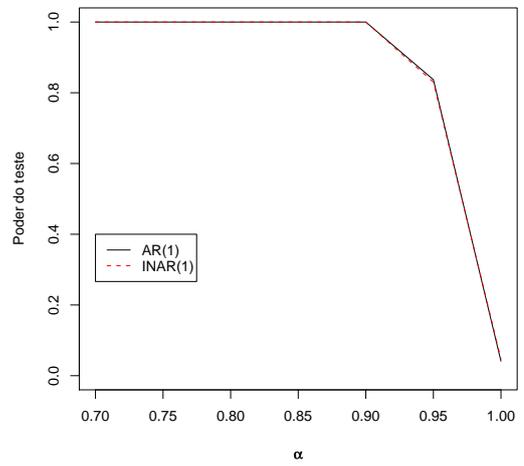
(a) $n=200$ e $\lambda = 0.1$



(b) $n=200$ e $\lambda = 3$



(c) $n=500$ e $\lambda = 0.05$



(d) $n=500$ e $\lambda = 0.5$

Figura 6.1: Poder do teste de DF

Capítulo 7

Aplicação

Com base nos estudos de simulações apresentados no capítulo anterior, as metodologias estudadas são aplicadas em dados reais. Os dados observados foram obtidos no Ipeadata e são referentes ao número anual de carros de passeio produzidos no Brasil (1957-2007), denotada por $\{X_t, t = 1, \dots, 51\}$. Os gráficos da série observada e as funções de autocorrelação são apresentadas a seguir, Figuras 7.1 e 7.2 respectivamente.

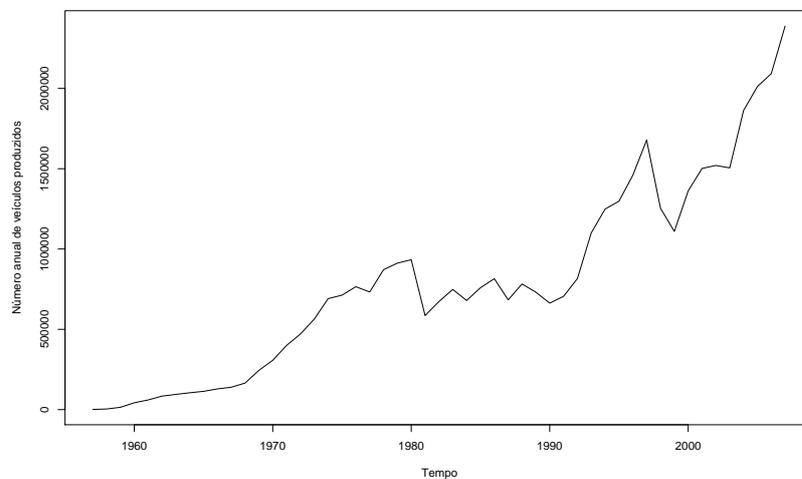


Figura 7.1: Série do número anual de veículos produzidos no Brasil.

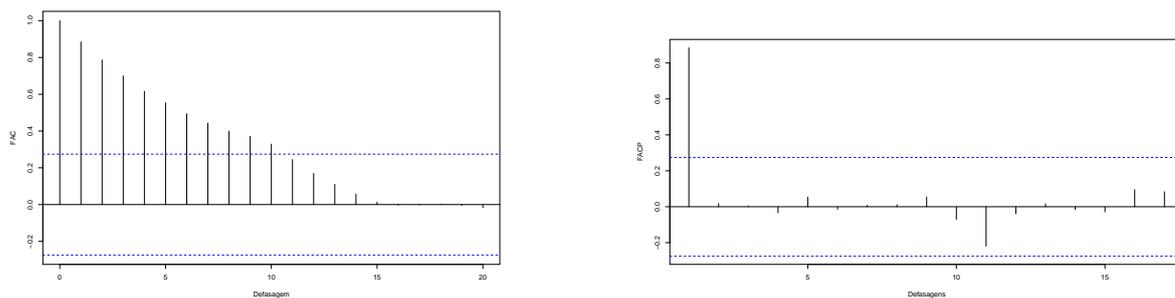


Figura 7.2: FAC (à esquerda) e FACP (à direita) da série do número anual de veículos produzidos no Brasil.

O gráfico da série estudada (Figura 7.1) apresenta uma tendência positiva, o que fornece indícios de presença de raiz unitária. A evidência de RU também é observada pela FAC (Figura 7.2).

Pelo teste de DF não rejeitou-se a hipótese nula de raiz unitária ($H_0 : \alpha = 1$), ao nível de significância de 5%. O valor obtido para a estatística $\hat{\tau}$ foi $-1,1112$, que quando comparado ao valor crítico $\tau = -2,93$ (Fuller, 1996, pp.642) encontra-se na região de não rejeição de H_0 . Então, aplica-se a primeira diferença na série de número anual de veículos produzidos (Figura 7.3).



Figura 7.3: Série diferenciada do número anual de veículos produzidos no Brasil.

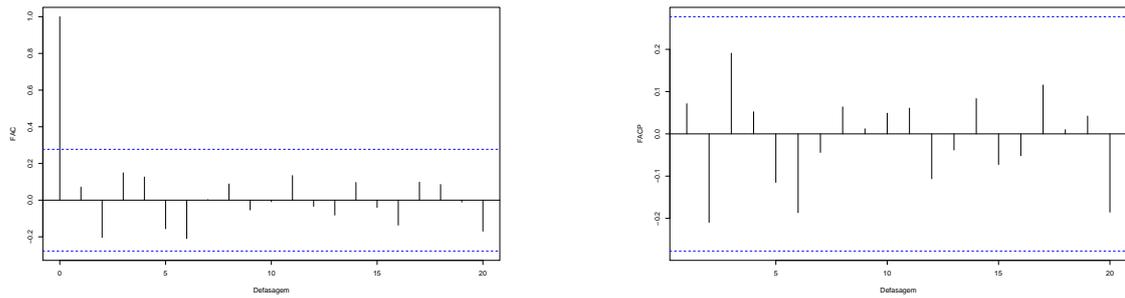


Figura 7.4: FAC (à esquerda) e FACP (à direita) da série diferenciada do número anual de veículos produzidos no Brasil.

Ao aplicar a 1ª diferença na série de produção de veículos, pode-se observar (Figura 7.3) que a série diferenciada é estacionária. As funções de autocorrelação (Figura 7.4) comportam-se como um ruído aleatório, indicando assim, que a série integrada de ordem 1 é descrita da mesma forma que um passeio aleatório (ver Capítulo 2). Assim é possível escrever o número anual de veículos produzido como um processo AR(1) integrado de ordem 1.

Capítulo 8

Conclusão

Pelos resultados empíricos apresentado nesse trabalho conclui-se que o teste de DF, para $\alpha = 1$, se comporta de maneira semelhante tanto no modelo INAR(1) quanto no modelo AR(1), para diferentes tamanho de amostra e diferentes λ . Variando o coeficiente α , percebe-se que o modelo INAR(1) apresenta significativa rejeição da hipótese nula para pequenos valores de λ . Porém, para grandes tamanhos amostrais e grandes valores de λ o teste se comporta de forma similar em ambos modelos. O estudo empírico mostrou que para pequenos tamanhos de séries temporais, o teste deve ser executado com a consideração do modelo especificado corretamente, pois o poder do modelo AR(1) é inferior ao poder do modelo INAR(1), mas com o aumento do tamanho da série o teste de DF pode ser aplicado em processo de contagem como forma aproximada do modelo AR(1).

Referências Bibliográficas

AL-OSH, M. A. e ALZAID, A. A.(1987). First-order integer valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis*, 8, 261:75.

AL-OSH, M. A. e ALZAID, A. A. (1988). Integer-valued moving average (INMA) process. *Statistical Papers*, 29, 281:300.

AL-OSH, M. A. e ALZAID, A. A.(1990) An integer-valued p-th order autoregressive structure (INAR(p)) process. *Journal of Applied Probability*, 27, 314:24.

BOX, G. E. P. e JENKINS, G. M. (1970). *Time series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

BRÄNNÄS, K. (1994). Estimation and Testing in Integer-valued AR(1) Models. *Umea Economic Studies* No. 335.

BRÄNNÄS, K. (1995a). Prediction and control of a time series count data model. *Int. J.Forecasting*, 11, 263:270.

BRÄNNÄS, K. (1995b). Explanatory variables in the AR(1) count data model. *Umea Economic Studies*, No. 381.

DICKEY, D.A e FULLER, W.A. (1979). Distribution of estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427:431.

DU, J.-G. e LI, Y. (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis*, 12,129:42.

FERLAND, R., LATOUR, A. e ORAICHI, D. (2006). Integer-Valued Garch Process. *Journal of Time Series Analysis*, 27, 6, 923:942.

FULLER, W.A. (1976). *Introduction to Statistical Time Series*. John Wiley, New York.

GAUTHIER, G. e LATOUR, A. (1994). Convergence Forte des Estimateurs des Parametres d'un Processus GENAR(p). *Ann. Sci. Math. Quebec*, 18,

49:71.

HELLSTRÖM, J. (2001). Unit root testing in integer-valued AR(1) models. *Economics Letters*, 70, 9:14.

JAMES, B.R. (1981). *Probabilidade: um curso em nível intermediário*. Impa, Rio de Janeiro.

JUNG, R.C., KUKUK, M. e LIESENFELD, R. (2006). Time series of count data: modeling, estimation and diagnostics. *Computational Statistics & Data Analysis*, 51(4), 2350:2364.

JØRGENSEN, B. (1983). Maximum likelihood estimation and large-sample inference for generalized linear and nonlinear regression models. *Biometrika*, 70, 19:28.

LEE, Y. e NELDER, J.A. (1996). Hierarchical Generalized Linear Models. *Journal of the Royal Statistical Society B*, 58, 619:678.

LIANG, K.Y. e ZEGER, S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13:22.

JUNG, R.C. e TREMAYNE, A. R. (2001). Testing for serial dependence in time series models of counts. *Journal Of Time Series Analysis*, 24, 1, 65:84.

KEDEM, B. e FOKIANOS, K. (2002) *Regression Models for Time Series Analysis*. Wiley, USA, 2nd.

MCCULLAGH, P. e NELDER, J.A. (1989). *Generalized Linear Models*. Chapman and Hall, Londres, 2nd.

MCKENZIE, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin* 21, 645:650.

MCKENZIE, E. (1988). Some ARMA models for dependent sequences of Poisson counts. *Adv. Appl. Prob.*, 20, 822:835.

MEYER, P. (2000). *Probabilidade. Aplicações à Estatística*. LTC, Rio de Janeiro.

PERRON, P. (1988). Trends and Random Walks in Macroeconomic Time Series. *Journal of Economic Dynamics and Control*, 12, 297:332.

SILVA, I. e OLIVEIRA, M.E.(2004). Seleção de ordem em modelos INAR. *Actas do X Congresso Anual da SPE*, Portugal.

STEUTAL, F. e VAN HARN, K. (1979). Discrete analogues of self-

decomposability and stability. *Ann. Prob.* 7, 893:9.

SUBBA RAO, T. (1994). *Developments in Time Series Analysis*, Chapman-Hall.

WEDDERBURN, R.W.M.(1974). Quasi-likelihood functions, generalized linear models and the Gauss-Newton method. *Biometrika*, 61, 439:447.

WEI, W. W. S. (2006). *Times Series Analysis*. Addison-Wesley Publishing Company, Inc., U.S.

ZEGER, S.L.(1988). A regression model for time series of counts. *Biometrika*, 75, 4, 621:629.

ZHENG, H., BASAWA, I.V. e DATTA, S. (2006). Inference for p th-Order Random Coefficient Integer-Valued Autoregressive Processes. *Journal Of Time Series Analysis*, Vol. 27, No. 3, 411:440.

Apêndice A

Sequências Absolutamente Somáveis

Esta seção apresenta uma revisão de alguns importantes resultados da convergência de sequências infinitas de números reais.

Definição 7.1.1

Uma sequência infinita, denotada por $\{a_j\}_{j=1}^{\infty}$ ou $\{a_j\}$, é uma função $a : \mathbb{N} \rightarrow \mathbb{R}$ que associa a cada número natural n um número real a_n chamado o n -ésimo termo da sequência.

Definição 7.1.2

Uma sequência duplamente infinita é uma função $a : \mathbb{Z} \rightarrow \mathbb{R}$ que associa a cada número inteiro n um número real a_n . A sequência duplamente infinita é denotada por $\{a_j\}_{j=-\infty}^{\infty}$.

Dada uma sequência $\{a_j\}_{j=1}^{\infty}$ de números reais, a partir dela forma-se uma nova sequência $\{s_j\}_{j=1}^{\infty}$ onde

$$s_1 = a_1, s_2 = a_1 + a_2, \dots, s_j = a_1 + a_2 + \dots + a_j = \sum_{i=1}^j a_i$$

os s_j são chamados de somas parciais da série $\sum a_j$.

Se existir o limite

$$\lim_{n \rightarrow \infty} s_n = \lim_{n \rightarrow \infty} \sum_{j=1}^n a_j$$

diz-se que a série $\sum a_j$ é convergente, se o limite não é definido caso contrário é divergente.

Se

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n |a_j|$$

é finito, a série $\sum a_j$ é dita absolutamente convergente e $\sum_{j=1}^{\infty} |a_j| < \infty$

Se o limite

$$\lim_{n \rightarrow \infty} \sum_{j=-n}^n |a_j|$$

de uma sequência duplamente infinita existir e for finita, então a série é absolutamente convergente, denota-se $\sum_{j=-\infty}^{\infty} |a_j| < \infty$ e a sequência $\{a_j\}$ é dita ser absolutamente somável.

Propriedades da Sequência Absolutamente Somável

1. Se $\{a_j\}$ é absolutamente somável, então

$$\sum_{j=-\infty}^{\infty} a_j^2 < \infty.$$

De fato, como $\{a_j\}$ é absolutamente somável, então

$$\sum_{j=-\infty}^{\infty} |a_j| < \infty.$$

Seja $B = (\dots + |a_{-2}| + |a_{-1}| + |a_0| + |a_1| + |a_2| + \dots) < \infty$, então

$$B^2 = (\dots + |a_{-2}| + |a_{-1}| + |a_0| + |a_1| + |a_2| + \dots)^2 < \infty.$$

Como

$$\begin{aligned} \sum_{j=-\infty}^{\infty} a_j^2 &= (\dots + a_{-2}^2 + a_{-1}^2 + a_0^2 + a_1^2 + a_2^2 + \dots) \\ &= (\dots + |a_{-2}|^2 + |a_{-1}|^2 + |a_0|^2 + |a_1|^2 + |a_2|^2 + \dots) \\ &\leq (\dots + |a_{-2}|^2 + |a_{-1}|^2 + |a_0|^2 + |a_1|^2 + |a_2|^2 + k) \end{aligned}$$

onde k é tal que

$$(\dots + |a_{-2}|^2 + |a_{-1}|^2 + |a_0|^2 + |a_1|^2 + |a_2|^2 + k) = (\dots + |a_{-2}| + |a_{-1}| + |a_0| + |a_1| + |a_2| + \dots)^2.$$

Portanto

$$\sum_{j=-\infty}^{\infty} a_j^2 = B^2 < \infty.$$

2. Dado $\{a_j\}$ e $\{b_j\}$ seqüências absolutamente somáveis então as seqüências $\{a_j + b_j\}$ e $\{a_j b_j\}$ são absolutamente somáveis.

Observe que

$$\sum_{j=-\infty}^{\infty} |a_j + b_j| \leq \sum_{j=-\infty}^{\infty} (|a_j| + |b_j|) = \sum_{j=-\infty}^{\infty} |a_j| + \sum_{j=-\infty}^{\infty} |b_j| < \infty,$$

pois $\{a_j\}$ e $\{b_j\}$ são absolutamente somáveis e

$$\sum_{j=-\infty}^{\infty} |a_j b_j| = \sum_{j=-\infty}^{\infty} (|a_j| |b_j|) \leq \sum_{j=-\infty}^{\infty} (|a_j| + |b_j|)^2 < \infty,$$

pois $\{(|a_j| + |b_j|)^2\}$ é o quadrado de uma seqüência absolutamente somável.

3. A convolução de duas seqüências absolutamente somáveis $\{a_j\}$ e $\{b_j\}$ definida

por $c_j = \sum_{k=-\infty}^{\infty} a_k b_{j-k}$ é absolutamente somável. Pois

$$\sum_{j=-\infty}^{\infty} |c_j| \leq \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |a_k| |b_{j-k}| \leq \sum_{j=-\infty}^{\infty} |a_k| \sum_{s=-\infty}^{\infty} |b_s| < \infty.$$

Apêndice B

Provas dos Teoremas do Capítulo 2

Demonstração do Teorema 2.2.3:

Definindo

$$Y_k(\omega) = \sum_{j=-k}^k |X_j(\omega)|,$$

como $Y_k(\omega)$ é a soma de variáveis aleatórias, ela é uma variável aleatória bem definida.

Uma vez que $Y_0(\omega) = |X_0(\omega)|$, $Y_1(\omega) = |X_{-1}(\omega)| + Y_0(\omega) + |X_1(\omega)|, \dots$, $Y_k(\omega) = |X_{-k}(\omega)| + Y_{k-1}(\omega) + |X_k(\omega)|$, e assim por diante, tem-se que $\{Y_k(\omega)\}$ é uma sequência de variáveis aleatórias não-decrescente.

Como a esperança E é uma função não-decrescente, então a sequência

$$\sum_{j=-n}^n E\{X_j\} = E\left\{\sum_{j=-n}^n X_j\right\}$$

não seria limitada, o que é uma contradição.

Logo, a sequência $\{X_j\}$ é absolutamente somável e a série $\sum_{j=-n}^n X_j$ é quase sempre convergente e definida (uma vez que é a soma de variáveis aleatórias).

Como por hipótese, $E\{|X_j|\}$ é finita pra cada j , então

$$E\{Y_k\} = E\left\{\sum_{j=-k}^k |X_j|\right\} = \sum_{j=-k}^k E|X_j| < \infty, \forall k.$$

Aplicando o Teorema da Convergência Monótona (James, 1981) à sequência $\{Y_k\}$ tem-se que

$$\begin{aligned}
 E \left\{ \sum_{j=-\infty}^{\infty} |X_j| \right\} &= E \left\{ \lim_{n \rightarrow \infty} \sum_{j=-n}^n |X_j| \right\} \\
 &= \lim_{n \rightarrow \infty} E \left\{ \sum_{j=-n}^n |X_j| \right\} \\
 &= \lim_{n \rightarrow \infty} \sum_{j=-n}^n E\{|X_j|\} \\
 &= \sum_{j=-\infty}^{\infty} E\{|X_j|\} < \infty.
 \end{aligned}$$

Tomando $S_n = \sum_{j=-n}^n X_j$,

$$|S_n| = \left| \sum_{j=-n}^n X_j \right| \leq \sum_{j=-n}^n |X_j|.$$

Como $S_n \rightarrow \sum_{j=-\infty}^{\infty} X_j$, segue que

$$|S_n| \leq \sum_{j=-\infty}^{\infty} |X_j|.$$

Aplicando o Teorema da Convergência Dominada (James, 1981) na sequência $\{S_n\}$, tem-se

$$E \left\{ \sum_{j=-\infty}^{\infty} X_j \right\} = \sum_{j=-\infty}^{\infty} E\{X_j\}.$$

□

Demonstração do Teorema 2.2.4:

No Teorema 2.2.1 apresentado por Fuller (1976), tem-se

$$E \left\{ \sum_{j=-\infty}^{\infty} a_j Z_{t-j} \right\} = \sum_{j=-\infty}^{\infty} a_j E\{Z_{t-j}\},$$

logo,

$$E\{X_t\} = E\left\{\sum_{j=-\infty}^{\infty} a_j Z_{t-j}\right\} = \sum_{j=-\infty}^{\infty} a_j E\{Z_{t-j}\} = \lim_{n \rightarrow \infty} \sum_{j=-n}^n a_j E\{Z_{t-j}\},$$

e a primeira parte está provada.

Para provar a segunda parte, observa-se que

$$\sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} E\{|a_j b_k X_{t-j} X_{t-k}|\} = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |a_j b_k| E\{|X_{t-j} X_{t-k}|\}. \quad (8.1)$$

Mas como

$$0 \leq (|X_{t-j}| - |X_{t-k}|)^2 = |X_{t-j}|^2 - 2|X_{t-j}||X_{t-k}| + |X_{t-k}|^2,$$

então,

$$2|X_{t-j}||X_{t-k}| \leq |X_{t-j}|^2 + |X_{t-k}|^2.$$

Dessa forma,

$$2E\{|X_{t-j}||X_{t-k}|\} \leq E\{|X_{t-j}|^2\} + E\{|X_{t-k}|^2\} \leq 2K,$$

portanto

$$E\{|X_{t-j}||X_{t-k}|\} \leq K. \quad (8.2)$$

De (8.1) e (8.2) obtem-se

$$\begin{aligned} \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} E\{|a_j b_k X_{t-j} X_{t-k}|\} &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |a_j| |b_k| E\{|X_{t-j} X_{t-k}|\} \\ &\leq \left(\sum_{j=-\infty}^{\infty} |a_j|\right) \left(\sum_{k=-\infty}^{\infty} |b_k|\right) K < \infty, \end{aligned}$$

uma vez que $\{a_j\}$ e $\{b_j\}$ são absolutamente somáveis.

Aplicando o Teorema 2.2.3, tem-se

$$\begin{aligned}
E\{Y_t Z_t\} &= E \left\{ \sum_{j=-\infty}^{\infty} a_j X_{t-j} \sum_{k=-\infty}^{\infty} b_k X_{t-k} \right\} \\
&= E \left\{ \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} a_j b_k X_{t-j} X_{t-k} \right\} \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} E \{ a_j b_k X_{t-j} X_{t-k} \} \\
&= \lim_{n \rightarrow \infty} \sum_{j=-n}^n \sum_{k=-n}^n E \{ a_j b_k X_{t-j} X_{t-k} \}.
\end{aligned}$$

□

Demonstração do Corolário 1:

Como $\{a_j\}$ e $\{b_j\}$ são absolutamente somáveis, $\{e_t\}$ uma sequência de variáveis aleatórias e $E\{e_t^2\} = \text{var}\{e_t\} + E\{e_t\} = \sigma^2 = K$, pode-se aplicar o Teorema 2.2.4.

Então,

$$E\{Y_t\} = \lim_{n \rightarrow \infty} \sum_{j=-n}^n a_j E\{e_{t-j}\} = 0.$$

Como $\{e_t\}$ é uma sequência de variáveis aleatórias não-correlacionadas, então

$$E\{e_{t-j} e_{t-k}\} = 0,$$

para $j \neq k$.

Logo,

$$E\{Y_t Z_t\} = \lim_{n \rightarrow \infty} \sum_{j=-n}^n \sum_{k=-n}^n a_j b_k E\{e_{t-j} e_{t-k}\} = \sigma^2 \lim_{n \rightarrow \infty} \sum_{j=-n}^n a_j b_j = \sigma^2 \sum_{j=-\infty}^{\infty} a_j b_j.$$

□

Apêndice C

Demonstração da Proposição 5.2.1

Esta seção apresenta a demonstração de algumas das propriedades do operador *thinning*

$$(i) \ 0 \circ X_1 = \sum_{j=1}^{X_1} Y_{j,1} \text{ tal que } P(Y_{j,1} = 1) = \alpha = 0 \Rightarrow P(Y_{j,1} = 0) = 1 - \alpha = 1$$

Logo

$$0 \circ X_1 = \sum_{j=1}^{X_1} Y_{j,1} = \sum_{j=1}^{X_1} 0 = 0.$$

(v)

$$\begin{aligned} E[\alpha_1 \circ X_1] &= E[E[(\alpha_1 \circ X_1)|X_1]] \\ &= E \left[E \left[\sum_{j=1}^{X_1} Y_{j,1} | X_1 \right] \right] \\ &= E \left[\sum_{j=1}^{X_1} E[Y_{j,1} | X_1] \right] \\ &= E \left[\sum_{j=1}^{X_1} E[Y_{j,1}] \right] \\ &= E \left[\sum_{j=1}^{X_1} \alpha_1 \right] \\ &= E[\alpha_1 X_1] \\ &= \alpha_1 E[X_1] \end{aligned}$$

(vi)

$$\begin{aligned} E[(\alpha_1 \circ X_1)X_2] &= E [E [(\alpha_1 \circ X_1) X_2|X_2]] \\ &= E \left[E \left[\left(\sum_{j=1}^{X_1} Y_{j,1} \right) X_2|X_2 \right] \right] \\ &= E[E[(Y_{1,1}X_2 + Y_{2,1}X_2 + \dots + Y_{X_1,1}X_2)|X_2]] \\ &= E[E[Y_{1,1}X_2|X_2] + E[Y_{2,1}X_2|X_2] + \dots + E[Y_{X_1,1}X_2|X_2]] \\ &= E[E[Y_{1,1}|X_2]E[X_2|X_2] + E[Y_{2,1}|X_2]E[X_2|X_2] + \dots \\ &\quad + E[Y_{X_1,1}|X_2]E[X_2|X_2]] \\ &= E [E[Y_{1,1}]X_2 + E[Y_{2,1}]X_2 + \dots + E[Y_{X_1,1}]X_2] \\ &= E[\alpha_1 X_2 + \dots + \alpha_1 X_2] \\ &= E[\alpha_1 X_1 X_2] \\ &= \alpha_1 E[X_1 X_2] \end{aligned}$$

(viii)

$$\begin{aligned} E[(\alpha_1 \circ X_1)^2] &= E [E [(\alpha_1 \circ X_1)^2 |X_1]] \\ &= E \left[E \left[\left(\sum_{i=1}^{X_1} Y_{i,1}^2 + \sum_{j=1}^{X_1} \sum_{\substack{k=1 \\ j \neq k}}^{X_1} Y_{j,1} Y_{k,1} \right) |X_1 \right] \right] \\ &= E \left[\sum_{i=1}^{X_1} E[Y_{i,1}^2|X_1] + \sum_{j=1}^{X_1} \sum_{\substack{k=1 \\ j \neq k}}^{X_1} E[Y_{j,1} Y_{k,1}|X_1] \right] \\ &= E \left[\sum_{i=1}^{X_1} E[Y_{i,1}^2] + \sum_{j=1}^{X_1} \sum_{\substack{k=1 \\ j \neq k}}^{X_1} E[Y_{j,1} Y_{k,1}] \right] \\ &= E[X_1(\sigma_1^2 + \alpha_1^2) + X_1(X_1 - 1)(\alpha_1 \alpha_1)] \\ &= E[\sigma_1^2 X_1 + \alpha_1^2 X_1^2] \\ &= \sigma_1^2 E[X_1] + \alpha_1^2 E[X_1^2] \end{aligned}$$

(x)

$$\begin{aligned} E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)] &= E[E[(\alpha_1 \circ X_1)(\alpha_2 \circ X_2)|X_1, X_2]] \\ &= E \left[E \left[\left(\sum_{i=1}^{X_1} Y_{i,1} \right) \left(\sum_{j=1}^{X_2} Y_{j,2} \right) | X_1, X_2 \right] \right] \\ &= E \left[\sum_{i=1}^{X_1} \sum_{j=1}^{X_2} E[Y_{i,1} Y_{j,2} | X_1, X_2] \right] \\ &= E \left[\sum_{i=1}^{X_1} \sum_{j=1}^{X_2} E[Y_{i,1} Y_{j,2}] \right] \\ &= E[X_1 X_2 \alpha_1 \alpha_2] \\ &= \alpha_1 \alpha_2 E[X_1 X_2] \end{aligned}$$

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)