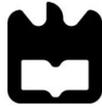




Universidade de Aveiro Departamento de Matemática
2010

Magda Sofia
Valério Monteiro

**Contributos para a Análise Estatística de Séries de
Contagem**



Universidade de Aveiro Departamento de Matemática
2010

**Magda Sofia
Valério Monteiro**

Contributos para a Análise Estatística de Séries de Contagem

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Matemática, realizada sob a orientação científica da Doutora Isabel Maria Simões Pereira, Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro e sob co-orientação científica do Doutor Manuel González Scotto, Professor Auxiliar do Departamento de Matemática da Universidade de Aveiro.

Dedico este trabalho ao Arlindo e ao João Pedro pelo apoio e amor incondicionais.

o júri

Presidente

Reitor da Universidade de Aveiro

Prof. Doutor Kamil Feridum Turkman

Professor Catedrático da Faculdade de Ciências da Universidade de Lisboa

Prof.^a Doutora Maria Esmeralda Elvas Gonçalves

Professora Associada da Faculdade de Ciências da Universidade de Coimbra

Prof.^a Doutora Maria Eduarda da Rocha Pinto Augusta da Silva

Professora Associada da Faculdade de Economia da Universidade do Porto

Prof.^a Doutora Andreia Oliveira Hall

Professora Associada da Universidade de Aveiro

Prof.^a Doutora Isabel Maria Simões Pereira

Professora Auxiliar da Universidade de Aveiro (Orientadora)

Prof. Doutor Manuel González Scotto

Professor Auxiliar da Universidade de Aveiro (Co-orientador)

agradecimentos

Aos meus orientadores, Professora Isabel Pereira e Professor Manuel Scotto, que, com o seu saber feito de experiência e com a sua incondicional amizade, souberam criar todas as condições para que eu pudesse concretizar os meus objectivos. A conclusão deste trabalho muito ficou a dever às suas valiosas sugestões, ao seu constante apoio e à sua profícua orientação científica.

À Universidade de Aveiro, que tem sido a minha casa nestes últimos anos e, em particular, à Escola Superior de Tecnologia e Gestão de Águeda, por tudo o que me proporcionou para a preparação desta dissertação. Ainda a nível institucional, deixo um agradecimento muito especial à Unidade de Investigação Matemática e Aplicações (UIMA) por ter disponibilizado recursos, nomeadamente financeiros, que me permitiram a experiência ímpar de apresentar e divulgar o meu trabalho em alguns congressos nacionais e internacionais.

Aos meus colegas da ESTGA, que, com os seus conselhos oportunos e a sua escuta sensível, me apoiaram durante a realização deste meu trabalho. A sua presença e a sua disponibilidade também contribuíram para o que aqui se apresenta, já que neles encontrei a inspiração e a força que, por vezes, me iam faltando.

À minha família, que, abnegada e pacientemente, tantas e tantas vezes compreendeu a minha ausência e me ajudou, ao longo de todo o processo, a superar sofrimentos e dúvidas.

palavras-chave

Dados de contagem, processos auto-regressivos, *binomial thinning*, sistemas de alarme óptimos, modelos com estrutura periódica, modelos definidos por limiares.

resumo

A análise das séries temporais de valores inteiros tornou-se, nos últimos anos, uma área de investigação importante, não só devido à sua aplicação a dados de contagem provenientes de diversos campos da ciência, mas também pelo facto de ser uma área pouco explorada, em contraste com a análise séries temporais de valores contínuos.

Uma classe que tem obtido especial relevo é a dos modelos baseados no operador *binomial thinning*, da qual se destaca o modelo auto-regressivo de valores inteiros de ordem p . Esta classe é muito vasta, pelo que este trabalho tem como objectivo dar um contributo para a análise estatística de processos de contagem que lhe pertencem. Esta análise é realizada do ponto de vista da predição de acontecimentos, aos quais estão associados mecanismos de alarme, e também da introdução de novos modelos que se baseiam no referido operador.

Em muitos fenómenos descritos por processos estocásticos a implementação de um sistema de alarmes pode ser fundamental para prever a ocorrência de um acontecimento futuro. Neste trabalho abordam-se, nas perspectivas clássica e bayesiana, os sistemas de alarme óptimos para processos de contagem, cujos parâmetros dependem de covariáveis de interesse e que variam no tempo, mais concretamente para o modelo auto-regressivo de valores inteiros não negativos com coeficientes estocásticos, DSINAR(1).

A introdução de novos modelos que pertencem à classe dos modelos baseados no operador *binomial thinning* é feita quando se propõem os modelos PINAR(1) $_{\tau}$ e o modelo SETINAR(2;1).

O modelo PINAR(1) $_{\tau}$ tem estrutura periódica, cujas inovações são uma sucessão periódica de variáveis aleatórias independentes com distribuição de Poisson, o qual foi estudado com detalhe ao nível das suas propriedades probabilísticas, métodos de estimação e previsão.

O modelo SETINAR(2;1) é um processo auto-regressivo de valores inteiros, definido por limiares auto-induzidos e cujas inovações formam uma sucessão de variáveis independentes e identicamente distribuídas com distribuição de Poisson. Para este modelo estudam-se as suas propriedades probabilísticas e métodos para estimar os seus parâmetros.

Para cada modelo introduzido, foram realizados estudos de simulação para comparar os métodos de estimação que foram usados.

keywords

Count data, autoregressive processes, binomial thinning, optimal alarm systems, models with periodic structure, threshold models.

abstract

The analysis of count processes has become an important area of research in the last two decades partially because of its wide applicability in different fields of science. Among the most successful integer-valued time series models proposed in the literature, we mention the binomial thinning based models class, which includes the autoregressive integer valued process of order p as a special case. This work aims to contribute to the statistical analysis of counting processes. In particular, the purpose of this thesis is two-folded: firstly, it explores the issue of event prediction associated with alarm mechanisms and secondly, it introduces two new models based on the binomial thinning operator.

In many phenomena described by stochastic processes, the implementation of an alarm system becomes fundamental to predict the occurrence of future events. In this work we develop an alarm system to predict whether a count process will upcross a certain level and give an alarm whenever the upcrossing level is predicted. We consider count models with parameters being functions of covariates of interest and varying on time. Classical and Bayesian methodologies are applied for producing optimal alarm systems. Both methodologies are illustrated and their performance compared through a simulation study. As an example an empirical application to a set of data concerning the number of sunspot on the surface of the sun is presented.

Within the binomial thinning based models class two new models are proposed and studied. The periodic integer-valued autoregressive model of order one with period T , driven by a periodic sequence of independent Poisson-distributed random variables, is studied in some detail. Basic probabilistic and statistical properties of this model are discussed. Moreover, parameter estimation and prediction are topics also addressed.

A class of self-exciting threshold integer-valued autoregressive models, driven by independent Poisson-distributed random variables, is also introduced. Basic probabilistic and statistical properties of this class of models are discussed. Moreover, parameter estimation is also addressed. Specifically, the methods of estimation under analysis are the least squares-type and likelihood-based ones. Their performances are compared through a simulation study.

Conteúdo

Introdução	1
1 Sistemas de alarme óptimos para processos de contagem	13
1.1 Sistemas de alarme óptimos	15
1.2 Sistemas de alarme óptimos para processos DSINAR(1)	18
1.3 Métodos de estimação	23
1.4 Estudo de simulação	25
1.5 Exemplo	38
1.6 Conclusões	41
2 Processos auto-regressivos de valores inteiros com estrutura periódica	43
2.1 O modelo PINAR(1) _T	46
2.2 Métodos de estimação	56
2.2.1 Estimadores baseados nos momentos	56
2.2.2 Estimadores dos mínimos quadrados condicionais	57
2.2.3 Estimadores dos mínimos quadrados ponderados	68
2.2.4 Estimadores de máxima verosimilhança condicional	69
2.3 Estudo de simulação	73
2.4 Predição no modelo PINAR _T (1)	83
2.4.1 Predição pontual	86
2.4.2 Predição intervalar	86
2.4.3 Estudo de simulação	87

2.5	Exemplo de aplicação	93
2.6	Conclusões	96
3	Modelos auto-regressivos de valores inteiros com limiares auto-induzidos	99
3.1	O modelo SETINAR(2;1)	100
3.2	Métodos de estimação	112
3.2.1	Estimadores dos mínimos quadrados condicionais	112
3.2.2	Estimadores de máxima verosimilhança condicional	114
3.3	Estudo de simulação	118
3.4	Conclusões	125
4	Considerações finais e trabalho futuro	127
	Apêndices	132
A	Funções implementadas no programa R	133
A.1	Funções utilizadas no Capítulo 1	133
A.1.1	Simulação dos modelos	133
A.1.2	Estimação dos parâmetros	134
A.1.3	Região de alarme e características operacionais	139
A.2	Funções utilizadas no Capítulo 2	149
A.2.1	Simulação do modelo	149
A.2.2	Estimação dos parâmetros do modelo	150
A.2.3	Previsão no modelo	163
A.3	Funções usadas no Capítulo 3	171
A.3.1	Simulação do modelo SETINAR(2;1)	171
A.3.2	Métodos de estimação	172
B	Tabelas e Figuras resultantes do estudo de simulação apresentado no	
	Capítulo 2	183

Glossário de Abreviaturas

AR	Auto-regressivo com suporte em \mathbb{R} .
ARMA	Auto-regressivo média móvel com suporte em \mathbb{R} .
DCDP	Desemprego de curta duração no Concelho de Penamacor.
DSINAR	Auto-regressivo com coeficientes estocásticos com suporte em \mathbb{Z}_0^+ .
EAPM	Erro absoluto percentual médio.
EAM	Erro absoluto médio.
FPTN	Função perda tudo ou nada.
EYW	Estimadores de Yule-Walker.
EQM	Erro quadrático médio.
FAC	Função de autocorrelação.
FACP	Função de autocorrelação parcial.
f.g.p.	Função geradora de probabilidades.
i.i.d.	Independente(s) e identicamente distribuída(s).
INAR	Auto-regressivo com suporte em \mathbb{Z}_0^+ .
INARMA	Auto-regressivo média móvel com suporte em \mathbb{Z}_0^+ .
MQC	Mínimos quadrados condicionais.
MQCP	Mínimos quadrados condicionais ponderados.
MVC	Máxima verosimilhança condicional.
PINAR	Auto-regressivo periódico com suporte em \mathbb{Z}_0^+ .
PARMA	Auto-regressivo média móvel periódico com suporte em \mathbb{R} .
SETINAR	Auto-regressivo com limiares auto-induzidos com suporte em \mathbb{Z}_0^+ .
v.a.(’s)	variável(eis) aleatória(s).

Notações

$\hat{}$ sobre uma letra maiúscula (ou grega)	estimador.
$\hat{}$ sobre uma letra minúscula	estimativa.
\circ	operador <i>binomial thinning</i> .
$\stackrel{d}{=}$	igualdade em distribuição.
\xrightarrow{d}	convergência em distribuição.
$\boldsymbol{\theta}$	vector de parâmetros.
$\hat{\boldsymbol{\theta}}_{YW}$	vector de estimadores obtido pelo método de Yule-Walker.
$\hat{\boldsymbol{\theta}}_{MQC}$	vector de estimadores obtido pelo método dos mínimos quadrados condicionais.
$\hat{\boldsymbol{\theta}}_{MVC}$	vector de estimadores obtido pelo método da máxima verosimilhança condicional.
$\hat{\boldsymbol{\theta}}_{MQCP}$	vector de estimadores obtido pelo método dos mínimos quadrados condicionais ponderados.
$\gamma(k)$	Função de autocorrelação.
$E[X_t]$	Esperança de X_t .
$V[X_t]$	Variância de X_t .
$P_{X_t}(s)$	Função geradora de probabilidades de X_t .
$P_{i,j}$	$P(X_t = j X_{t-1} = i)$.
$L(\boldsymbol{\theta})$	Função de verosimilhança.
$l(\boldsymbol{\theta})$	Logaritmo da função de verosimilhança.
L^2	Espaço das variáveis com momentos finitos até à segunda ordem.

Lista de Figuras

1.1	Regiões de alarme para valores de $k \in [0.151, 0.175]$, considerando as estimativas MVC, no instante $t = 200$	35
1.2	Gráfico da probabilidade da região de alarme em função de k , versus B_1 , B_2 e B_3 no instante $t = 200$ e $j = 2$, considerando a perspectiva clássica com as estimativas MVC.	36
1.3	Série temporal do número de grupos das manchas solares e a área total corrigida, em milionésimos de área visível do sol /100, para o período de 1 de Dezembro de 1998 a 30 Abril de 1999.	39
1.4	FAC e FACP da série área total corrigida em milionésimos da área visível do sol/100.	40
2.1	Gráfico da série DCDP, de Janeiro 1997 a Dezembro de 2007.	44
2.2	Gráficos de algumas medidas amostrais da série DCDP, de Janeiro de 1997 a Dezembro de 2007.	45
2.3	Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (4, 1, 3, 2)$	79
2.4	Viés das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (5, 3, 6, 1)$	80
2.5	Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (2, 3, 1, 4)$	81
2.6	Valores de x_{t+h} (preto), \hat{x}_{t+h} (azul), \hat{m}_{t+h} (vermelho) e $\hat{m}o_{t+h}$ (verde) de uma realização de M1.	90
2.7	Preditores intervalares, a 95% confiança, de x_{t+h} (● representa \hat{x}_{t+h}), $h = 1, \dots, 24$	92
2.8	Valores estimados pelo modelo PINAR(1) ₁₂ ajustado.	95

2.9	FAC e FACP amostrais dos resíduos do modelo PINAR(1) ₁₂ ajustado.	96
3.1	Gráficos de algumas realizações do modelo SETINAR(2;1).	102
3.2	Viés das estimativas de θ para A1, A2, B1 e B2, com $\theta = (0.2, 0.1, 3)$, $\theta = (0.2, 0.65, 3)$, $\theta = (0.2, 0.1, 7)$ e $\theta = (0.2, 0.65, 7)$ respectivamente.	122
3.3	Viés das estimativas de θ para A3, A4, B3 e B4, com $\theta = (0.8, 0.1, 3)$, $\theta = (0.8, 0.65, 3)$, $\theta = (0.8, 0.1, 7)$ e $\theta = (0.8, 0.65, 7)$ respectivamente.	123
B.1	Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (2, 4, 1, 3)$	193
B.2	Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (6, 1, 5, 3)$	194
B.3	Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (3, 6, 1, 5)$	195
B.4	Viés das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (3, 2, 4, 1)$	196
B.5	Viés das estimativas de $\alpha = (0.6, 0.4, 0.8, 0.25)$ e $\lambda = (2, 4, 1, 3)$	197
B.6	Viés das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (1, 6, 3, 5)$	198
B.7	Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (4, 1, 3, 2)$	199
B.8	Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (6, 1, 5, 3)$	200
B.9	Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (3, 5, 1, 6)$	201

Lista de Tabelas

1	Propriedades do operador <i>binomial thinning</i>	2
2	Propriedades do modelo INAR(1).	4
1.1	Características operacionais no instante $t = 200$	28
1.2	Resultados para $\lambda = 2$ no instante $t = 200$ (percentagens em parêntesis).	31
1.3	Resultados para $\lambda = 3$ no instante $t = 200$ (percentagens em parêntesis).	32
1.4	Resultados para $\lambda = 4$ no instante $t = 200$ (percentagens em parêntesis).	33
1.5	Resultados para $\lambda = 5$ no instante $t = 200$ (percentagens em parêntesis).	34
1.6	Características operacionais para uma realização do modelo DSINAR(1), com $\lambda = 4$ e $\omega_1 = 0.3$, no instante $t = 200$	36
1.7	Sistema de alarme óptimo <i>online</i> para os instantes $t = 199, \dots, 205$, segundo B_3	37
1.8	Resultados do sistema de alarme com $u = 4$ (percentagens em parêntesis).	41
2.1	Valores de α e λ usados na simulação	75
2.2	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.5, 0.76, 0.63)$ e $\lambda = (4, 1, 3, 2)$	76
2.3	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.6, 0.4, 0.8, 0.25)$ e $\lambda = (5, 3, 6, 1)$	77
2.4	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (2, 3, 1, 4)$	78
2.5	Percentagens de estimativas dos parâmetros (α, λ) que se encontram dentro do espaço dos parâmetros.	82

2.6	Predições pontuais para uma amostra, de dimensão $n = 400$ com $x_{400} = 15$, do modelo M1 para $h = j + 4k$, $j = \{1, 2, 3, 4\}$ $k = \{0, \dots, 9\}$	89
2.7	EQM, EAM, FPTN e EAPM de 24 predições h passos à frente para M1, M2 e M3.	91
2.8	EQM, EAM, FPTN e EAPM de 24 predições um passo à frente para M1, M2 e M3.	91
2.9	Probabilidades de cobertura a partir das predições intervalares (95% confiança), h passos à frente, $h = 1, \dots, 24$	92
2.10	Probabilidades de cobertura a partir das 24 predições intervalares (95% confiança), 1 passo à frente.	93
2.11	Estimativas dos parâmetros do modelo PINAR(1) ₁₂	94
3.1	Parâmetros dos modelos SETINAR(2; 1).	119
3.2	Média e EQM (em parêntesis) das estimativas de θ para os modelos A1, A2, B1 e B2.	120
3.3	Média e EQM (em parêntesis) das estimativas de θ para os modelos A3, A4, B3 e B4.	121
3.4	Percentagens das estimativas de θ no espaço dos parâmetros.	124
B.1	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (2, 4, 1, 3)$	184
B.2	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (6, 1, 5, 3)$	185
B.3	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (3, 6, 1, 5)$	186
B.4	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.6, 0.4, 0.8, 0.25)$ e $\lambda = (3, 2, 4, 1)$	187
B.5	Média e EQM (em parêntesis) das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (2, 4, 1, 3)$	188

B.6	Média e EQM (em parêntesis) das estimativas de $\boldsymbol{\alpha} = (0.60, 0.40, 0.80, 0.25)$ e $\boldsymbol{\lambda} = (1, 6, 3, 5)$	189
B.7	Média e EQM (em parêntesis) das estimativas de $\boldsymbol{\alpha} = (0.4, 0.2, 0.5, 0.1)$ e $\boldsymbol{\lambda} = (4, 1, 3, 2)$	190
B.8	Média e EQM (em parêntesis) das estimativas de $\boldsymbol{\alpha} = (0.40, 0.20, 0.50, 0.1)$ e $\boldsymbol{\lambda} = (6, 1, 5, 3)$	191
B.9	Média e EQM (em parêntesis) das estimativas de $\boldsymbol{\alpha} = (0.40, 0.20, 0.50, 0.1)$ e $\boldsymbol{\lambda} = (3, 5, 1, 6)$	192

Introdução

As séries temporais de valores inteiros surgem em diversos contextos, habitualmente como contagens de acontecimentos, objectos ou indivíduos em intervalos de tempo especificados. Alguns exemplos são o número mensal de acidentes de trabalho numa fábrica, o número de pacientes atendidos em cada hora numa urgência hospitalar, o número de linhas ocupadas numa rede telefónica num intervalo de tempo, o número de quartos ocupados diariamente num hotel, a população anual de uma dada espécie, o número mensal de pessoas que possuem determinada doença, entre muitos outros. Durante muitos anos as séries temporais de valores inteiros, usualmente designadas de séries de contagem, foram tratadas como se o seu suporte fosse o conjunto dos números reais. Nos casos em que as séries representam contagens de valores elevados este procedimento poderá funcionar pela aplicação do teorema limite central, mas em muitas situações as séries temporais de valores inteiros representam contagens de valores reduzidos, o que torna questionável o uso de modelos de valores contínuos.

A área de investigação relativa à análise de processos de contagem desenvolveu-se muito nas últimas duas décadas sobretudo devido à sua aplicação em diversos campos da ciência como a biologia experimental (Zhou e Basawa, 2005), as ciências sociais (McCabe e Martin, 2005), a procura do turismo internacional (Nordström, 1996, Garcia-Ferrer e Queralt, 1997, Brännäs *et al.*, 2002, e Brännäs e Nordström, 2006), o ambiente (Thyregod *et al.*, 1999), os sistemas de filas de espera (Ahn *et al.*, 2000), a economia (Quoreshi, 2006), as telecomunicações (Weiß, 2008a), e os sistemas de alarme óptimos (Monteiro *et al.*, 2008).

Uma boa revisão dos trabalhos realizados nesta área pode ser encontrada em McDonald e Zucchini (1997, §1) e McKenzie (2003). Para desenvolvimentos mais recentes consultar Jung e Tremayne (2006) e Weiß (2008b).

De entre os modelos propostos para séries temporais de valores inteiros destaca-se a classe dos modelos baseados no operador *binomial thinning*. Este operador foi sugerido por Steutel e Van Harn (1979) e a sua aplicação a uma v.a. discreta X com suporte no conjunto $A = \{0, \dots, n\}$, ou \mathbb{N}_0 , com parâmetro *thinning* $\alpha \in [0, 1]$, resulta na v.a. $\alpha \circ X$ definida por $\sum_{i=1}^X U_i$, sendo os U_i 's v.a.'s i.i.d., independentes de X , com distribuição de Bernoulli cuja probabilidade de sucesso é α . A interpretação deste operador é a seguinte: considere-se uma população que tem X elementos e a probabilidade de qualquer um dos elementos possuir uma determinada característica é α . Se os indivíduos dessa população possuem essa característica de forma independente uns dos outros então o número de elementos da população que possui essa característica é dado por $\alpha \circ X$. As principais propriedades do operador *binomial thinning* são apresentadas na Tabela 1 e as suas demonstrações podem ser consultadas, por exemplo, em Silva (2005). Como se pode ver pela referida

Tabela 1: Propriedades do operador *binomial thinning*.

-
1. $\alpha \circ (\beta \circ X) \stackrel{d}{=} (\alpha\beta) \circ X \quad \alpha, \beta \in [0, 1]$
 2. $(\alpha + \beta) \circ X \stackrel{d}{\neq} \alpha \circ X + \beta \circ X \quad \alpha, \beta \in [0, 1]$
 3. $E[\alpha \circ X] = \alpha E[X] \quad \alpha \in [0, 1]$
 4. $E[X(\alpha \circ Y)] = \alpha E[XY] \quad \alpha \in [0, 1]$
 5. $E[(\alpha \circ X)^2] = \alpha^2 E[X^2] + \alpha(1 - \alpha)E(X) \quad \alpha \in [0, 1]$
 6. $V[\alpha \circ X] = \alpha^2 V[X] + \alpha(1 - \alpha)E[X] \quad \alpha \in [0, 1]$
 7. $E[(\alpha \circ X)^3] = \alpha^3 E[X^3] + 3\alpha^2(1 - \alpha)E[X^2] + \alpha(1 - \alpha)(1 - 2\alpha)E[X] \quad \alpha \in [0, 1]$
 8. $P_{\alpha \circ X}(s) = P_X(1 - \alpha + \alpha s)$
-

tabela, o operador *binomial thinning* partilha de algumas propriedades com a multiplicação usual, como é o caso da associatividade entre parâmetros *thinning* (propriedade 1), em

termos de igualdade em distribuição, e também das propriedades relativas a momentos de primeira ordem (propriedades 3 e 4). No entanto, a multiplicação usual goza da propriedade distributiva da soma de escalares relativamente à multiplicação com uma v.a. X (em termos de igualdade em distribuição), propriedade esta que deixa de ser válida quando a multiplicação é substituída pelo operador *binomial thinning* (propriedade 2). Os momentos de ordem superior a um que envolvem o operador *binomial thinning* também não são iguais aos respectivos momentos quando se usa a multiplicação usual em vez do referido operador (propriedades 5, 6 e 7).

Embora o operador *binomial thinning* tenha sido originalmente introduzido por razões teóricas, para adaptar os termos auto-decomponibilidade e estabilidade a séries temporais de valores inteiros, o conceito é agora amplamente utilizado na análise de séries temporais de valores inteiros de modo a adaptar a estrutura dos modelos ARMA de valores contínuos. Tendo em conta que o operador *binomial thinning* retorna sempre um inteiro, a multiplicação usual nos modelos ARMA contínuos foi substituída por este operador que em conjunto com a adopção de uma distribuição discreta para os erros, originou a família de modelos ARMA de valores inteiros (INARMA). Elementos desta classe são os modelos auto-regressivos de valores inteiros de ordem p (INAR(p)), $p \in \mathbb{N}$, e os modelos de médias móveis de valores inteiros de ordem q (INMA(q)), $q \in \mathbb{N}$. Os primeiros foram introduzidos por McKenzie (1986) e Al-Osh e Alzaid (1987) para o caso $p = 1$ e definidos para a ordem p por Alzaid e Al-Osh(1990) e Du e Li(1991). O modelo INAR(p) é definido pela equação recursiva

$$X_t = \alpha_1 \circ X_{t-1} + \alpha_2 \circ X_{t-2} + \cdots + \alpha_p \circ X_{t-p} + Z_t, \quad \alpha_p > 0, \quad t \in \mathbb{Z},$$

onde $(Z_t)_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s i.i.d. de valores inteiros não negativos com $E[Z_t] = \mu_Z$ e $V[Z_t] = \sigma_Z^2$, designada de sucessão das inovações ou dos erros. O modelo INAR(1) foi amplamente estudado por McKenzie (1988), Al-Osh e Alzaid (1987), Alzaid e Al-Osh (1988) e Brännäs e Hellström (2001), nomeadamente quanto à estacionaridade do processo, valor médio, variância e funções de autocovariância e autocorrelação. Para $p = 1$, a interpretação deste modelo é bastante intuitiva: o número de indivíduos que está no sistema

no instante t é a soma do número de indivíduos que permaneceram no sistema, cada um com probabilidade de sobrevivência α_1 , até ao instante t com o número de indivíduos que entram no sistema durante o intervalo de tempo $(t - 1, t]$. Algumas das propriedades do modelo INAR(1) são apresentadas na Tabela 2, e como se pode observar este modelo partilha algumas das propriedades do modelo AR(1) de valores contínuos, como a média condicional e a função de autocorrelação.

Tabela 2: Propriedades do modelo INAR(1).

1. $E[X_t] = \frac{\mu_Z}{1 - \alpha}$
2. $V[X_t] = \frac{\alpha\mu_Z + \sigma_Z^2}{1 - \alpha^2}$
3. $E[X_t X_{t-1}] = \alpha X_{t-1} + \mu_Z$
4. $V[X_t X_{t-1}] = \alpha(1 - \alpha)X_{t-1} + \sigma_Z^2$
5. $\gamma(j) = Cov(X_t, X_{t+j}) = \alpha^j V[X_t]$
6. $\rho(j) = Corr(X_t, X_{t+j}) = \alpha^j$
7. $P_{X_t}(s) = P_{Z_t}(s)P_{X_t}(1 - \alpha + \alpha s)$

Al-Osh e Alzaid (1987) mostraram que o modelo INAR(1) tem distribuição de Poisson se e só se o processo das inovações Z_t tem distribuição de Poisson, o que traduz o papel relevante do modelo INAR(1) cujas inovações têm distribuição de Poisson. Esta distribuição desempenha aqui um papel análogo ao da distribuição normal nos modelos ARMA de valores contínuos. Outras propriedades do modelo INAR(1) com distribuição de Poisson podem ser encontradas em Freeland (1998) e Freeland e McCabe (2004).

No caso do modelo INAR(p) proposto por Alzaid e Al-Osh (1990), os autores supuseram que a soma dos coeficientes dos operadores *binomial thinning* é inferior a um, $\sum_{i=1}^p \alpha_i < 1$, e que cada uma das séries de contagem $\alpha_i \circ X_t$, dado X_t , é independente da história passada do processo X_{t-k} e dos seus sobreviventes $\alpha_j \circ X_{t-k}$, com $i, j = 1, \dots, p$ e $k > 0$, e assumiram que a distribuição condicionada do vector $(\alpha_1 \circ X_t, \alpha_2 \circ X_t, \dots, \alpha_p \circ X_t)$, dado

X_t , é multinomial de parâmetros $(\alpha_1, \dots, \alpha_p, X_t)$. Para este modelo estudaram as suas propriedades e provaram, sob determinadas condições, a existência de uma distribuição limite. Sob esta formulação, o modelo INAR(p) tem a mesma estrutura de segunda ordem de um modelo ARMA($p, p - 1$). Uma abordagem diferente foi seguida por Du e Li (1991). Os autores admitiram que todas as séries de contagem $\alpha_i \circ X_{t-i}$, $\alpha_i \in [0, 1]$; $i = 1, \dots, p$, são mutuamente independentes e independentes de Z_t e mostraram que as condições de estacionaridade (de segunda ordem) do modelo INAR(p) assim definido são as mesmas do modelo AR(p), isto é, existe solução estacionária quando as raízes do polinómio auto-regressivo de ordem p estão fora do círculo unitário. Uma representação alternativa para o modelo INAR(p) foi obtida por Franke e Subba Rao (1995) traduzindo-o como um processo INAR(1) multivariado.

Para o modelo INAR(p), foram sugeridas extensões empíricas relevantes por Brännäs (1995, variáveis explicativas), Brännäs e Hellström (2001, estrutura de dependência), Blundel *et al.* (2002, dados em painel) e mais recentemente por Silva *et al.* (2005, dados replicados). Zheng *et al.* (2006, 2007) estenderam o modelo INAR(1) a coeficientes aleatórios, para o qual investigaram as suas propriedades probabilísticas e ilustraram a sua aplicação na análise da série de contagem de ataques epilépticos (e.g. Latour, 1988) e na análise do número mensal de infectados com poliomielite nos Estados Unidos no período de 1970 a 1983. Alosch (2009) analisou o modelo INAR(1) para dados longitudinais provenientes de ensaios clínicos com lacunas e cujo parâmetros *thinning* são funções de covariáveis. Com este trabalho o autor analisou o impacto de diferentes mecanismos de falta de dados na estimação dos parâmetros do modelo.

O modelo INMA(1) foi introduzido por McKenzie (1986) para distribuições marginais do tipo binomial negativa e geométrica e dois anos depois McKenzie (1988) introduziu o modelo INMA(1) com distribuição marginal de Poisson. O modelo de ordem q foi proposto por Al-Osh e Alzaid (1988) e McKenzie (1988) e satisfaz a equação

$$X_t = Z_t + \alpha_1 \circ Z_{t-1} + \dots + \alpha_q \circ Z_{t-q}, \quad \alpha_q > 0, \quad t \in \mathbb{Z}.$$

No trabalho de Al-Osh e Alzaid (1988) são estudadas algumas das propriedades básicas do modelo, nomeadamente o valor médio, a variância e funções de autocovariância e autocorrelação. Este modelo foi posteriormente estudado por Brännäs e Hall (2001), que trataram a estimação dos parâmetros do modelo, através dos métodos de Yule-Walker e mínimos quadrados condicionais e propuseram ainda um novo método de estimação de momentos generalizado baseado na função geradora de probabilidades. A classe dos modelos INMA foi aplicada por Brännäs *et al.* (2002) no estudo do número de hóspedes num determinado hotel e por Quoreshi (2006) no estudo do número de transacções de acções. Modelos relacionados com o modelo INMA(q) foram introduzidos por Aly e Bouzar (1994, 2005), Zhu e Joe (2003) e mais recentemente por Neal e Subba Rao (2007).

Os modelos mistos auto-regressivos médias móveis de valores inteiros não negativos, INARMA(p, q), foram inicialmente estudados por McKenzie (1986) com distribuições marginais geométricas e posteriormente McKenzie (1998) considerou distribuições marginais de Poisson. Alzaid e Al-Osh (1993) estudaram este tipo de modelos usando margens com distribuição Poisson generalizada. Jung e Tremaine (2006) fazem uma revisão do que já foi analisado para estes modelos.

É de salientar que todos os modelos referidos anteriormente são baseados no operador *binomial thinning* ou em modificações deste operador. Para além deste operador existe uma diversidade de operadores *thinning*. Por exemplo, Gomes (2005) e Gomes e Canto e Castro (2009) propuseram a generalização do operador *binomial thinning*, o operador *thinning* generalizado, \circ^G , que opera sobre duas variáveis aleatórias. A aplicação deste operador às v.a.'s α , com suporte em \mathbb{R}^+ , e X , com suporte em \mathbb{Z}_0^+ , resulta na v.a. $\alpha \circ^G X$, que condicional a X e a α terá a distribuição discreta G com média αX e variância $\sigma^2 = \delta X$ finita, sendo $\delta = \delta(\alpha, X)$ mensurável em \mathbb{R}^+ . Por exemplo, se G for binomial de parâmetros X e α , obtém-se o operador *binomial thinning*. Neste caso $\delta = \alpha(1 - \alpha)$. Com base neste operador generalizado Gomes (2005) apresentou e estudou o modelo DSINAR(1) generalizado e Gomes e Canto e Castro (2009) analisaram o modelo RCINAR(1) generalizado. O

modelo DSINAR(1) satisfaz a equação

$$X_t = \alpha_t \circ^G X_{t-1} + Z_t,$$

onde $(\alpha_t)_{t \in \mathbb{Z}}$ é um processo estocástico com suporte em \mathbb{R}^+ com momentos finitos até à segunda ordem e o processo das inovações $(Z_t)_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s, inteiras não negativas, não correlacionadas que têm uma dada distribuição discreta com média e variância finitas. Para além disso o processo das inovações $(Z_t)_{t \in \mathbb{Z}}$ é independente de $(\alpha_t)_{t \in \mathbb{Z}}$ e da operação *thinning* generalizada e é não correlacionado com X_s , para $s < t$. O modelo RCINAR(1) generalizado é um caso particular do modelo anterior quando se considera que o processo $(\alpha_t)_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s i.i.d.'s. Outros operadores *thinning* que já foram usados na modelação de séries de contagem podem ser encontrados em Weiß (2008b), onde é feito um levantamento destes operadores e dos modelos onde estes são usados.

O estudo bibliográfico efectuado permitiu concluir que dentro da classe dos modelos baseados no operador *binomial thinning*, e suas extensões, ainda não tinham sido abordados temas como os sistemas de alarme óptimos para modelos desta classe, modelos periodicamente correlacionados e modelos definidos através de limiares auto-induzidos. Esta tese aborda estes temas. O Capítulo 1 aborda a predição de acontecimentos para processos de contagem e conseqüentemente a construção de sistema de alarmes óptimos. Como o preditor média condicional não dá necessariamente preditores óptimos, no sentido de dar o menor número de falsos alarmes quanto possível e detectar o maior número de ocorrências de um dado evento, o objectivo é olhar para outros preditores que possuam essas características óptimas. No Capítulo 2 apresenta-se o modelo auto-regressivo de valores inteiros com estrutura periódica, estudam-se as suas propriedades probabilísticas e o problema da inferência e previsão para este modelo, enquanto que no Capítulo 3 é apresentado o modelo auto-regressivo de valores inteiros com limiares auto-induzidos. Para facilitar a leitura e a consistência da tese cada capítulo será independente e auto-contido, cada um deles com uma parte introdutória e portanto não se utilizará um capítulo geral introdutório. De seguida expor-se-á, de forma sucinta, o conteúdo de cada capítulo referindo com especial

ênfase as contribuições inovadores deste trabalho.

Capítulo 1: Sistema de alarme óptimos para processos de contagem

Quando se fala em previsão, no contexto de série temporais, é usual falar-se apenas na predição pontual de um processo num instante futuro, com base nos valores passados do processo. O preditor usualmente utilizado é o que minimiza o erro quadrático médio, ou seja a média do valor pretendido condicionado ao passado. No entanto quando o objectivo não é a predição de um valor futuro do processo mas o de prever se um determinado acontecimento ocorrerá futuramente, a minimização do erro quadrático médio não será a metodologia mais adequada para dar resposta a esse objectivo. Actualmente, um tema importante na análise de uma grande variedade de fenómenos aleatórios consiste na detecção e alerta da ocorrência de uma catástrofe ou algum outro evento relacionado com um mecanismo de alarme. Os exemplos vão desde a previsão do aumento do nível médio do mar e as frequências de inundação devido ao aquecimento global, decorrente do aumento das concentrações de gases com efeito de estufa, até a avaliação do impacto da poluição atmosférica na saúde. Uma forma de o fazer é usando um sistema de alarme simples, designado na literatura por sistema de alarme ingénuo, com base no preditor $\hat{X}_{t+h} = E[X_{t+h}|X_s, -\infty < s \leq t]$, para $h > 0$, onde um alarme é dado sempre que o preditor excede o nível previamente fixado. No entanto este sistema de alarme está longe de ser óptimo porque tem um fraco desempenho na capacidade de detectar os acontecimentos, localizá-los com precisão no tempo e dar o menor número de falsos alarmes quanto possível. Com o objectivo de dar respostas a algumas destas questões Lindgren (1975a, b, 1980, 1985) e de Maré (1980) estabeleceram os princípios para os sistemas de alarme óptimos e obtiveram os resultados básicos para a previsão de cruzamentos de nível; mais tarde, Svensson *et al.* (1996) aplicaram estes resultados na aplicação de sistemas de alarme

óptimos para prever níveis elevados das águas do Báltico. Uma grande desvantagem dos sistemas de alarme introduzidos pelos referidos autores é o facto de não incorporar a variação das estimativas dos parâmetros do modelo. Para superar esta limitação Amaral Turkman e Turkman (1990) sugeriram uma abordagem bayesiana e aplicaram-na a modelos AR(1), embora nenhuma tentativa tenha sido feita para resolver os difíceis problemas computacionais associados. Antunes *et al.* (2003) estenderam os resultados obtidos por Amaral Turkman e Turkman (1990) a modelos auto-regressivos de ordem p e mostraram como obter numericamente as características operacionais do sistema de alarme.

Vale a pena referir que todas as referências citadas no parágrafo anterior tratam o caso de processos estocásticos de valores contínuos. Um problema interessante que está relacionado com o anterior, e que ainda não foi abordado, é o desenvolvimento de sistemas de alarme para séries de contagens que sejam modeladas através de modelos auto-regressivos de valores inteiros. Este capítulo tem como objectivo dar um contributo neste sentido.

Capítulo 2: Processos auto-regressivos de valores inteiros com estrutura periódica

Os processos periodicamente correlacionados, também designados por processos ciclo-estacionários, são processos cujos momentos até à segunda ordem são funções periódicas e cuja estrutura de autocorrelação depende também de uma função periódica. Para além disso, dentro de cada período a estacionaridade de segunda ordem é usualmente preservada para estes processos. Estes processos foram introduzidos por Bennett (1958) e Gladyshev (1961, 1963) e desde então muitos trabalhos foram desenvolvidos sobre este tema, sobretudo devido à sua aplicação nas áreas da hidrologia (Vecchia, 1985), meteorologia (Lund *et al.*, 1995), economia (Franses e Paap, 2001) e processamento de sinal (Gardner *et al.*, 2006), onde existem muitos fenómenos de natureza cíclica. Outros exemplos podem ser encontrados no trabalho de Hurd e Miamee (2007) e nas referências que estes citam. Grande parte

da literatura sobre este tema é dedicada aos modelos auto-regressivos de médias móveis periódicos (PARMA), que são extensões dos modelos ARMA comuns, com parâmetros a variar periodicamente no tempo. A análise das propriedades probabilísticas dos modelos PARMA bem como as técnicas de inferência e predição foram desenvolvidas, entre outros autores, por Bentarzi e Hallin (1994), Lund e Basawa (2000), Basawa e Lund (2001), Basawa *et al.* (2004), Shao e Ni (2004), Lund *et al.* (2006) e Shao (2006).

É de referir que todos os trabalhos mencionados se referem ao caso de processos periodicamente correlacionados de valores contínuos. Por outro lado, a análise de processos de valores inteiros com estrutura periódica não tem recebido grande atenção por parte dos investigadores. Este facto aliado à necessidade de ter em conta a natureza discreta de certos conjuntos de dados, que surgem como contagens de acontecimentos, serviu de motivação para abordar este tema no Capítulo 2, que pretende ser um contributo para o desenvolvimento desta área. Nele será introduzido o modelo $\text{PINAR}(1)_T$ cujas propriedades probabilísticas serão analisadas em detalhe, serão abordados métodos, no domínio da frequência, para estimar os seus parâmetros e será dado destaque à predição para o modelo em análise.

Capítulo 3: Processos auto-regressivos de valores inteiros definidos por limiares auto-induzidos

Na análise de séries estacionárias de valores inteiros a classe INARMA com distribuição de Poisson tem desempenhado um papel fundamental. No entanto esta classe de modelos poderá não ser suficientemente ampla que permita captar de uma forma precisa características exibidas por séries de contagens, tais como súbitos picos de grande magnitude, mudanças de volatilidade no tempo, limiares de grandes excedências que aparecem em grupos e fenómenos do tipo linearidade por troços. Para dar resposta a algumas destas questões Hall (2001) e Hall *et al.* (2009) introduziram uma classe geral de modelos concebidos para

modelar séries temporais de valores inteiros assumindo, com alta probabilidade, valores pequenos mas apresentando ao mesmo tempo súbitos picos de valores positivos de grande magnitude. Outro grande contributo foi o de Doukhan *et al.* (2006) ao introduzirem a classe de modelos bilineares de valores inteiros. Recentemente Drost *et al.* (2008) propuseram extensões ao trabalho desenvolvido por Doukhan e co-autores. Porém, no campo das séries temporais de valores inteiros pouco se tem feito no desenvolvimento de modelos para lidar com séries de contagens que exibam padrões por troços. No levantamento bibliográfico efectuado, surgiu apenas o trabalho de Thyregod *et al.* (1999), no qual foi introduzido um modelo do tipo INAR, baseado em limiares auto-induzidos, para analisar medições da precipitação. No entanto, neste trabalho as questões relacionadas com a existência da distribuição estacionária do processo, a existência de momentos e as propriedades assintóticas dos estimadores dos parâmetros não foram abordadas. Este capítulo tem como objectivo dar um contributo neste sentido.

Os modelos definidos por limiares foram introduzidos por Tong (1977) e desde então muita atenção tem sido dada a esta classe de modelos muito devido à sua aplicação nas áreas da economia e finanças (Boero e Marrocu, 2004, Pai e Pedersen 1999, Potter, 1995), hidrologia (Fu *et al.* 2004) e fenómenos físicos (Tong, 1990). De entre os modelos definidos por limiares o mais divulgado é o modelo ARMA definido por limiares auto-induzidos (SETARMA). O modelo SETARMA de ordem $(k; p_1, \dots, p_k; q_1, \dots, q_k)$ toma a forma

$$X_t = \sum_{i=1}^k \left[\phi_0^{(i)} + \sum_{j=1}^{p_i} \phi_j^{(i)} X_{t-j} + Z_t - \sum_{r=1}^{q_i} \psi_r^{(i)} Z_{t-r} \right] I(X_{t-d} \in R_i), \quad t \in \mathbb{Z},$$

com $(Z_t)_{t \in \mathbb{Z}}$ uma sucessão de v.a.'s i.i.d., $R_i = [r_{i-1}, r_i)$ forma uma partição da recta real tal que $-\infty = r_0 < r_1 < \dots < r_k = +\infty$, sendo r_i o valor do limiar, d representa o atraso do limiar, p_i e q_i são respectivamente as ordens da parte auto-regressiva e média móvel de cada um dos troços, e $\phi_j^{(i)}$ e $\psi_r^{(i)}$ são parâmetros desconhecidos para $j = 1, \dots, p_i$ e $r = 1, \dots, q_i$, $i = 1, \dots, k$. Finalmente $I(\cdot)$ é um processo estocástico de Bernoulli. Note-se que o modelo SETARMA é caracterizado por uma estrutura de linearidade por troços,

com $\bigcup_{i=1}^k R_i = \mathbb{R}$. Refira-se que o modelo é definido por limiares auto-induzidos, isto é o nível do limiar é definido pelo espaço da variável e não pelo tempo.

Todas as referências mencionadas no parágrafo anterior tratam do caso dos modelos de valores contínuos definidos por limiares. A análise de modelos de valores inteiros definidos por limiares auto-induzidos não tem sido alvo de grande interesse na literatura. Este capítulo pretende contribuir para o desenvolvimento desta classe de modelos através da introdução do modelo SETINAR(2;1), do estudo das suas propriedades probabilísticas e da análise dos métodos para estimar os seus parâmetros.

Capítulo 1

Sistemas de alarme óptimos para processos de contagem

Na análise de certos processos estocásticos a implementação de um sistema de alarmes é fundamental para a predição de acontecimentos futuros. Conseguir predizer a ocorrência de um acontecimento que poderá ter consequências graves para algum sector, de forma precisa no tempo, é uma mais valia para esse sector.

Para processos de valores contínuos o desenvolvimento de sistemas de alarme óptimos foi primeiramente abordado por Lindgren (1975a, b, 1980, 1985) e de Maré (1980), que estabeleceram os seus princípios, e mais tarde aplicados por Svensson (1996) a processos auto-regressivos. Antunes *et al.* (2003) usaram uma abordagem bayesiana para implementar um sistema de alarme óptimo a processos auto-regressivos de ordem p , como extensão do trabalho efectuado por Turkman e Turkman (1990) que fizeram essa implementação no caso AR(1). No contexto dos processos de contagem esta temática não tem sido alvo de investigação na literatura. Potenciais aplicações dos sistemas de alarme óptimos para processos de contagem podem ser encontradas no estudo dos efeitos de factores ambientais, tais como poluentes (ozono, nitrogénio, dióxido, etc) e variáveis climatéricas (pressão

atmosférica, temperatura, humidade relativa, etc), na incidência de algumas doenças respiratórias e também na mortalidade. Grande parte dos trabalhos realizados nesta área baseia-se na utilização de modelos lineares generalizados e modelos aditivos generalizados usando técnicas não paramétricas. Por exemplo existem estudos da relação entre a mortalidade e a poluição do ar (Katsouyanni *et al.*, 2002), o número de admissões hospitalares e a poluição do ar (Touloumi *et al.*, 2004), entre a pressão atmosférica e a mortalidade (Campbell *et al.*, 2001, Braga *et al.*, 2001), e doenças gastrointestinais relacionadas com a água potável (Schwartz *et al.*, 1997); para mais referências consultar o trabalho de Koop e Tale (2004). As abordagens utilizadas nas referências supracitadas, não são directamente aplicáveis para prever com antecedência futuros cruzamentos de nível, no sentido de poderem detectar com antecedência um grande número de ocorrências. É neste contexto que os sistemas de alarme podem dar um contributo importante. Questões semelhantes surgem na modelação do número diário ou mensal de hóspedes num hotel (Brännäs *et al.*, 2002 e Brännäs e Nordström, 2006) quando se utilizam covariáveis económicas tais como o nível de receitas do país de origem dos hóspedes, preços e taxas de câmbio. Os modelos sugeridos por Brännäs e seus co-autores não dão resposta à predição de catástrofes, como por exemplo, a procura de quartos exceder a capacidade do hotel.

Neste capítulo pretendem-se salientar os seguintes aspectos:

1. Enquanto que no caso dos modelos de valores contínuos considerados por Svensson *et al.* (1996) as percentagens de alarmes correctos, quando se considera um acontecimento de cruzamento de nível de valor extremo, é aproximadamente 50%, no caso discreto, para os modelos em consideração, essas percentagens são de 30%. É importante referir que os resultados obtidos neste capítulo não são directamente comparáveis com os de Antunes *et al.* (2003), uma vez que estes autores consideram para catástrofes acontecimentos de cruzamentos de nível não extremos;
2. Não existe um critério universal para escolher sistemas de alarme óptimos. Os considerados por Antunes *et al.* (2003) e Svensson *et al.* (1996) não são directamente

aplicáveis ao caso em estudo devido principalmente à natureza discreta dos dados. Neste capítulo é apresentada uma abordagem diferente;

3. A predição de acontecimentos considerada neste trabalho permite a predição *on-line* no sentido em que, as estimativas dos parâmetros do modelo, as regiões de alarme e as características operacionais são actualizadas, em cada instante.

Assim este capítulo está organizado da seguinte forma:

A Secção 1.1 descreve os conceitos fundamentais da predição de acontecimentos. A abordagem aí apresentada é semelhante à de Antunes *et al.* (2003). Na Secção 1.2 é feito o desenvolvimento de um sistema de alarme óptimo para o modelo DSINAR(1) (do inglês *Doubly Stochastic INteger-valued AutoRegressive*). São dadas as expressões das características operacionais do sistema. A estimação dos parâmetros do modelo é discutida na Secção 1.3 sob as perspectivas clássica e bayesiana. Na Secção 1.4 apresenta-se um estudo de simulação onde ambas as perspectivas são comparadas. A Secção 1.5 descreve um exemplo de aplicação dos sistemas de alarme óptimos a um conjunto de dados reais relativos ao número de manchas solares na superfície do sol. Finalmente são apresentadas algumas conclusões na Secção 1.6.

1.1 Sistemas de alarme óptimos

Considere-se um processo estocástico de contagem $(X_t)_{t \in \mathbb{N}}$, com espaço de parâmetro $\Theta \subset \mathbb{R}^k$, para algum $k \in \mathbb{N}$ fixo. Suponha-se a divisão da sequência temporal $\{1, 2, \dots, t-1, t, t+1, \dots\}$ nas três secções $\{1, 2, \dots, t-q\}$, $\{t-q+1, \dots, t\}$ e $\{t+1, \dots\}$ que correspondem ao passado, presente e futuro de modo a que, para algum $q > 0$, os conjuntos $D_t = \{X_1, X_2, \dots, X_{t-q}\}$, $\mathbf{X}_2 = \{X_{t-q+1}, \dots, X_t\}$ e $\mathbf{X}_3 = \{X_{t+1}, \dots\}$ representem, respectivamente, a experiência informativa, a experiência presente e a experiência futura no instante t .

Definição 1.1. Designa-se por catástrofe qualquer acontecimento de interesse, $C_{t,j}$, definido na σ -álgebra gerada por \mathbf{X}_3 .

Ao longo deste trabalho considerar-se-á para catástrofe o cruzamento de um nível u no instante $t + j$, isto é, $C_{t,j} = \{X_{t+j-1} \leq u < X_{t+j}\}$, para algum $j \in \mathbb{N}$.

Definição 1.2. Qualquer acontecimento $A_{t,j}$, definido na σ -álgebra gerada por \mathbf{X}_2 que seja preditor de $C_{t,j}$ designa-se por região de alarme.

Diz-se que é dado um alarme no instante t , para a catástrofe $C_{t,j}$, se o valor observado de \mathbf{X}_2 pertence à região de alarme. Além disso, o alarme é considerado correcto se a ocorrência de $A_{t,j}$ for seguida da ocorrência de $C_{t,j}$. Por outro lado, um falso alarme é definido como a ocorrência de $A_{t,j}$ sem a de $C_{t,j}$. Se um alarme é dado quando uma catástrofe ocorre diz-se que a catástrofe é detectada. Ademais a região de alarme $A_{t,j}$ diz-se que tem dimensão $\alpha_{t,j}$ se $\alpha_{t,j} = P(A_{t,j}|D_t)$.

Definição 1.3. A região de alarme é óptima de dimensão $\alpha_{t,j}$ se

$$P(A_{t,j}|C_{t,j}, D_t) = \sup_{B \in \sigma_{\mathbf{X}_2}} P(B|C_{t,j}, D_t), \quad (1.1)$$

com $P(B|D_t) = \alpha_{t,j}$.

É de notar que esta região de alarme é também o supremo, entre todos os acontecimentos de $\sigma_{\mathbf{X}_2}$, da probabilidade de alarme correcto; para mais detalhes ver o Lema 1 de Antunes *et al.* (2003).

Definição 1.4. Um sistema de alarme óptimo de dimensão $\{\alpha_{t,j}\}$ é uma família de regiões de alarme $\{A_{t,j}\}$ que satisfazem (1.1).

O Lema 1.5 que se segue é uma adaptação do Lema 2 em Antunes *et al.* (2003) para processos de contagem.

Lema 1.5. *Sejam $p(\mathbf{x}_2|D_t)$ e $p(\mathbf{x}_2|C_{t,j}, D_t)$ as probabilidades preditivas de \mathbf{X}_2 e $\mathbf{X}_2|C_{t,j}$, respectivamente. O sistema de alarme $\{A_{t,j}\}$ definido por*

$$A_{t,j} = \{\mathbf{x}_2 \in \mathbb{N}_0^q : \frac{P(C_{t,j}|\mathbf{x}_2, D_t)}{P(C_{t,j}|D_t)} \geq k_{t,j}\},$$

para um valor fixo $k_{t,j} : P(\mathbf{X}_2 \in A_{t,j}|D_t) = \alpha_{t,j}$, é óptimo de dimensão $\alpha_{t,j}$.

O lema anterior garante que a região de alarme assim definida produz a maior probabilidade de detecção. Para reforçar o facto da região de alarme depender da escolha de $k_{t,j}$ é importante salientar que, uma vez que $P(C_{t,j}|D_t)$ não depende de \mathbf{x}_2 , a região de alarme pode ser reescrita na forma

$$A_{t,j} = \{\mathbf{x}_2 \in \mathbb{N}_0^q : P(C_{t,j}|\mathbf{x}_2, D_t) \geq k\}, \quad (1.2)$$

onde $k = k_{t,j}P(C_{t,j}|D_t)$ é escolhido por forma a obter características operacionais razoáveis.

Definição 1.6. (Características operacionais) As seguintes probabilidades designam-se por características operacionais de um sistema de alarme:

1. Dimensão do alarme: $P(A_{t,j}|D_t)$;
2. Probabilidade de alarme correcto: $P(C_{t,j}|A_{t,j}, D_t)$;
3. Probabilidade de detectar a catástrofe: $P(A_{t,j}|C_{t,j}, D_t)$;
4. Probabilidade de falso alarme: $P(\bar{C}_{t,j}|A_{t,j}, D_t)$;
5. Probabilidade de não detectar a catástrofe: $P(\bar{A}_{t,j}|C_{t,j}, D_t)$.

O valor de k a escolher deverá ser aquele que, por um lado, maximize as probabilidades de detecção da catástrofe e de alarme correcto e, por outro lado, minimize as probabilidades de falso alarme e de não detecção da catástrofe.

A maximização simultânea das probabilidades de detecção da catástrofe e de alarme correcto em geral não é possível, sendo portanto necessário assumir um compromisso entre

estas características operacionais. Svensson *et al.* (1996) sugerem que se considere o valor de k que corresponda à igualdade das probabilidades de detecção e de alarme correcto. Neste trabalho será considerada uma abordagem diferente cujos detalhes serão abordados na Secção 1.4.

1.2 Sistemas de alarme óptimos para processos DSINAR(1)

Nesta secção desenvolve-se um sistema de alarme óptimo para o modelo DSINAR(1) definido pela equação recursiva

$$X_t = \alpha_t \circ X_{t-1} + Z_t, \quad t \in \mathbb{N} \quad (1.3)$$

onde o operador *thinning* \circ é definido como

$$\alpha_t \circ X_{t-1} \stackrel{d}{=} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_t),$$

sendo $(U_{i,t}(\alpha_t))$, para $i = 1, 2, \dots$, uma sucessão i.i.d. de variáveis aleatórias de Bernoulli com probabilidade de sucesso $P(U_{i,t}(\alpha_t) = 1) = \alpha_t$. Adicionalmente $(Z_t)_{t \in \mathbb{N}}$ constitui uma sucessão de variáveis aleatórias i.i.d. com distribuição de Poisson de parâmetro λ , que se assume, para cada instante t ser independente de X_{t-1} , α_t e de $\alpha_t \circ X_{t-1}$. É adoptado para α_t a seguinte especificação

$$\alpha_t = \frac{e^{\boldsymbol{\omega}' \mathbf{Y}_{t-s}}}{1 + e^{\boldsymbol{\omega}' \mathbf{Y}_{t-s}}}, \quad s \geq 0,$$

sendo $\mathbf{Y}_t = (Y_{1,t}, \dots, Y_{l,t})$ um vector de covariáveis de interesse e $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_l)$ o correspondente vector de parâmetros desconhecidos. Vale a pena referir que, embora para a maioria dos processos que envolvem séries de contagens o pressuposto de $s > 0$ seja muito utilizado, a motivação para também se considerar $s = 0$ surge dos trabalhos de Brännäs *et al.* (2002) e Brännäs e Nordström (2006), onde esta hipótese é mais adequada em termos

de modelação, para analisar a procura turística internacional bem como outros processos relacionados.

O modelo definido em (1.3) foi inicialmente proposto por Gomes (2005), onde se assume que α_t é um processo estocástico com suporte em \mathbb{R}^+ com esperança finita. No trabalho supracitado podem-se encontrar as propriedades deste modelo, entre as quais condições de existência e estacionariedade do processo, cálculo de momentos e também a estimação de parâmetros do modelo.

Para o modelo DSINAR(1) o cálculo das características operacionais é feito com base na probabilidade de catástrofe condicional ao passado D_t e ao presente \mathbf{X}_2 , i.e. $P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y)$, e na probabilidade de catástrofe condicional apenas ao passado D_t , $P(C_{t,j}|D_t, \mathbf{S}_Y)$ (daqui para frente designadas por P_1 e P_2 respectivamente), onde \mathbf{S}_Y é definida como sendo a σ -álgebra gerada por $(\mathbf{Y}_t)_{t \in \mathbb{N}}$. Estas duas probabilidades são obtidas à custa de probabilidades condicionais do tipo $P(x_{t+h}|x_t, \lambda, \omega, \mathbf{S}_Y)$, com $h \in \mathbb{N}$. De modo a obter $P(x_{t+h}|x_t, \lambda, \omega, \mathbf{S}_Y)$ é necessário o próximo resultado.

Proposição 1.7. *Para o modelo DSINAR(1) definido em (1.3) é válida a seguinte relação*

$$X_{t+h}|\mathbf{S}_Y, \lambda, \omega \stackrel{d}{=} \left(\prod_{i=0}^{h-1} \alpha_{t+h-i} \right) \circ X_t + \left[Z_{t+h} + \sum_{j=1}^{h-1} \left(\prod_{m=0}^{j-1} \alpha_{t+h-m} \right) \circ Z_{t+h-j} \right], \quad (1.4)$$

convencionando-se que $\sum_{i=1}^0 a_i = 0, \forall a_i$.

Demonstração A demonstração é feita por indução e assenta nas propriedades do operador *binomial thinning* (ver Silva e Oliveira, 2004). Para $h = 1$ tem-se $X_{t+1} \stackrel{d}{=} \alpha_{t+1} \circ X_t + Z_{t+1}$.

Considere-se que o resultado é válido para $h = p - 1$, ($p > 1$), então para $h = p$, tem-se

$$\begin{aligned}
X_{t+p} | \mathbf{S}_Y, \lambda, \omega &\stackrel{d}{=} \\
&\stackrel{d}{=} \alpha_{t+p} \circ X_{t+p-1} + Z_{t+p} \\
&\stackrel{d}{=} \alpha_{t+p} \circ \left[\left(\prod_{i=0}^{p-2} \alpha_{t+p-1-i} \right) \circ X_t + Z_{t+p-1} + \sum_{j=1}^{p-2} \left(\prod_{m=0}^{j-1} \alpha_{t+p-1-m} \right) \circ Z_{t+p-1-j} \right] \\
&+ Z_{t+p} \\
&\stackrel{d}{=} \left(\alpha_{t+p} \prod_{i=0}^{p-2} \alpha_{t+p-1-i} \right) \circ X_t + \alpha_{t+p} \circ Z_{t+p-1} + \sum_{j=1}^{p-2} \left(\alpha_{t+p} \prod_{m=0}^{j-1} \alpha_{t+p-1-m} \right) \circ Z_{t+p-1-j} \\
&+ Z_{t+p} \\
&\stackrel{d}{=} \left(\prod_{i=0}^{p-1} \alpha_{t+p-i} \right) \circ X_t + \alpha_{t+p} \circ Z_{t+p-1} + \sum_{j=1}^{p-2} \left(\prod_{m=0}^j \alpha_{t+p-m} \right) \circ Z_{t+p-1-j} + Z_{t+p} \\
&\stackrel{d}{=} \left(\prod_{i=0}^{p-1} \alpha_{t+p-i} \right) \circ X_t + \alpha_{t+p} \circ Z_{t+p-1} + \sum_{j=2}^{p-1} \left(\prod_{m=0}^{j-1} \alpha_{t+p-m} \right) \circ Z_{t+p-j} + Z_{t+p} \\
&\stackrel{d}{=} \left(\prod_{i=0}^{p-1} \alpha_{t+p-1-i} \right) \circ X_t + Z_{t+p} + \sum_{j=1}^{p-1} \left(\prod_{m=0}^{j-1} \alpha_{t+p-m} \right) \circ Z_{t+p-j}.
\end{aligned}$$

■

O segundo membro da expressão (1.4) é a soma de duas componentes independentes, uma que depende apenas de X_t e outra que depende das inovações. Tendo em conta as propriedades do operador *binomial thinning* pode-se verificar que a componente que se encontra entre parentesis (em (1.4)) tem distribuição de Poisson de parâmetro $\lambda^* =$

$\lambda \sum_{j=0}^{h-1} \beta_{t+h,j}$ sendo

$$\beta_{t,s} = \begin{cases} \prod_{m=0}^{s-1} \alpha_{t-m} & s \geq 1 \\ 1 & s = 0 \end{cases}.$$

Por outro lado, a componente $\left(\prod_{i=0}^{h-1} \alpha_{t+h-i} \right) \circ X_t$ condicional a $X_t = x_t$ tem distribuição Binomial de parâmetros x_t e $\beta_{t+h,h}$. Portanto a função de probabilidade de X_{t+h} condicional a X_t e \mathbf{S}_Y , é a convolução entre a distribuição Binomial e a distribuição de Poisson, pelo

que toma a forma

$$P(x_{t+h}|x_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) = e^{-\lambda^*} \sum_{i=0}^{\min(x_t, x_{t+h})} \frac{(\lambda^*)^{x_{t+h}-i}}{(x_{t+h}-i)!} C_i^{x_t} (\beta_{t+h,h})^i (1-\beta_{t+h,h})^{x_t-i}. \quad (1.5)$$

Uma vez encontrada $P(x_{t+h}|x_t, \lambda, \boldsymbol{\omega}, \mathbf{S}_Y)$, com $h \in \mathbb{N}$, podem ser calculadas as probabilidades P_1 e P_2 e posteriormente as características operacionais. A expressão de P_1 pode ser calculada usando a seguinte expressão:

$$\begin{aligned} P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) &= \\ &= P(X_{t+j-1} \leq u < X_{t+j}|\mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \\ &= \sum_{x_{t+j-1}=0}^u P(X_{t+j-1} = x_{t+j-1}, X_{t+j} > u|\mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \\ &= \sum_{x_{t+j-1}=0}^u [P(X_{t+j-1} = x_{t+j-1}|\mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \times P(X_{t+j} > u|x_{t+j-1}, \mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega})] \\ &= \sum_{x_{t+j-1} \leq u} p(x_{t+j-1}|x_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \times \left(1 - \sum_{x_{t+j} \leq u} p(x_{t+j}|x_{t+j-1}, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \right), \end{aligned} \quad (1.6)$$

enquanto que a expressão de P_2 é dada por

$$\begin{aligned} P(C_{t,j}|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) &= \\ &= P(X_{t+j-1} \leq u < X_{t+j}|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \\ &= \sum_{x_{t+j-1}=0}^u P(X_{t+j-1} = x_{t+j-1}, X_{t+j} > u|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \\ &= \sum_{x_{t+j-1}=0}^u [P(X_{t+j-1} = x_{t+j-1}|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \times P(X_{t+j} > u|x_{t+j-1}, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega})] \\ &= \sum_{x_{t+j-1} \leq u} p(x_{t+j-1}|x_{t-q}, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \times \left(1 - \sum_{x_{t+j} \leq u} p(x_{t+j}|x_{t+j-1}, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \right). \end{aligned} \quad (1.7)$$

Uma vez que $(X_t)_{t \in \mathbb{N}}$ é um processo de Markov as probabilidades $p(x_{t+j}|\lambda, \boldsymbol{\omega}, x_{t+j-1}, \mathbf{S}_Y)$ e $p(x_{t+j-1}|\lambda, \boldsymbol{\omega}, x_1, \dots, x_t, \mathbf{S}_Y)$ são calculadas através de (1.5). Após estes cálculos preliminares, as características operacionais podem agora ser facilmente calculáveis.

1. Dimensão do alarme

$$\begin{aligned}
 P(A_{t,j}|D_t, \mathbf{S}_Y, \lambda, \omega) &= \sum_{\mathbf{x}_2 \in A_{t,j}} P(\mathbf{X}_2 = \mathbf{x}_2 | D_t, \mathbf{S}_Y, \lambda, \omega) \\
 &= \sum_{\mathbf{x}_2 \in A_{t,j}} \prod_{i=0}^{q-1} p(x_{t-i} | x_{t-i-1}, \mathbf{S}_Y, \lambda, \omega).
 \end{aligned}$$

2. Probabilidade de alarme correcto

$$\begin{aligned}
 P(C_{t,j}|A_{t,j}, D_t, \mathbf{S}_Y, \lambda, \omega) &= \\
 &= \frac{P(C_{t,j} \cap A_{t,j} | D_t, \mathbf{S}_Y, \lambda, \omega)}{P(A_{t,j} | D_t, \mathbf{S}_Y, \lambda, \omega)} \\
 &= \frac{\sum_{\mathbf{x}_2 \in A_{t,j}} P(\mathbf{X}_2 = \mathbf{x}_2 | D_t, \mathbf{S}_Y, \lambda, \omega) P(C_{t,j} | \mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \omega)}{P(\mathbf{X}_2 = \mathbf{x}_2 | D_t, \mathbf{S}_Y, \lambda, \omega)} \\
 &= \frac{\sum_{\mathbf{x}_2 \in A_{t,j}} \prod_{i=0}^{q-1} p(x_{t-i} | x_{t-i-1}, \mathbf{S}_Y, \lambda, \omega) P(C_{t,j} | \mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \omega)}{\sum_{\mathbf{x}_2 \in A_{t,j}} \prod_{i=0}^{q-1} p(x_{t-i} | x_{t-i-1}, \mathbf{S}_Y, \lambda, \omega)}.
 \end{aligned}$$

A expressão da dimensão do alarme em conjunto com a expressão (1.6) permite reescrever a probabilidade de alarme correcto na seguinte forma

$$\begin{aligned}
 P(C_{t,j}|A_{t,j}, D_t, \mathbf{S}_Y, \lambda, \omega) &= \\
 &= \sum_{\mathbf{x}_2 \in A_{t,j}} \left[\prod_{i=0}^{q-1} p(x_{t-i} | x_{t-i-1}, \mathbf{S}_Y, \lambda, \omega) \sum_{x_{t+j-1}=0}^u p(x_{t+j-1} | x_t, \mathbf{S}_Y, \lambda, \omega) \times \right. \\
 &\times \left. \left(1 - \sum_{x_{t+j}=0}^u p(x_{t+j} | x_{t+j-1}, \mathbf{S}_Y, \lambda, \omega) \right) \right] \left[\sum_{\mathbf{x}_2 \in A_{t,j}} \prod_{i=0}^{q-1} p(x_{t-i} | x_{t-i-1}, \mathbf{S}_Y, \lambda, \omega) \right]^{-1}.
 \end{aligned}$$

3. Probabilidade de detectar a catástrofe

$$\begin{aligned}
P(A_{t,j}|C_{t,j}, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) &= \\
&= \frac{P(C_{t,j} \cap A_{t,j}|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega})}{P(C_{t,j}|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega})} \\
&= \frac{\sum_{\mathbf{x}_2 \in A_{t,j}} P(\mathbf{X}_2 = \mathbf{x}_2|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega})}{P(C_{t,j}|D_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega})} \\
&= \sum_{\mathbf{x}_2 \in A_{t,j}} \left[\prod_{i=0}^{q-1} p(x_{t-i}|x_{t-i-1}, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \sum_{x_{t+j-1}=0}^u p(x_{t+j-1}|x_t, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \times \right. \\
&\times \left. \left(1 - \sum_{x_{t+j}=0}^u p(x_{t+j}|x_{t+j-1}, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \right) \right] \times \\
&\times \left[\sum_{x_{t+j-1} \leq u} p(x_{t+j-1}|x_{t-q}, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \left(1 - \sum_{x_{t+j} \leq u} p(x_{t+j}|x_{t+j-1}, \mathbf{S}_Y, \lambda, \boldsymbol{\omega}) \right) \right]^{-1}.
\end{aligned}$$

1.3 Métodos de estimação

Nesta secção considera-se a estimação do vector dos parâmetros desconhecidos $\boldsymbol{\theta} = (\lambda, \boldsymbol{\omega})$ do ponto de vista clássico e bayesiano. No contexto clássico o método de estimação considerado é o método de Máxima Verosimilhança Condicional (MVC). As estimativas MVC são obtidas maximizando a log-verosimilhança relativamente a $\boldsymbol{\theta}$, recorrendo ao método iterativo de Newton-Raphson. Os valores iniciais para inicializar o algoritmo são as estimativas de mínimos quadrados condicionais. Nesta abordagem, os cálculos associados às características operacionais de um sistema de alarme, definidas na secção anterior, são efectuados através do conhecido método *plug-in*. Este procedimento consiste na substituição dos verdadeiros parâmetros do modelo pelas estimativas MVC para cada uma das expressões. Os valores de Y_{t+j} são estimados usando os preditores lineares j passos à frente dados por $\hat{Y}_{t+j} = E[Y_{t+j}|\mathbf{y}]$, sendo $\mathbf{y} = (y_t, y_{t-1}, \dots)$.

Na perspectiva bayesiana é necessário considerar uma distribuição *a priori* para o vector

dos parâmetros $\boldsymbol{\theta}$. Esta distribuição tem como objectivo quantificar a informação que existe sobre os valores dos parâmetros antes se ter acesso aos dados. Para o modelo DSINAR(1) considera-se que $\lambda \sim \text{Gamma}(a, b)$, $a, b > 0$ e $\omega_i \sim N(\mu_i, \tau_i^{-1})$ com $\mu_i \in \mathbb{R}$ e $\tau_i > 0$, $i = 1, \dots, l$. Assumindo que todos os parâmetros envolvidos são independentes entre si, a distribuição *a priori*, $h(\boldsymbol{\theta})$, é proporcional a

$$h(\boldsymbol{\theta}) \propto \frac{1}{\prod_{i=1}^l \sqrt{2\pi\tau_i^{-1}}} \lambda^{a-1} \exp \left\{ -b\lambda - \frac{1}{2} \sum_{i=1}^l \tau_i (\omega_i - \mu_i)^2 \right\}. \quad (1.8)$$

Note-se que esta distribuição é vaga quando os hiperparâmetros tendem para zero. Além disso, a distribuição de D_t condicional a x_0 é a convolução das distribuições Binomial e Poisson que toma a forma

$$f_{D_t}(d_t|x_0, \boldsymbol{\theta}, \mathbf{S}_Y) = \prod_{n=1}^{t-q} e^{-\lambda} \sum_{i=0}^{\min(x_{n-1}, x_n)} \frac{\lambda^{x_n-i}}{(x_n-i)!} C_i^{x_{n-1}} \alpha_n^i (1-\alpha_n)^{x_{n-1}-i}. \quad (1.9)$$

Conjugando as equações (1.8) e (1.9), a distribuição *a posteriori* é proporcional a

$$\begin{aligned} h(\boldsymbol{\theta}|d_t, \mathbf{S}_Y) &\propto \frac{1}{\prod_{i=1}^l \sqrt{2\pi\tau_i^{-1}}} \lambda^{a-1} \exp \left\{ -b\lambda - \frac{1}{2} \sum_{i=1}^l \tau_i (\omega_i - \mu_i)^2 - (t-q)\lambda \right\} \\ &\times \prod_{n=1}^{t-q} \sum_{i=0}^{\min(x_{n-1}, x_n)} \frac{\lambda^{x_n-i}}{(x_n-i)!} C_i^{x_{n-1}} \alpha_n^i (1-\alpha_n)^{x_{n-1}-i}. \end{aligned} \quad (1.10)$$

A probabilidade P_1 é dada por

$$P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y) = \int_{\Gamma} \int P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y, \boldsymbol{\theta}) h(\boldsymbol{\theta}|D_t, \mathbf{S}_Y) d\boldsymbol{\omega} d\lambda, \quad (1.11)$$

onde $\Gamma = \{(\omega_1, \omega_2, \dots, \omega_l) \in (-\infty, \infty)^l\}$ enquanto que a probabilidade de P_2 toma a forma

$$P(C_{t,j}|D_t, \mathbf{S}_Y) = \int_{\Gamma} \int P(C_{t,j}|D_t, \mathbf{S}_Y, \boldsymbol{\theta}) h(\boldsymbol{\theta}|D_t, \mathbf{S}_Y) d\boldsymbol{\omega} d\lambda, \quad (1.12)$$

com $P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y, \boldsymbol{\theta})$ e $P(C_{t,j}|D_t, \mathbf{S}_Y, \boldsymbol{\theta})$ são calculadas através de (1.6) e (1.7) respectivamente.

É de notar que a complexidade das expressões (1.11) e (1.12) impossibilita o seu cálculo

analítico, mesmo no caso $j = 1$. Atendendo à expressão (1.11), como a probabilidade da catástrofe é por definição $P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y) = E_{\theta|D_t=d_t, \mathbf{S}_Y} [P(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y, \theta)]$, facilmente se obtém a sua aproximação de Monte Carlo através da expressão

$$\hat{P}(C_{t,j}|\mathbf{x}_2, D_t, \mathbf{S}_Y) = \frac{1}{m} \sum_{i=1}^m P(C_{t,j}|\mathbf{x}_2, D_t, \theta_i, \mathbf{S}_Y), \quad (1.13)$$

onde os valores de $\theta_i = (\lambda_i, \omega_i)$, $i = 1, \dots, m$ constituem uma amostra da distribuição *a posteriori* $h(\theta|d_t, \mathbf{S}_Y)$. Tendo em conta que não é possível gerar valores directamente a partir desta distribuição, usa-se a metodologia Gibbs com passo Metropolis, disponível no programa WINBUGS, para gerar amostras de (1.10). É adoptado um procedimento semelhante para estimar a probabilidade P_2 e as características operacionais do sistema de alarme.

1.4 Estudo de simulação

Nesta secção apresenta-se um estudo de simulação para ilustrar o desempenho do sistema de alarme óptimo usando conjuntos de dados gerados a partir do modelo DSINAR(1) definido em (1.3) com uma covariável. Como o modelo em consideração é auto-regressivo de ordem 1, a sequência de tempo foi dividida de modo a que o presente contenha apenas uma observação, isto é, $q = 1$. O estudo de simulação contempla 8 diferentes combinações dos parâmetros (λ, ω_1) , nomeadamente $\lambda = 2, 3, 4, 5$ e $\omega_1 = 0.2, 0.3$. Para a covariável Y_t é adoptado o modelo auto-regressivo de ordem 1

$$Y_t - 3 = 0.6(Y_{t-1} - 3) + \epsilon_t, \quad t = 1, 2, \dots,$$

com $\epsilon_t \sim N(0, 1)$. Simulam-se 200 realizações de dimensão 250 para cada combinação de (λ, ω_1) . Embora seja possível a construção de sistemas de alarme para $j > 2$ utiliza-se um sistema de alarme apenas dois passos à frente para reduzir o esforço computacional. Assim, nesta análise considera-se o sistema de alarme óptimo dois passos à frente analisado no instante $t = 200$, i.e. $\mathbf{X}_2 = \{x_{200}\}$ e define-se como catástrofe o cruzamento do nível u no

instante $t+2$, i.e., $C_{200,2} = \{X_{201} \leq u < X_{202}\}$. A escolha de u é feita em quatro etapas: (a) para cada combinação fixa de (λ, ω_1) e para cada realização i , das 200 realizações obtidas, calcula-se a correspondente probabilidade $P_2^{(i)}$ em (1.7), considerando os verdadeiros valores dos parâmetros, dada por

$$P_2^{(i)} = \sum_{x_{t+j-1}^{(i)} \leq u} p(x_{t+j-1}^{(i)} | x_{t-1}^{(i)}, \mathbf{S}_Y) \left(1 - \sum_{x_{t+j}^{(i)} \leq u} p(x_{t+j}^{(i)} | x_{t+j-1}^{(i)}, \mathbf{S}_Y) \right),$$

com $i = 1, \dots, 200$, para um valor fixo de $u \in [10, 25]$; (b) Para o conjunto de probabilidades $(P_2^{(1)}, \dots, P_2^{(200)})$ calcula-se a correspondente média amostral $(P_2^{(1)} + \dots + P_2^{(200)})/200$; (c) repetem-se os passos (a) e (b) para todos os valores inteiros de 10 a 25; (d) escolhe-se o valor apropriado para u cujo valor da média amostral em (b) esteja mais próximo de 0.1. A escolha deste valor é justificado pelo facto de se ter interesse em acontecimentos considerados raros. O procedimento anterior levou à selecção de $u = 8$ para $(\lambda, \omega_1) = (2, 0.2)$, $u = 9$ para $(\lambda, \omega_1) = (2, 0.3)$, $u = 11$ para $(\lambda, \omega_1) = (3, 0.2)$, $u = 13$ para $(\lambda, \omega_1) = (3, 0.3)$, $u = 12$ para $(\lambda, \omega_1) = (4, 0.2)$, $u = 17$ para $(\lambda, \omega_1) = (4, 0.3)$ e $(\lambda, \omega_1) = (5, 0.2)$, e $u = 20$ para $(\lambda, \omega_1) = (5, 0.3)$.

Para uma combinação fixa de (λ, ω_1) e para cada uma das 200 realizações é construída a região de alarme óptima através da expressão (1.2) para valores de k a variar desde P_2 até $P_2+0.1$. Para cada região óptima são calculadas as respectivas características operacionais. Este procedimento é usado nas abordagens clássica (usando os verdadeiros parâmetros e as estimativas de máxima verosimilhança) e bayesiana. Na perspectiva bayesiana e usando a metodologia Gibbs com o passo de Metropolis, são geradas amostras de dimensão 35000 dos parâmetros a partir da distribuição *a posteriori*, que incluem um período de aquecimento de 15000 observações. Apenas são guardadas observações dos parâmetros de 20 em 20 iterações, para se obter uma amostra com observações aproximadamente não correlacionadas.

Como já foi mencionado anteriormente, a escolha de k desempenha um papel fundamental na obtenção das melhores características operacionais. Para além do critério proposto por

Svesson *et al.* (1996), já mencionado na secção 1.1, outro procedimento para lidar com este problema inclui escolher k de modo a que:

- (a) a dimensão do alarme iguale aproximadamente a probabilidade $P(C_{t,2}|D_t)$;
- (b) esse valor satisfaça a condição $P(A_{t,2}|D_t) \approx 2P(C_{t,2}|D_t)$ (para detalhes ver Antunes *et al.*, 2003).

Quando lidamos com processos de contagem estes procedimentos não podem ser aplicados directamente devido à natureza discreta dos dados. Para dar resposta a este problema consideram-se os seguintes critérios para a selecção de k :

Critério 1:

$$(B1) \quad 0.5 < \frac{P(A_{t,2}|D_t)}{P(C_{t,2}|D_t)} \leq 1.5;$$

$$(B2) \quad 1.5 < \frac{P(A_{t,2}|D_t)}{P(C_{t,2}|D_t)} \leq 2.5;$$

$$(B3) \quad 2.5 < \frac{P(A_{t,2}|D_t)}{P(C_{t,2}|D_t)} \leq 3.5;$$

Para o critério B_j , com $j = 1, 2, 3$, o valor óptimo de $k = k^*$ é

$$k^* = \min_{k: j-1/2 < \frac{P(A_{t,2}|D_t)}{P(C_{t,2}|D_t)} \leq j+1/2} \left| \frac{P(A_{t,2}|D_t)}{P(C_{t,2}|D_t)} - j \right|.$$

Critério 2:

$$(B4) \quad P(A_{t,2}|C_{t,2}, D_t) \approx 0.5.$$

Como para todos os valores de k considerados a probabilidade de alarme correcto é muito reduzida, introduz-se o Critério 2 alternativo que beneficia a probabilidade de detecção e que se traduz na escolha de k que verifique $P(A_{t,2}|C_{t,2}, D_t) \approx 0.5$.

A Tabela 1.1 apresenta as características operacionais do sistema de alarme segundo o critério B_2 , considerando as perspectivas clássica (com as estimativas MVC) e bayesiana, para 6 realizações diferentes de cada uma das combinações de $\lambda = 2, 3, 4, 5$ e $\omega_1 = 0.2, 0.3$.

Tabela 1.1: Características operacionais no instante $t = 200$.

Parâmetros	x_{t-1}	k	estimativas MVC				Perspectiva Bayesiana				
			$P(C_{t,2} D)$	$P(A_{t,2,k} D)$	$P(C_{t,2,k} A_{t,2,k})$	$P(A_{t,2,k} C_{t,2})$	k	$P(C_{t,2} D)$	$P(A_{t,2,k} D)$	$P(C_{t,2} A_{t,2,k})$	$P(A_{t,2,k} C_{t,2})$
$(\lambda, \omega_1) = (2, 0.2)$ $u = 8$	2	0.078	0.051	0.109	0.094	0.200	0.091	0.052	0.106	0.094	0.194
	4	0.110	0.082	0.140	0.116	0.198	0.110	0.083	0.140	0.116	0.198
	7	0.109	0.093	0.207	0.111	0.245	0.108	0.093	0.206	0.111	0.245
	9	0.113	0.095	0.172	0.117	0.212	0.113	0.094	0.171	0.117	0.212
	10	0.110	0.096	0.144	0.112	0.179	0.109	0.089	0.143	0.112	0.178
14	0.089	0.040	0.065	0.102	0.166	0.088	0.040	0.071	0.102	0.169	
$(\lambda, \omega_1) = (2, 0.3)$ $u = 9$	2	0.108	0.071	0.137	0.149	0.286	0.109	0.072	0.139	0.149	0.287
	4	0.122	0.080	0.168	0.128	0.270	0.122	0.080	0.168	0.129	0.270
	8	0.134	0.114	0.211	0.140	0.259	0.135	0.113	0.211	0.148	0.261
	9	0.137	0.107	0.194	0.144	0.260	0.136	0.106	0.193	0.143	0.259
	13	0.109	0.049	0.112	0.131	0.296	0.109	0.049	0.114	0.130	0.296
14	0.100	0.035	0.054	0.141	0.220	0.099	0.035	0.058	0.141	0.222	
$(\lambda, \omega_1) = (3, 0.2)$ $u = 11$	3	0.092	0.048	0.076	0.120	0.189	0.092	0.048	0.077	0.120	0.189
	5	0.079	0.055	0.134	0.089	0.217	0.088	0.056	0.128	0.090	0.206
	10	0.118	0.103	0.168	0.120	0.196	0.119	0.103	0.169	0.120	0.197
	12	0.109	0.090	0.159	0.113	0.201	0.109	0.089	0.158	0.113	0.201
	16	0.108	0.059	0.136	0.118	0.274	0.108	0.057	0.137	0.118	0.273
17	0.095	0.048	0.095	0.111	0.218	0.098	0.047	0.094	0.111	0.211	
$(\lambda, \omega_1) = (3, 0.3)$ $u = 13$	6	0.118	0.079	0.183	0.122	0.282	0.118	0.081	0.183	0.124	0.282
	9	0.124	0.107	0.245	0.129	0.297	0.124	0.107	0.244	0.129	0.298
	12	0.158	0.125	0.253	0.165	0.336	0.158	0.124	0.253	0.165	0.335
	14	0.129	0.103	0.247	0.134	0.322	0.129	0.101	0.246	0.133	0.322
	18	0.116	0.065	0.157	0.128	0.312	0.117	0.063	0.157	0.129	0.311
22	0.070	0.013	0.022	0.110	0.184	0.069	0.014	0.025	0.108	0.185	

(continua na página seguinte)

Tabela 1.1-Continuação

Parâmetros	x_{t-1}	estimativas MVC						Perspectiva Bayesiana					
		k	$P(C_{t,2} D)$	$P(A_{t,2,k} D)$	$P(C_{t,2} A_{t,2,k})$	$P(A_{t,2,k} C_{t,2})$	k	$P(C_{t,2} D)$	$P(A_{t,2,k} D)$	$P(C_{t,2} A_{t,2,k})$	$P(A_{t,2,k} C_{t,2})$		
$(\lambda, \omega_1) = (4, 0.2)$ $u = 12$	4	0.099	0.075	0.179	0.101	0.240	0.098	0.075	0.178	0.101	0.240		
	5	0.089	0.067	0.112	0.092	0.155	0.090	0.068	0.113	0.093	0.155		
	11	0.116	0.102	0.250	0.118	0.291	0.116	0.102	0.250	0.119	0.291		
	12	0.137	0.117	0.261	0.141	0.316	0.138	0.116	0.206	0.141	0.250		
	18	0.115	0.073	0.158	0.123	0.267	0.115	0.072	0.158	0.122	0.265		
19	0.117	0.071	0.113	0.137	0.219	0.116	0.070	0.113	0.137	0.217			
$(\lambda, \omega_1) = (4, 0.3)$ $u = 17$	6	0.102	0.057	0.130	0.120	0.273	0.101	0.057	0.130	0.120	0.272		
	8	0.139	0.109	0.215	0.153	0.303	0.139	0.108	0.214	0.152	0.302		
	16	0.144	0.117	0.245	0.154	0.321	0.144	0.117	0.243	0.155	0.321		
	18	0.138	0.102	0.165	0.145	0.235	0.138	0.101	0.164	0.145	0.234		
	22	0.127	0.059	0.141	0.145	0.345	0.127	0.058	0.140	0.144	0.343		
23	0.094	0.034	0.065	0.129	0.248	0.093	0.033	0.067	0.128	0.248			
$(\lambda, \omega_1) = (5, 0.2)$ $u = 17$	6	0.125	0.089	0.160	0.151	0.271	0.125	0.089	0.160	0.150	0.270		
	8	0.149	0.112	0.179	0.166	0.266	0.150	0.112	0.181	0.167	0.267		
	16	0.169	0.149	0.258	0.171	0.296	0.170	0.149	0.258	0.172	0.297		
	17	0.181	0.160	0.374	0.186	0.433	0.181	0.160	0.373	0.186	0.434		
	20	0.161	0.135	0.239	0.166	0.293	0.162	0.134	0.238	0.166	0.293		
21	0.147	0.124	0.218	0.152	0.268	0.148	0.124	0.217	0.153	0.268			
$(\lambda, \omega_1) = (5, 0.3)$ $u = 20$	10	0.157	0.116	0.236	0.182	0.370	0.157	0.115	0.234	0.181	0.369		
	11	0.129	0.097	0.191	0.152	0.301	0.129	0.097	0.191	0.152	0.301		
	18	0.181	0.158	0.321	0.186	0.379	0.180	0.157	0.321	0.185	0.380		
	19	0.204	0.185	0.353	0.208	0.397	0.203	0.184	0.353	0.207	0.397		
	23	0.171	0.152	0.310	0.176	0.359	0.172	0.152	0.309	0.177	0.359		
25	0.150	0.133	0.318	0.157	0.374	0.150	0.133	0.317	0.157	0.374			

Na Tabela 1.2 apresentam-se os resultados para $\lambda = 2$. Para $\omega_1 = 0.2$ o rácio alarmes/catástrofes está, em geral, acima do intervalo definido pelo respectivo critério, enquanto que para $\omega_1 = 0.3$ este rácio está abaixo do intervalo considerado. O número de falsos alarmes é bastante elevado, sendo a sua percentagem sempre superior a 80% para ambos os valores de ω_1 . Para $\omega_1 = 0.2$ o critério B_3 tem a percentagem mais alta de detecção, independentemente da abordagem adoptada, enquanto que para $\omega_1 = 0.3$ o critério B_4 evidencia o melhor conjunto de características operacionais. Na Tabela 1.3, para $\lambda = 3$ e $\omega_1 = 0.2$ o rácio alarmes/catástrofes cai no respectivo intervalo, para cada um dos critérios utilizados e para todas as abordagens consideradas. No entanto, quando se consideram os verdadeiros parâmetros para $\lambda = 3$ e $\omega_1 = 0.3$, isso já não acontece para os critérios B_3 e B_2 . Em termos de falsos alarmes os resultados são bastante análogos aos obtidos na Tabela 1.2 com percentagens superiores a 80%. Para $\lambda = 3$ o critério B_4 parece ser o que garante o melhor conjunto de características operacionais. Os resultados na Tabela 1.4 mostram que o rácio alarmes/catástrofes está, em geral, abaixo dos intervalos definidos para o respectivo critério. Salienta-se ainda que para $\lambda = 4$ e $\omega_1 = 0.2$ o critério mais equilibrado parece ser B_2 . As percentagens de falso alarme estão abaixo de 70% e as percentagens de detecção variam entre 30% e 44.4%. No que diz respeito a $\lambda = 4$ e $\omega_1 = 0.3$, o critério que parece obter o melhor conjunto de características operacionais é B_3 ; as percentagens de falsos alarmes estão abaixo dos 70% e a percentagem de detecção na perspectiva clássica (considerando os verdadeiros parâmetros) é de 30% enquanto que nos outros dois casos é igual a 44.4%.

Relativamente à Tabela 1.5 pode-se concluir que, para $\lambda = 5$ e $\omega_1 = 0.2$ o rácio alarmes/catástrofes pertence sempre ao intervalo do respectivo critério. Parece que neste caso o melhor critério é B_3 , pois a percentagem de catástrofes detectadas é superior a 45% embora a percentagem de falsos alarmes seja superior a 80% em ambas as abordagens em consideração. Para $\lambda = 5$ e $\omega_1 = 0.3$, as percentagens de falsos alarmes são sempre muito elevadas e o critério que parece ser mais razoável é B_2 com uma percentagem de catástrofes detectadas superior a 30%.

Tabela 1.2: Resultados para $\lambda = 2$ no instante $t = 200$ (percentagens em parêntesis).

Perspectiva	Critério	$\omega_1 = 0.2$				$\omega_1 = 0.3$			
		Alarmes		Catástrofes		Alarmes		Catástrofes	
		Falsos	Total	Detectadas	Total	Falso	Total	Detectadas	Total
Class. Par. verdadeiros	B_1	1	1	0	4	11	0	6	
	B_2	32 (84%)	38	6 (35%)	17	14 (88%)	2 (13%)	16	
	B_3	22 (85%)	26	4 (67%)	6	19 (83%)	4 (33%)	12	
	B_4	81 (90%)	90	9 (50%)	18	57 (84%)	11 (58%)	19	
Est. MVC	B_1	4	4	0	5	10	0	3	
	B_2	28 (88%)	32	4 (31%)	13	19 (90%)	2 (12%)	17	
	B_3	16 (84%)	19	3 (60%)	5	21 (84%)	4 (31%)	13	
	B_4	83 (90%)	92	9 (50%)	18	108 (89%)	13 (68%)	19	
Bayesiana	B_1	4	4	0	6	10	0	3	
	B_2	28 (88%)	32	4 (20%)	20	19 (90%)	2 (11%)	18	
	B_3	18 (82%)	22	4 (67%)	6	21 (84%)	4 (31%)	13	
	B_4	82 (90%)	91	9 (50%)	18	62 (85%)	11 (58%)	19	

Tabela 1.3: Resultados para $\lambda = 3$ no instante $t = 200$ (percentagens em parêntesis).

Perspectiva	Critério	$\omega_1 = 0.2$				$\omega_1 = 0.3$			
		Alarmes		Catástrofes		Alarmes		Catástrofes	
		Falsos	Total	Detectadas	Total	Falso	Total	Detectadas	Total
Class. Par. verdadeiros	B_1	11 (85%)	13	2 (20%)	10	23 (85%)	27	4 (14%)	28
	B_2	23 (92%)	25	2 (15%)	13	15 (94%)	16	1 (33%)	3
	B_3	45 (80%)	56	11 (52%)	21	37 (80%)	46	9 (35%)	26
	B_4	74 (85%)	87	13 (54%)	24	65 (84%)	77	12 (43%)	28
Class. Est. MVC	B_1	15 (88%)	17	2 (17%)	12	18 (82%)	22	4 (20%)	20
	B_2	25 (89%)	28	3 (21%)	14	22 (88%)	25	3 (23%)	13
	B_3	37 (84%)	44	7 (47%)	15	36 (80%)	45	9 (38%)	24
	B_4	78 (85%)	92	14 (58%)	24	66 (85%)	78	12 (43%)	28
Bayesiana	B_1	14 (88%)	16	2 (20%)	10	19 (83%)	23	4 (20%)	20
	B_2	25 (89%)	28	3 (21%)	14	21 (88%)	24	3 (27%)	11
	B_3	40 (85%)	47	7 (50%)	14	36 (81%)	44	8 (35%)	23
	B_4	78 (85%)	92	14 (58%)	24	65 (84%)	77	12 (43%)	28

Tabela 1.4: Resultados para $\lambda = 4$ no instante $t = 200$ (percentagens em parêntesis).

Perspectiva	Critério	$\omega_1 = 0.2$				$\omega_1 = 0.3$			
		Alarmes		Catástrofes		Alarmes		Catástrofes	
		Falsos	Total	Detectadas	Total	Falso	Total	Detectadas	Total
Class. Par. verdadeiros	B_1	0	0	0	27	13 (81%)	16	3 (10%)	30
	B_2	25 (76%)	33	8 (30%)	27	23 (76%)	30	7 (23%)	30
	B_3	0	0	0	27	21 (70%)	30	9 (30%)	30
	B_4	76 (81%)	93	17 (63%)	27	46 (75%)	61	15 (50%)	30
Class. Est. MVC	B_1	3	3	0	6	8 (80%)	10	2 (7%)	28
	B_2	14 (66%)	21	7 (36%)	19	18 (90%)	20	2 (10%)	20
	B_3	21 (80%)	26	5 (33%)	15	23 (65%)	35	12 (44%)	27
	B_4	71 (80%)	88	17 (63%)	27	50 (75%)	66	16 (53%)	30
Bayesiana	B_1	4	4	0	6	9 (81%)	11	2 (9%)	27
	B_2	14 (63%)	22	8 (44%)	18	19 (90%)	21	2 (28%)	7
	B_3	21 (75%)	28	7 (50%)	14	23 (67%)	35	12 (44%)	27
	B_4	72 (81%)	89	17 (63%)	27	49 (75%)	65	16 (53%)	30

Tabela 1.5: Resultados para $\lambda = 5$ no instante $t = 200$ (percentagens em parêntesis).

Perspectiva	Critério	$\omega_1 = 0.2$				$\omega_1 = 0.3$			
		Alarmes		Catástrofes		Alarmes		Catástrofes	
		Falsos	Total	Detectadas	Total	Falso	Total	Detectadas	Total
Class. Par. verdadeiros	$B1$	21 (87%)	24	3 (12%)	25	10 (90%)	11	1 (12%)	8
	$B2$	33 (80%)	41	8 (32%)	25	15 (83%)	18	3 (34%)	8
	$B3$	60 (81%)	74	14 (56%)	25	17	17	0	8
	$B4$	62 (81%)	76	14 (56%)	25	35 (89%)	39	4 (50%)	8
Class. Est. MVC	$B1$	13 (72%)	18	5 (22%)	22	10 (76%)	13	3 (37%)	8
	$B2$	41 (84%)	46	7 (30%)	23	20 (83%)	24	4 (50%)	8
	$B3$	48 (82%)	58	10 (45%)	22	30 (88%)	34	4 (50%)	8
	$B4$	65 (85%)	77	12 (48%)	25	40 (91%)	44	4 (50%)	8
Bayesiana	$B1$	12 (70%)	17	5 (22%)	22	9 (75%)	12	3 (37%)	8
	$B2$	35 (85%)	41	6 (26%)	23	21 (84%)	25	4 (50%)	8
	$B3$	46 (82%)	56	10 (45%)	22	30 (88%)	34	4 (50%)	8
	$B4$	63 (84%)	75	12 (48%)	25	40 (90%)	44	4 (50%)	8

Da análise das Tabelas 1.2 a 1.5 pode-se afirmar que não existe um critério universal para o conjunto de parâmetros analisados.

Para ilustrar como é que a predição *online* é efectuada na prática, considera-se uma realização do modelo DSINAR(1) com os parâmetros $\lambda = 4$ e $\omega_1 = 0.3$. A catástrofe que se pretende prever é $C_{t,2} = \{X_{t+1} \leq 17 < X_{t+2}\}$, $t = 199, \dots, 205$. Antes de mais analisa-se a região de alarme no instante $t = 200$. Assim, são calculadas para este exemplo as regiões de alarme considerando as estimativas de máxima verosimilhança e as estimativas bayesianas, com valores de k a variar entre 0.151 e 0.175. O gráfico que contém as regiões de alarme usando as estimativas MVC está representado na Figura 1.1. As regiões de alarme que se obtêm quando se utiliza a perspectiva bayesiana são idênticas.

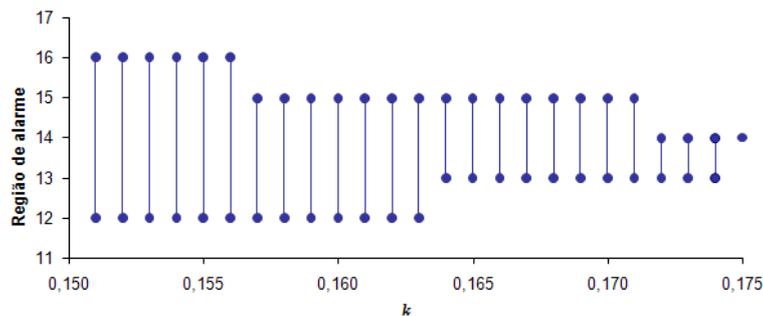


Figura 1.1: Regiões de alarme para valores de $k \in [0.151, 0.175]$, considerando as estimativas MVC, no instante $t = 200$.

A Tabela 1.6 contém as correspondentes características operacionais considerando a perspectiva clássica usando as estimativas de máxima verosimilhança e a perspectiva bayesiana.

Na Figura 1.2 está representada a probabilidade da região de alarme, considerando as estimativas MVC. Os critérios B_1 , B_2 e B_3 também se encontram representados na mesma figura. Para o critério B_1 , a região de alarme óptima e as suas características operacionais são obtidas com $k = 0.175$, enquanto que para o critério B_2 o valor de k é 0.172 e para o critério B_3 k é 0.164.

Tabela 1.6: Características operacionais para uma realização do modelo DSINAR(1), com $\lambda = 4$ e $\omega_1 = 0.3$, no instante $t = 200$.

Perspectiva	k	$P(C_{t,2} D)$	$P(A_{t,2,k} D)$	$P(C_{t,2} A_{t,2,k})$	$P(A_{t,2,k} C_{t,2})$
Est. MVC	0.151	0.151	0.633	0.170	0.712
	0.157	0.151	0.537	0.172	0.612
	0.164	0.151	0.410	0.175	0.476
	0.172	0.151	0.286	0.176	0.334
	0.175	0.151	0.142	0.178	0.168
Bayesiana	0.151	0.150	0.633	0.169	0.714
	0.157	0.150	0.538	0.172	0.615
	0.164	0.150	0.411	0.175	0.478
	0.172	0.150	0.287	0.176	0.336
	0.175	0.150	0.143	0.178	0.169

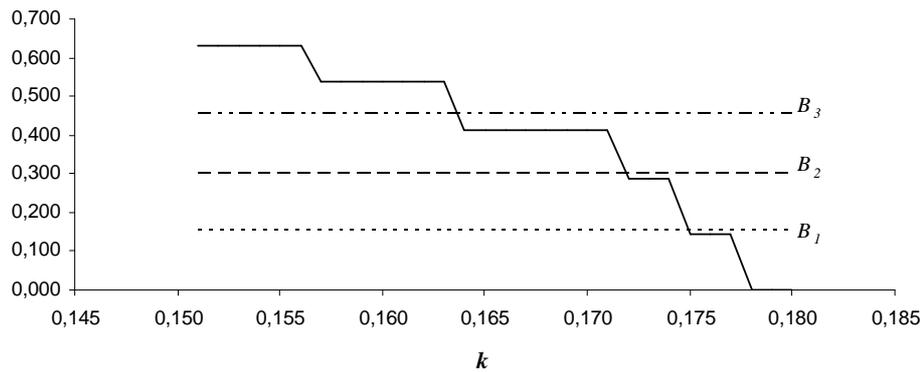


Figura 1.2: Gráfico da probabilidade da região de alarme em função de k , versus B_1 , B_2 e B_3 no instante $t = 200$ e $j = 2$, considerando a perspectiva clássica com as estimativas MVC.

Na Tabela 1.7 apresentam-se as regiões de alarme do sistema de alarme óptimo para os instantes $t = 199, 200, \dots, 205$ e as correspondentes características operacionais considerando o critério B_3 . Como se pode ver, para cada instante t a região de alarme vai sendo adaptada consoante o valor de X_{t-1} , assim como as respectivas características operacionais.

Tabela 1.7: Sistema de alarme óptimo *online* para os instantes $t = 199, \dots, 205$, segundo B_3 .

Perspectiva	t	x_{t-1}	k	$P(C_{t,2} D)$	$P(A_{t,2,k} D)$	$P(C_{t,2} A_{t,2,k})$	$P(A_{t,2,k} C_{t,2})$	Região de Alarme
Class. Par. verdadeiros	199	13	0.192	0.182	0.487	0.207	0.554	{14,15,16}
	200	12	0.138	0.120	0.347	0.149	0.432	{14, 15, 16}
	201	15	0.134	0.119	0.310	0.140	0.365	{15, 16}
	202	18	0.163	0.149	0.438	0.185	0.545	{14, 15, 16, 17}
	203	14	0.119	0.101	0.305	0.122	0.368	{14, 15}
	204	15	0.103	0.092	0.297	0.107	0.344	{16, 17}
	205	13	0.127	0.106	0.297	0.145	0.403	{15, 16, 17, 18}
Class. Est. MVC	199	13	0.198	0.197	0.585	0.224	0.666	{12, 13, 14, 15, 16}
	200	12	0.164	0.151	0.410	0.175	0.476	{13, 14, 15}
	201	15	0.151	0.140	0.449	0.165	0.530	{13, 14, 15, 16}
	202	18	0.171	0.159	0.459	0.203	0.587	{12, 13, 14, 15, 16, 17}
	203	14	0.140	0.128	0.395	0.150	0.465	{13, 14, 15}
	204	15	0.131	0.119	0.385	0.136	0.441	{14, 15, 16}
	205	13	0.156	0.148	0.442	0.171	0.51	{14, 15, 16, 17}
Bayesiana	199	13	0.197	0.195	0.587	0.223	0.669	{12, 13, 14, 15, 16}
	200	12	0.164	0.150	0.411	0.175	0.478	{13, 14, 15}
	201	15	0.151	0.139	0.448	0.165	0.531	{13, 14, 15, 16}
	202	18	0.170	0.157	0.459	0.202	0.587	{12, 13, 14, 15, 16, 17}
	203	14	0.139	0.127	0.396	0.150	0.466	{13, 14, 15}
	204	15	0.131	0.119	0.385	0.137	0.442	{14, 15, 16}
	205	13	0.157	0.148	0.442	0.171	0.512	{14, 15, 16, 17}

1.5 Exemplo

Nesta secção é feita uma aplicação dos sistemas de alarme óptimos a um conjunto de dados reais constituído pelo número de grupos de manchas solares (classificação de Zurich), disponível no boletim anual de observações solares publicado pelo Observatório astrofísico de Catania. Estes relatórios acompanham o nascimento, crescimento e decaimento das manchas solares que surgem na superfície do sol.

As manchas solares são manchas escuras que surgem na superfície do sol; são regiões magnéticas e que podem assumir diversas formas. O tamanho das manchas varia bastante, sendo geralmente maior que o nosso planeta. Elas são medidas em termos de milionésimos da área visível do Sol. Uma mancha é considerada grande quando mede entre 300 e 500 milionésimos do disco solar. As temperaturas no centro das manchas baixam para 3700K (kelvin) (em comparação com os 5700K nas regiões ausentes de manchas). As manchas podem durar vários dias, e as de maior dimensão podem durar várias semanas. Os campos magnéticos criados por estas regiões são milhares de vezes mais fortes do que o campo magnético criado pela Terra. As manchas solares surgem normalmente em grupos de dois conjuntos de manchas. Um conjunto formará um campo magnético positivo e o outro formará um campo magnético negativo. O campo magnético é mais forte na parte mais escura das manchas, designada por sombra e mais fraco na parte mais clara designada de penumbra.

O conjunto de dados considerado é a série diária de grupos de manchas solares observada entre o período de 1 de Dezembro de 1998 e 30 de Abril de 1999. A covariável considerada neste caso é a área total corrigida, em milionésimos da área visível do sol. O primeiro problema a ultrapassar é a existência de lacunas nos dados obtidos. O método usado para preencher os dados omissos em séries temporais depende basicamente do número consecutivo de falhas. Nas séries temporais em análise, as pequenas lacunas foram preenchidas usando o método de interpolação com as observações vizinhas. É de notar que este método produz valores que à partida não são inteiros como a série em estudo. Para obter valores

coerentes com os dados, os valores interpolados foram arredondados para o inteiro mais próximo. Ambas as séries são apresentadas na Figura 1.3.

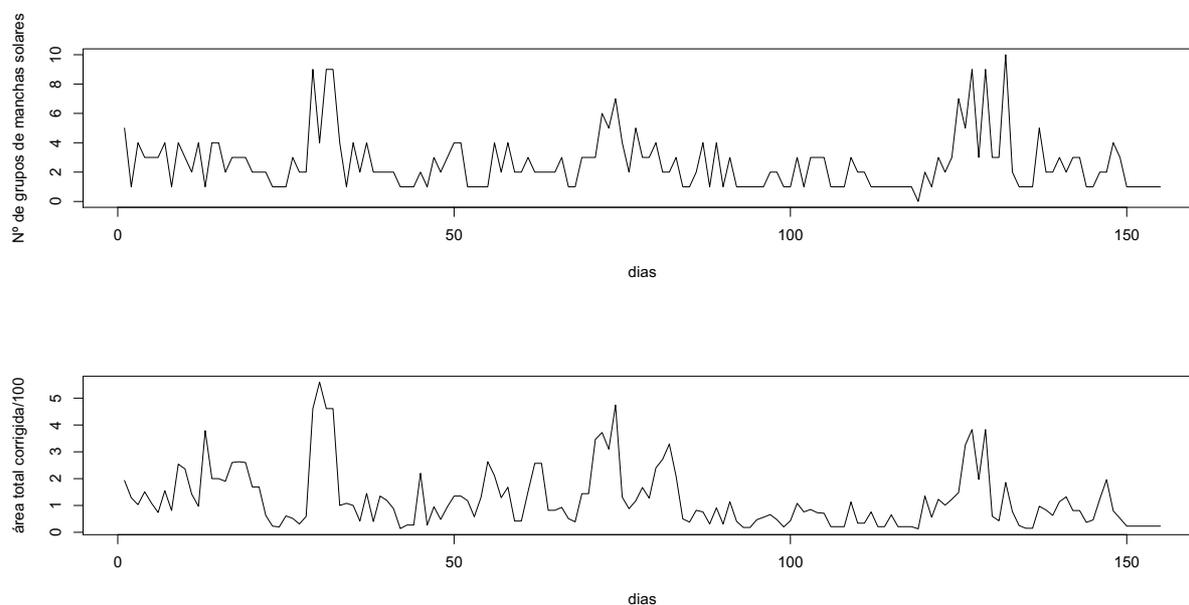


Figura 1.3: Série temporal do número de grupos das manchas solares e a área total corrigida, em milionésimos de área visível do sol /100, para o período de 1 de Dezembro de 1998 a 30 Abril de 1999.

As funções de autocorrelação e de autocorrelação parcial (FAC e FACP) da covariável, área total corrigida (dividida por 100), estão representadas na Figura 1.4. As linhas a tracejado são as aproximações dos limites dos intervalos de confiança a 95%.

Os gráficos da Figura 1.4 são consistentes com os de um processo auto-regressivo de ordem 1, e portanto o passo seguinte é a estimação dos parâmetros e dos acontecimentos de interesse. Para fazer a ilustração do desempenho do sistema de alarme óptimo online, começou-se por usar os dados de 1 de Dezembro de 1998 a 12 de Fevereiro de 1999 para estimar os parâmetros do modelo DSINAR(1). Seguidamente fez-se a aplicação considerando os dados referentes aos dias entre 13 de Fevereiro e 30 de Abril de 1999. Definiu-se para catástrofe o cruzamento de nível extremal $C_{t,2} = \{X_{t+1} \leq 4 \leq X_{t+2}\}$ com $t = 75, 76, \dots, 149$,

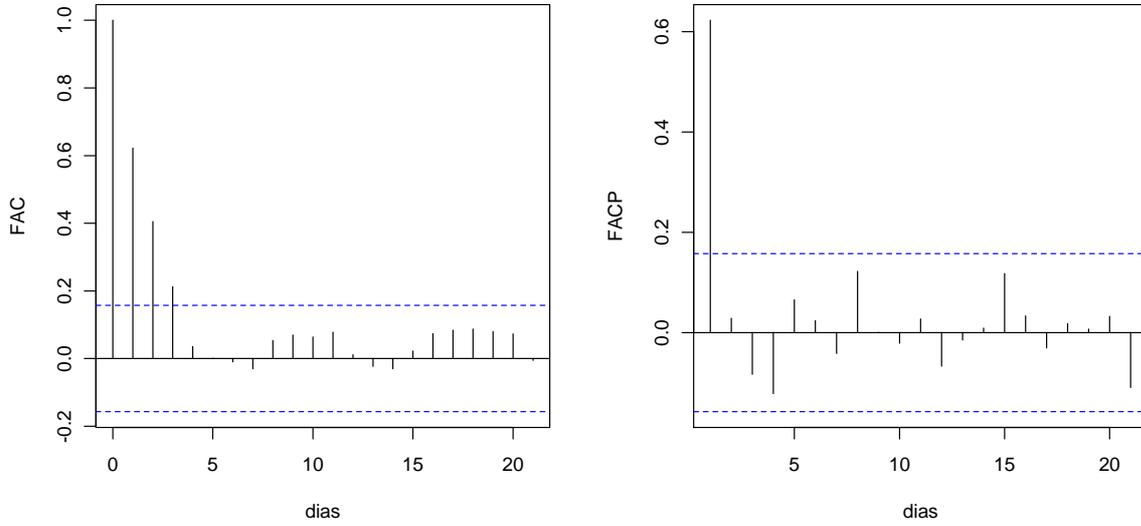


Figura 1.4: FAC e FACP da série área total corrigida em milionésimos da área visível do sol/100.

onde em cada instante considerado a probabilidade da catástrofe condicional ao passado D_t é $P(C_{t,j}|D_t, \mathbf{S}_Y) \approx 0.09$. Para calcular a região óptima e as suas características operacionais, em cada instante t , os valores de Y_{t+j} , $j = 0, 1, 2$, são substituídos pelos respectivos preditores em ambas as perspectivas. Assim, do ponto de vista clássico utilizam-se os preditores $\hat{Y}_{t+j} = \hat{\mu}_{t-1} + \hat{\phi}_{t-1}^j(Y_{t-1} - \hat{\mu}_{t-1})$, onde $(\hat{\mu}_{t-1}, \hat{\phi}_{t-1})$ são as estimativas de máxima verosimilhança dos parâmetros $\mu = E(Y_t)$ e ϕ do modelo auto-regressivo de ordem 1, enquanto que no caso bayesiano os preditores são obtidos através da média da correspondente distribuição preditiva. Na Tabela 1.8 são apresentados os resultados para o sistema de alarme óptimo *online* do modelo DSINAR(1) considerando o critério B_4 .

Analisando a Tabela 1.8 verifica-se que em Fevereiro o sistema de alarme prevê correctamente a catástrofe enquanto que em Abril a percentagem de catástrofes detectadas é de 33%. Por outro lado em Março não existem catástrofes, portanto existe 100% de falsos alarmes. Em todos os meses analisados a percentagem de falsos alarmes é elevada. Este número poderá ser inferior se forem aceites alarmes que foram dados um passo mais cedo e se não forem contabilizados os alarmes que são dados enquanto se está no estado de

catástrofe. Portanto pode ser discutível se estes falsos alarmes deverão ser, num contexto prático, considerados falsos.

Tabela 1.8: Resultados do sistema de alarme com $u = 4$ (percentagens em parêntesis).

Mês	Perspectiva	Alarmes		Catástrofes	
		Falso	Total	Detectada	Total
Fevereiro	Est. MVC	8 (88%)	9	1 (100%)	1
	Bayesiana	8 (88%)	9	1 (100%)	1
Março	Est. MVC	8 (100%)	8	0	0
	Bayesiana	8 (100%)	8	0	0
Abril	Est. MVC	13 (92%)	14	1 (33%)	3
	Bayesiana	13 (92%)	14	1 (33%)	3

Vale a pena referir que, no que diz respeito ao número de falsos alarmes, foram encontrados noutros estudos resultados semelhantes. Por exemplo Svensson e Holst (1998) obtiveram uma taxa de cerca de 95% de falsos alarmes na análise dos níveis do mar Báltico junto à costa dinamarquesa (para mais detalhes consultar Svensson e Holst, 1998).

1.6 Conclusões

Neste capítulo foi apresentado um sistema de alarme óptimo para modelos descritos por processos auto-regressivos de valores inteiros cujos parâmetros são funções de covariáveis de interesse. A técnica dos alarmes óptimos conduz à predição óptima de acontecimentos na medida em que dá o menor número de falsos alarmes para uma pré-determinada dimensão.

Em ambas as perspectivas usadas, clássica e bayesiana, os resultados que se obtiveram foram muito semelhantes, em termos do estudo de simulação bem como no exemplo utilizado. Isto deve-se ao facto de, na abordagem bayesiana, se ter assumido que não existia

informação prévia ao acesso aos dados, tendo sido por isso usadas distribuições *a priori* não informativas.

Tal como foi referido ao longo do capítulo o número de falsos alarmes é elevado quer no estudo de simulação quer no exemplo real. Uma possibilidade de diminuir o número de falsos alarmes pode consistir na inclusão de informação externa adicional no modelo ou supor que o nível da catástrofe possa depender do tempo. Este tópico será analisado em trabalho futuro.

Capítulo 2

Processos auto-regressivos de valores inteiros com estrutura periódica

A modelação de processos periódicos de valores contínuos tornou-se objecto de estudo a partir dos trabalhos de Bennett (1958) e Gladyshev (1961, 1963). De entre os modelos periódicos de valores contínuos os modelos auto-regressivos médias móveis periódicos (PARMA) são os mais estudados ao nível das suas propriedades probabilísticas, da estimação dos parâmetros do modelo e da previsão. O desenvolvimento desta classe de modelos deve-se por um lado, ao facto de serem extensões dos modelos ARMA e por outro, à sua grande utilidade para modelar fenómenos nas áreas da hidrologia, economia, climatologia e de telecomunicações, entre outras. A modelação de processos auto-regressivos de valores inteiros não negativos com estrutura periódica é uma área que se encontra por explorar, o que motivou o desenvolvimento do trabalho que se apresenta neste capítulo. Potenciais aplicações de processos INAR com estrutura periódica surgem por exemplo nas áreas da demografia, epidemiologia, turismo e mercado de trabalho. A título ilustrativo na Figura 2.1 apresenta-se o gráfico da série do número mensal de desempregados de curta duração do Concelho de Penamacor (DCDP), no período de Janeiro de 1997 a Dezembro de 2007. Na Figura 2.2 são representadas as médias e as variâncias mensais e a função de

autocorrelação, onde é bem patente a existência de periodicidade, de período 12, nestas funções.

Neste capítulo introduzir-se-á a classe de modelos auto-regressivos de valores inteiros com estrutura periódica de primeira ordem e período T , que se irá designar por $\text{PINAR}(1)_T$ (do inglês *Periodic INteger-valued AutoRegressive*). Na Secção 2.1 apresenta-se o modelo, demonstra-se a existência e unicidade de um processo estacionário que satisfaz a equação que o define. Também se apresentam expressões para a média, variância e função de autocorrelação da distribuição ciclo-estacionária. A estimação dos parâmetros do modelo é abordada na Secção 2.2 e os resultados obtidos são ilustrados através de um estudo de simulação na Secção 2.3. A predição de novas observações é abordada na Secção 2.4 e a Secção 2.5 descreve, a título ilustrativo, a aplicação do modelo $\text{PINAR}(1)_T$ aos dados da série DCDP que estão representados na Figura 2.1. Finalmente o capítulo termina com algumas considerações gerais.

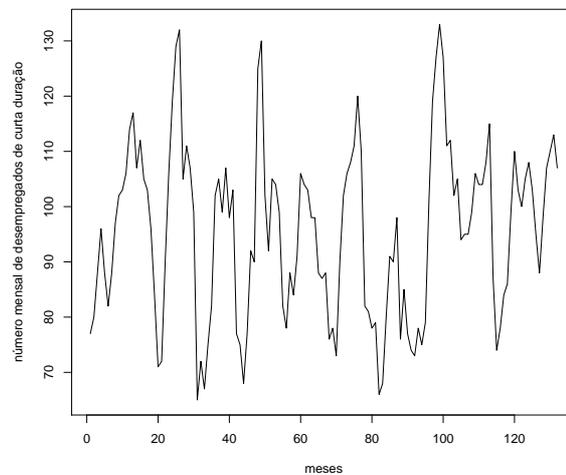
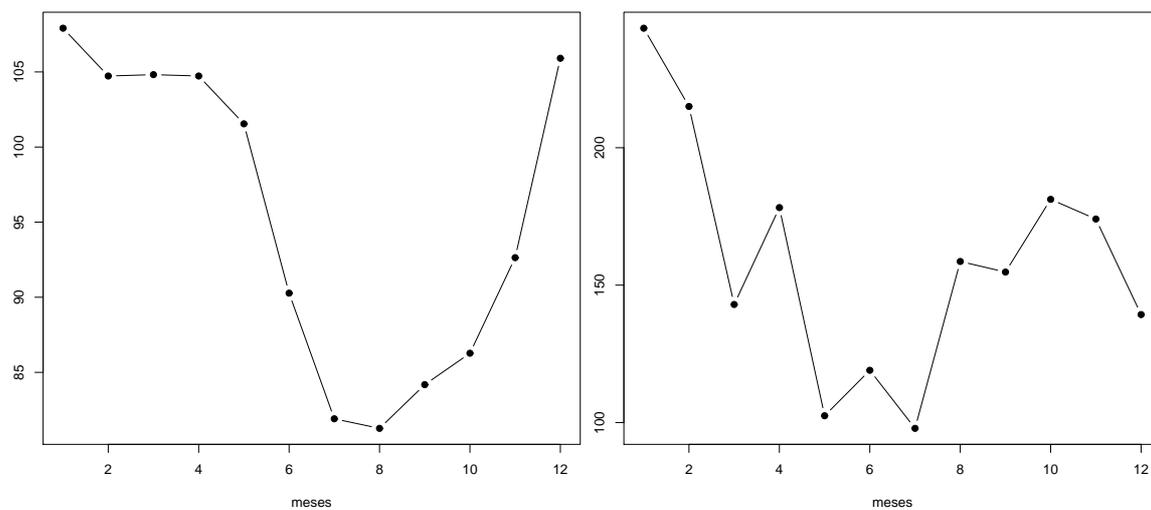
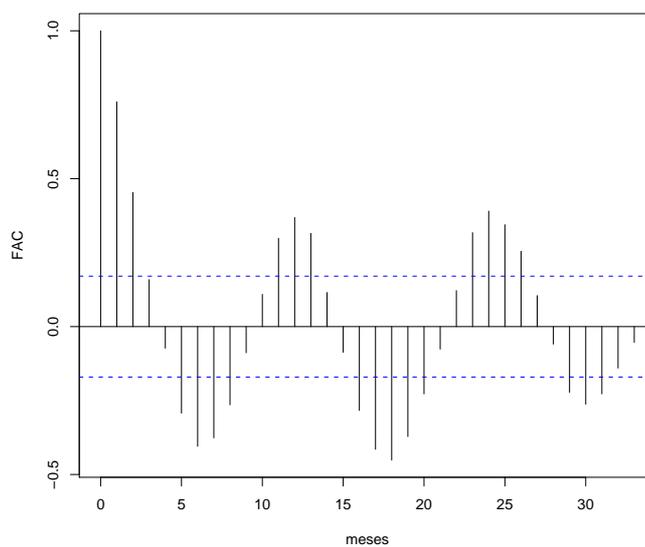


Figura 2.1: Gráfico da série DCDP, de Janeiro 1997 a Dezembro de 2007.



(a) Médias mensais

(b) Variâncias mensais



(c) Correlograma

Figura 2.2: Gráficos de algumas medidas amostrais da série DCDP, de Janeiro de 1997 a Dezembro de 2007.

2.1 O modelo PINAR(1)_T

O modelo PINAR(1)_T é definido pela equação recursiva

$$X_t = \phi_t \circ X_{t-1} + Z_t, t \in \mathbb{N}, \quad (2.1)$$

com $\phi_t = \alpha_j \in (0, 1)$ para $t = j + kT, (j = 1, \dots, T, k \in \mathbb{N}_0)$. A operação *thinning* \circ é definida por

$$\phi_t \circ X_{t-1} \stackrel{d}{=} \sum_{i=1}^{X_{t-1}} U_{i,t}(\phi_t),$$

com $(U_{i,t}(\phi_t)), i = 1, 2, \dots$, uma sucessão periódica de variáveis aleatórias de Bernoulli com probabilidade de sucesso $P(U_{i,t}(\phi_t) = 1) = \phi_t$. Assume-se que o processo das inovações $(Z_t)_{t \in \mathbb{N}}$ é uma sucessão periódica de v.a.'s independentes com distribuição de Poisson de média v_t ($Z_t \sim P(v_t)$), com $v_t = \lambda_j$ para $t = j + kT (j = 1, \dots, T, k \in \mathbb{N}_0)$, onde para cada t , Z_t é independente de X_{t-1} e de $\phi_t \circ X_{t-1}$. Para evitar ambiguidade, T é o menor inteiro positivo que satisfaz (2.1). É importante realçar que a natureza discreta de (X_t) fica garantida pela operação *thinning* que desempenha o papel da multiplicação usual nos modelos ARMA contínuos.

A interpretação deste modelo é bastante intuitiva: para cada período $j, j = 1, \dots, T$, o número de indivíduos que se encontra no sistema é a soma do número de indivíduos que sobreviveram, do período $j - 1$ para o período j (cada um com probabilidade de sobrevivência α_j), com o número de indivíduos que chega ao sistema entre os períodos $j - 1$ e j que é em média λ_j .

Por uma questão de simplicidade define-se a sucessão $\beta_{t,i}$ que irá ter grande utilidade ao longo do capítulo.

Definição 2.1. Para $i \leq t$,

$$\beta_{t,i} = \begin{cases} \prod_{j=0}^{i-1} \phi_{t-j} & i > 0 \\ 1 & i = 0 \end{cases}.$$

Note-se que esta sucessão é periódica de período T , $\beta_{t+kT,i} = \beta_{t,i}$, $t > i$, verificando, para $j = 1, 2, \dots, T$, $k \in \mathbb{N}_0$, $\beta_{t,j+kT} = \beta_{t,j}\beta_{T,T}^k$, $\beta_{j+T,j+a} = \beta_{j,j}\beta_{T,a}$ e $\beta_{T,T} = \prod_{i=1}^T \alpha_i$.

A análise da existência e unicidade do processo PINAR(1)_T pode ser obtida escrevendo-o na forma de processo auto-regressivo de valores inteiros multivariado, como foi introduzido por Latour (1997). A equação (2.1) admite a representação vectorial

$$\mathbf{Y}_t = \mathbf{A} \circ \mathbf{Y}_{t-1} + \boldsymbol{\zeta}_t, \quad (2.2)$$

com $\mathbf{Y}_t = [X_{1+tT} \ X_{2+tT} \ \dots \ X_{T+tT}]'$ onde $'$ denota a transposta de uma matriz,

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & \dots & \beta_{1,1} \\ 0 & 0 & \dots & \beta_{2,2} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \beta_{T,T} \end{pmatrix}$$

e

$$\boldsymbol{\zeta}_t = \mathbf{B} \circ \mathbf{Z}_t = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ \beta_{2,1} & 1 & 0 & \dots & 0 & 0 \\ \beta_{3,2} & \beta_{3,1} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \beta_{T-1,T-2} & \beta_{T-1,T-3} & \beta_{T-1,T-4} & \dots & 1 & 0 \\ \beta_{T,T-1} & \beta_{T,T-2} & \beta_{T,T-3} & \dots & \beta_{T,1} & 1 \end{pmatrix} \circ \begin{pmatrix} Z_{1+tT} \\ Z_{2+tT} \\ \vdots \\ Z_{T+tT} \end{pmatrix},$$

sendo $(\boldsymbol{\zeta}_t)$ uma sucessão i.i.d. de vectores aleatórios de média finita $E[\boldsymbol{\zeta}_t] = \mathbf{B}[\lambda_1 \ \lambda_2 \ \dots \ \lambda_T]'$ e $V[\boldsymbol{\zeta}_t] = \mathbf{B} \boldsymbol{\Sigma}_{\mathbf{Z}} \mathbf{B}'$ onde $\boldsymbol{\Sigma}_{\mathbf{Z}} = Cov(\mathbf{Z}_t, \mathbf{Z}_{t+j}) = \text{diag}(\lambda_1, \dots, \lambda_T)$. Tendo em conta que os valores próprios da matriz \mathbf{A} são inferiores a 1 e que $\boldsymbol{\zeta}_t$ é independente de \mathbf{Y}_s , $s < t$ segue da proposição 3.1 em Latour (1997, p. 236) a existência quase certa de um processo estacionário único de valores inteiros não negativos que satisfaz (2.1). No resultado que se segue introduz-se a distribuição (ciclo)estacionária de X_t para cada um dos T períodos.

Teorema 2.2. *O processo (X_t) , com $t = j + kT$ (para um valor fixo de $j = 1, \dots, T$ e $k \in \mathbb{N}_0$) é uma cadeia de Markov irreduzível aperiódica e recorrente positiva (portanto*

ergódica) e a sua distribuição estacionária é a mesma distribuição de

$$V_j = \sum_{m=1}^{+\infty} \sum_{a=0}^{T-1} (\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m-1}) \circ Z_{T-a+Tm} + \sum_{m=0}^{j-1} \beta_{j,m} \circ Z_{j-m},$$

onde a série converge quase certamente e também em média quadrática.

Demonstração O processo X_t é uma cadeia de Markov em \mathbb{N}_0 com as seguintes probabilidades de transição

$$P_{ab}^{(1)} = P(X_{j+kT} = b | X_{j-1+kT} = a) = \sum_{m=0}^{\min(b,a)} \alpha_j^m (1 - \alpha_j)^{a-m} e^{-\lambda_j} \frac{\lambda_j^{b-m}}{(b-m)!} > 0, \quad (2.3)$$

com $a, b \in \mathbb{N}_0$.

A partir da expressão (2.3) conclui-se que o processo (X_{j+kT}) é uma cadeia de Markov irreduzível e aperiódica que, ou é recorrente positiva ou $\lim_{k \rightarrow +\infty} P_{ab}^{(j+kT)} = 0$, para qualquer $a, b \in \mathbb{N}_0$; ver por exemplo Ross (1983, p. 109). Aplicando sucessivamente a expressão recursiva definida por (2.1) tem-se

$$X_{j+kT} \stackrel{d}{=} \beta_{j+kT, j+kT} \circ X_0 + \sum_{m=0}^{kT-1} \beta_{j+kT, m} \circ Z_{j-m+kT} + \sum_{m=0}^{j-1} \beta_{j, m} \beta_{T, T}^k \circ Z_{j-m},$$

que, após o rearranjo de alguns termos (os coeficientes da operação *thinning* associados a $Z_1, \dots, Z_j, \dots, Z_{1+kT}, \dots, Z_{j+kT}$ passam a estar associados a $Z_{1+kT}, \dots, Z_{j+kT}, \dots, Z_1, \dots, Z_j$ respectivamente) dá origem a

$$\begin{aligned} X_{j+kT} &\stackrel{d}{=} (\beta_{j,j} \beta_{T,T}^k) \circ X_0 + \sum_{m=1}^k \sum_{a=0}^{T-1} \beta_{j+kT, j+a+T(m-1)} \circ Z_{T(m+1)-a} + \sum_{m=0}^{j-1} \beta_{j, m} \circ Z_{j-m} \\ &\stackrel{d}{=} (\beta_{j,j} \beta_{T,T}^k) \circ X_0 + \sum_{m=1}^k \sum_{a=0}^{T-1} (\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m-1}) \circ Z_{T-a+Tm} + \sum_{m=0}^{j-1} \beta_{j, m} \circ Z_{j-m} \quad (2.4) \\ &\stackrel{d}{=} Y_{j+kT}. \end{aligned}$$

Note-se que a primeira parcela da expressão (2.4) tende em probabilidade para zero, portanto a igualdade em distribuição não depende de estar condicionado a X_0 . De seguida

mostra-se que Y_{j+kT} converge quase certamente usando o critério de Cauchy (ver Rohatgi, 1976, p. 270). De facto, para todo o $\epsilon > 0$ e $k, n \in \mathbb{N}_0$ tem-se

$$\begin{aligned}
& P\left(\max_{1 \leq l \leq n} |Y_{j+kT} - Y_{j+T(k+l)}| > \epsilon\right) = \\
& = P\left(\max_{1 \leq l \leq n} \left| \begin{aligned} & (\beta_{j,j}\beta_{T,T}^k) \circ X_0 - (\beta_{j,j}\beta_{T,T}^{k+l}) \circ X_0 - \\ & - \sum_{m=1}^l \sum_{a=0}^{T-1} (\beta_{j,a}\beta_{T,a}\beta_{T,T}^{m+k-1}) \circ Z_{T(m+k+1)-a} \end{aligned} \right| > \epsilon\right) \\
& \leq P\left(\max_{1 \leq l \leq n} \left\{ \begin{aligned} & |(\beta_{j,j}\beta_{T,T}^k) \circ X_0 - (\beta_{j,j}\beta_{T,T}^{k+l}) \circ X_0| + \\ & + \sum_{m=1}^l \sum_{a=0}^{T-1} (\beta_{j,a}\beta_{T,a}\beta_{T,T}^{m+k-1}) \circ Z_{T(m+k+1)-a} \end{aligned} \right\} > \epsilon\right) \\
& \leq P\left(|(\beta_{j,j}\beta_{T,T}^k) \circ X_0 - (\beta_{j,j}\beta_{T,T}^{k+n}) \circ X_0| + \sum_{m=1}^n \sum_{a=0}^{T-1} (\beta_{j,a}\beta_{T,a}\beta_{T,T}^{m+k-1}) \circ Z_{T(m+k+1)-a} > \epsilon\right) \\
& \quad \text{uma vez que } \alpha \circ X \leq X \\
& \leq \frac{E\left[\beta_{j,j}\beta_{T,T}^k \circ X_0 - \beta_{j,j}\beta_{T,T}^{k+n} \circ X_0 + \sum_{m=1}^n \sum_{a=0}^{T-1} (\beta_{j,a}\beta_{T,a}\beta_{T,T}^{m+k-1}) \circ Z_{T(m+k+1)-a}\right]}{\epsilon} \\
& = \frac{\beta_{j,j}\beta_{T,T}^k(1 - \beta_{T,T}^n)E[X_0] + \beta_{j,j}\beta_{T,T}^{k-1} \sum_{m=1}^n \sum_{a=0}^{T-1} \beta_{T,a}\beta_{T,T}^m \lambda_{T-a}}{\epsilon} \longrightarrow 0, \text{ quando } k \rightarrow +\infty.
\end{aligned}$$

Note-se que a variável $V_j^{(k)} = \sum_{m=1}^k \sum_{a=0}^{T-1} \beta_{j,j}\beta_{T,a}\beta_{T,T}^{m-1} \circ Z_{T(m+1)-a} + \sum_{m=0}^{j-1} \beta_{j,m} \circ Z_{j-m}$ converge em distribuição para o mesmo limite que Y_{j+kT} , quer condicional ou incondicionalmente a X_0 , uma vez que $V_j^{(k)}$ é independente de X_0 . Consequentemente, tem-se

$$\begin{aligned}
\lim_{k \rightarrow +\infty} P(X_{j+kT} = b | X_0 = a) & = \lim_{k \rightarrow +\infty} P(Y_{j+kT} = b | X_0 = a) \\
& = \lim_{k \rightarrow +\infty} P(V_j^{(k)} = b | X_0 = a) \\
& = \lim_{k \rightarrow +\infty} P(V_j^{(k)} = b) \\
& = \lim_{k \rightarrow +\infty} P(Y_{j+kT} = b) = \lim_{k \rightarrow +\infty} P(X_{j+kT} = b).
\end{aligned}$$

Pretende-se mostrar por redução ao absurdo que X_t é recorrente positiva. Considere-se que a cadeia não é recorrente positiva; então para quaisquer $a, b \in \mathbb{N}_0$ tem-se

$\lim_{k \rightarrow +\infty} P_{ab}^{(j+kT)} = 0 = P(Y_j^* = b)$, onde Y_j^* representa o limite quase certo de Y_{j+kT} . Mas isto contradiz o facto de $\sum_i P(Y_j^* = i) = 1$, logo X_{j+kT} é recorrente positiva.

A convergência quase certa de $V_j^{(k)}$ é obtida usando os mesmos argumentos utilizados para mostrar a convergência quase certa de Y_{j+kT} . Para terminar a prova falta apenas mostrar que $V_j^{(k)}$ converge em média quadrática para V_j , e para tal mostra-se que V_{j+kT} é uma sucessão de Cauchy em L^2 .

$$\begin{aligned}
E \left[\left(V_j^{(k)} - V_j^{(k+n)} \right)^2 \right] &= \\
&= \sum_{m=1}^n \sum_{\substack{l=1 \\ l \neq m}}^n \sum_{a=0}^{T-1} \sum_{b=0}^{T-1} \beta_{j,j}^2 \beta_{T,a} \beta_{T,b} \beta_{T,T}^{m+l+2k-2} E \left[Z_{T(m+k+1)-a} Z_{T(l+k+1)-b} \right] \\
&+ \sum_{m=1}^n \sum_{a=0}^{T-1} \left(\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m+k-1} \lambda_{T-a} + \beta_{j,j}^2 \beta_{T,a}^2 \beta_{T,T}^{2m+2k-2} \lambda_{T-a}^2 \right) \\
&+ \sum_{m=1}^n \sum_{a=0}^{T-1} \sum_{\substack{b=0 \\ b \neq a}}^{T-1} \beta_{j,j}^2 \beta_{T,a} \beta_{T,b} \beta_{T,T}^{2m+2k-2} E \left[Z_{T(m+k+1)-a} Z_{T(m+k+1)-b} \right] \\
&= \sum_{m=1}^n \sum_{\substack{l=1 \\ l \neq m}}^n \sum_{a=0}^{T-1} \sum_{b=0}^{T-1} \beta_{j,j}^2 \beta_{T,a} \beta_{T,b} \beta_{T,T}^{m+l+2k-2} \lambda_{T-a} \lambda_{T-b} \\
&+ \sum_{m=1}^n \sum_{a=0}^{T-1} \left(\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m+k-1} \lambda_{T-a} + \beta_{j,j}^2 \beta_{T,a}^2 \beta_{T,T}^{2m+2k-2} \lambda_{T-a}^2 \right) \\
&+ \sum_{m=1}^n \sum_{a=0}^{T-1} \sum_{\substack{b=0 \\ b \neq a}}^{T-1} \beta_{j,j}^2 \beta_{T,a} \beta_{T,b} \beta_{T,T}^{2m+2k-2} \lambda_{T-a} \lambda_{T-b}.
\end{aligned}$$

Após alguns cálculos morosos, que apenas exigem manipulação algébrica, obtém-se

$$\begin{aligned}
E \left[\left(V_j^{(k)} - V_j^{(k+n)} \right)^2 \right] &= \beta_{T,T}^{2k+1} \beta_{j,j}^2 \mu_T^2 (1 - \beta_{T,T}^{n-1})^2 + \\
&+ \beta_{T,T}^k \beta_{j,j} \mu_T (1 - \beta_{T,T}^n) + \beta_{T,T}^{2k} \beta_{j,j}^2 (1 - \beta_{T,T}^{2n}) \mu_T^2 \\
&\longrightarrow 0, \quad k \rightarrow +\infty.
\end{aligned}$$

Portanto, para todo $n > 0$, $E \left[\left| V_j^{(k)} - V_j^{(k+n)} \right|^2 \right] \longrightarrow 0$, quando $k \rightarrow +\infty$, o que conclui a prova. \blacksquare

A partir do resultado anterior obtêm-se a média periódica, a variância periódica e a função de autocovariância de X_t .

Lema 2.3. Para um valor fixo de $j = 1, \dots, T$, com $T \in \mathbb{N}$, $t = j + kT$ e $k \in \mathbb{N}_0$, tem-se

1. A média $E[X_t]$ e a variância $V[X_t]$ são iguais a μ_j , dada por

$$\mu_j = \frac{\sum_{k=0}^{j-1} \beta_{j,k} \lambda_{j-k} + \beta_{j,j} \sum_{i=0}^{T-j-1} \beta_{T,i} \lambda_{T-i}}{1 - \beta_{T,T}}, \quad (2.5)$$

com a convenção $\sum_{i=0}^{-1} a_i = 0$, $\forall a_i$.

2. A função de autocovariância de X_t é

$$\gamma_{j+kT}(h) = \gamma_j(h) = \begin{cases} \beta_{j+h,h} \mu_j & h \geq 0 \\ \beta_{j+kT,|h|} \mu_{j+kT-|h|} & h < 0 \end{cases}. \quad (2.6)$$

Demonstração Como a distribuição estacionária de X_t é a mesma de V_j então é suficiente determinar a média e a variância desta variável.

$$\begin{aligned} E[V_j] &= \lim_{k \rightarrow +\infty} \sum_{m=1}^k \sum_{a=0}^{T-1} E((\beta_{j,j} \beta_{T,a} \beta_{T,T}^{m-1}) \circ Z_{T-a+Tm}) + \sum_{m=0}^{j-1} E(\beta_{j,m} \circ Z_{j-m}) \\ &= \lim_{k \rightarrow +\infty} \sum_{m=1}^k \sum_{a=0}^{T-1} \beta_{j,j} \beta_{T,a} \beta_{T,T}^{m-1} \lambda_{T-a} + \sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m} \\ &= \sum_{a=0}^{T-1} \beta_{j,j} \beta_{T,a} \lambda_{T-a} \lim_{k \rightarrow +\infty} \sum_{m=1}^k \beta_{T,T}^{m-1} + \sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m} \\ &= \beta_{j,j} \sum_{a=0}^{T-1} \beta_{T,a} \lambda_{T-a} \frac{1}{1 - \beta_{T,T}} + \sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m} \\ &= \frac{\beta_{j,j} \sum_{a=0}^{T-1} \beta_{T,a} \lambda_{T-a} + (1 - \beta_{T,T}) \sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m}}{1 - \beta_{T,T}}. \end{aligned}$$

Desdobrando o primeiro somatório e juntando os termos pretendidos obtém-se

$$\begin{aligned}
 E[V_j] &= \\
 &= \frac{\sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m} + \beta_{j,j} \sum_{a=0}^{T-j-1} \beta_{T,a} \lambda_{T-a} + \beta_{j,j} \sum_{a=T-j}^{T-1} \beta_{T,a} \lambda_{T-a} - \beta_{T,T} \sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m}}{1 - \beta_{T,T}} \\
 &= \frac{\sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m} + \beta_{j,j} \sum_{a=0}^{T-j-1} \beta_{T,a} \lambda_{T-a} + \beta_{j,j} \sum_{m=0}^{j-1} \beta_{T,T-j+m} \lambda_{j-m} - \beta_{j,j} \beta_{T,T-j} \sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m}}{1 - \beta_{T,T}} \\
 &= \frac{\sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m} + \beta_{j,j} \sum_{a=0}^{T-j-1} \beta_{T,a} \lambda_{T-a} + \beta_{j,j} \sum_{m=0}^{j-1} \beta_{T,T-j+m} \lambda_{j-m} - \beta_{j,j} \sum_{m=0}^{j-1} \beta_{T,T-j+m} \lambda_{j-m}}{1 - \beta_{T,T}} \\
 &= \frac{\sum_{m=0}^{j-1} \beta_{j,m} \lambda_{j-m} + \beta_{j,j} \sum_{a=0}^{T-j-1} \beta_{T,a} \lambda_{T-a}}{1 - \beta_{T,T}} \\
 &= \mu_j.
 \end{aligned}$$

Como V_j é definida à custa da sucessão de v.a.'s independentes Z_t que têm distribuição de Poisson, significa que cada parcela de V_j tem distribuição de Poisson o que permite afirmar que a variância de V_j é igual à sua média.

Considere-se agora a função de autocovariância e $h > 0$. Esta função não é simétrica, como acontece no caso dos modelos usuais INAR e AR, sendo necessário determinar a expressão para valores positivos e negativos. Para valores positivos toma a forma

$$\begin{aligned}
 \gamma_t(h) &= Cov(X_t, X_{t+h}) \\
 &= Cov\left(X_t, \beta_{t+h,h} \circ X_t + \sum_{i=0}^{h-1} \beta_{t+h,i} \circ Z_{t+h-i}\right) \\
 &= \beta_{t+h,h} Cov(X_t, X_t) \\
 &= \beta_{t+h,h} \gamma_t(0) \\
 &= \beta_{t+h,h} \mu_t;
 \end{aligned}$$

ou seja, para $t = j + kT$, com $j = 1, \dots, T$ e $k \in \mathbb{N}$, é possível escrever a função de

autocovariância na forma

$$\begin{aligned}\gamma_{j+kT}(h) &= Cov(X_{j+kT}, X_{j+kT+h}) \\ &= \beta_{j+kT+h,h}\mu_{j+kT} \\ &= \beta_{j+h,h}\mu_j.\end{aligned}$$

Para valores negativos obtém-se

$$\begin{aligned}\gamma_t(-h) &= Cov(X_t, X_{t-h}) \\ &= Cov\left(\beta_{t,h} \circ X_{t-h} + \sum_{i=0}^{h-1} \beta_{t,i} \circ Z_{t-i}, X_{t-h}\right) \\ &= \beta_{t,h}Cov(X_{t-h}, X_{t-h}) \\ &= \beta_{t,h}\gamma_{t-h}(0) \\ &= \beta_{t,h}\mu_{t-h},\end{aligned}$$

ou seja, para $t = j + kT$, com $j = 1, \dots, T$ e $k \in \mathbb{N}$ tem-se

$$\begin{aligned}\gamma_{j+kT}(-h) &= Cov(X_{j+kT}, X_{j+kT-h}) \\ &= \beta_{j+kT,h}\mu_{j+kT-h}.\end{aligned}$$

■

É de realçar que a média é uma combinação linear dos parâmetros λ 's cujos coeficientes são funções dos parâmetros α 's e que pode ser calculada através da expressão alternativa

$$\mu_j = \beta_{j,j}(\mu_T + \frac{1}{\beta_{j,j}} \sum_{k=0}^{j-1} \beta_{j,k}\lambda_{j-k}), \quad j = 1, \dots, T.$$

A função de autocovariância $\gamma_j(\cdot)$ não é simétrica; no entanto tem-se que $\gamma_t(-h) = \gamma_{t-h}(h)$ e $\gamma_t(h) = \gamma_{t+h}(-h)$. Tendo em conta que h pode ser reescrito na forma $h = i + mT$, para algum $i \in \{1, \dots, T\}$ e $m \in \mathbb{N}_0$, a função de autocovariância pode assumir a forma $\gamma_j(h) = \beta_{T,T}^m \beta_{j+i,i} \mu_j$ e $\gamma_j(-h) = \beta_{T,T}^m \beta_{j+T,i} \mu_{j+T-i}$, o que significa que, para cada j fixo o decaimento da função autocovariância é exponencial.

O Teorema 2.4 que se apresenta define a distribuição marginal do processo (X_t) , descrito por (2.1).

Teorema 2.4. *A distribuição marginal de (X_t) , com $t = j + kT$ para um valor fixo de $j = 1, \dots, T$, com $T \in \mathbb{N}$, e $k \in \mathbb{N}_0$ é Poisson de média μ_j se e só se (Z_t) for uma sucessão de v.a.'s independentes com distribuição de Poisson de parâmetro λ_j .*

Demonstração Para provar este teorema utiliza-se a função geradora de probabilidades associada a X_{j+kT} , $j = 1, \dots, T$. Aplicando recursivamente a equação (2.1) T vezes, X_{j+kT} pode ser escrito na forma

$$\begin{aligned} X_{j+kT} &= \beta_{T,T} \circ X_{j+T(k-1)} + \sum_{i=0}^{T-1} \beta_{j+kT,i} \circ Z_{j+kT-i} \\ &= \beta_{T,T} \circ X_{j+T(k-1)} + \underbrace{\sum_{i=0}^{j-1} \beta_{j,i} \circ Z_{j-i+kT} + \sum_{i=0}^{T-j-1} \beta_{j,j} \beta_{T,i} \circ Z_{T-i+(k-1)T}}_{Y_j^*}. \end{aligned}$$

Tendo em conta a f.g.p. de uma v.a. que resulta da aplicação do operador *binomial thinning*, a função geradora de probabilidades de X_{j+kT} é dada por

$$P_{X_{j+kT}}(r) = P_{X_{j+T(k-1)}}(1 - \beta_{T,T} + \beta_{T,T}r) \times P_{Y_j^*}(r). \quad (2.7)$$

" \Rightarrow "

Assuma-se que $Z_{j+kT} \sim P(\lambda_j)$. Como (Z_{j+kT}) é uma sucessão de v.a.'s independentes com distribuição de Poisson de parâmetro λ_j então, pelas propriedades da operação *thinning*, todas as parcelas de Y_j^* são independentes e têm distribuição de Poisson, garantindo assim que a distribuição de Y_j^* é Poisson com média $\sum_{i=0}^{j-1} \beta_{j,i} \lambda_{j-i} + \sum_{i=0}^{T-j-1} \beta_{j,j} \beta_{T,i} \lambda_{T-i}$. Tendo em conta a expressão da média periódica (2.5), pode-se afirmar que Y_j^* tem média $(1 - \beta_{T,T})\mu_j$, o que permite concluir a partir de (2.7) que

$$\frac{P_{X_{j+kT}}(r)}{P_{X_{j+T(k-1)}}(1 - \beta_{T,T} + \beta_{T,T}r)} = e^{-\mu_j(1-\beta_{T,T})(1-r)} = \frac{e^{-\mu_j(1-r)}}{e^{-\mu_j(1-(1-\beta_{T,T}+\beta_{T,T}r))}}.$$

A partir desta última igualdade conclui-se que $P_{X_{j+kT}}(r) = e^{-\mu_j(1-r)}$, que é a função geradora de probabilidades de uma v.a. com distribuição de Poisson de parâmetro μ_j .

"←"

Se X_{j+kT} tem distribuição de Poisson de parâmetro μ_j então por (2.7) tem-se

$$\begin{aligned} e^{-\mu_j(1-r)} &= P_{X_{j+T(k-1)}}(1 - \beta_{T,T} + \beta_{T,T}r) \times P_{Y_{j^*}}(r) \\ &= e^{-\mu_j\beta_{T,T}(1-r)} \times P_{Y_{j^*}}(r), \end{aligned}$$

o que significa que

$$\begin{aligned} P_{Y_{j^*}}(r) &= e^{-\mu_j(1-\beta_{T,T})(1-r)} \\ &= \prod_{i=0}^{j-1} e^{-\beta_{j,i}\lambda_{j-i}(1-r)} \prod_{i=0}^{T-j-1} e^{-\beta_{j,j}\beta_{T,i}\lambda_{T-i}(1-r)} \end{aligned} \quad (2.8)$$

$$= \prod_{i=0}^{j-1} P_{Z_{j-i+kT}}(1 - \beta_{j,i} + \beta_{j,i}r) \prod_{i=0}^{T-j-1} P_{Z_{T-i+kT}}(1 - \beta_{j,j}\beta_{T,i} + \beta_{j,j}\beta_{T,i}r), \quad (2.9)$$

que é a função geradora de probabilidades de uma v.a. com distribuição de Poisson com média $\mu_j(1 - \beta_{T,T}) = \sum_{m=0}^{j-1} \beta_{j,m}\lambda_{j-m} + \beta_{j,j} \sum_{m=0}^{T-j-1} \beta_{T,m}\lambda_{T-m}$.

De (2.8) e (2.9) conclui-se que

$$\begin{aligned} P_{Z_{j-i+kT}}(1 - \beta_{j,i} + \beta_{j,i}r) &= e^{-\beta_{j,i}\lambda_{j-i}(1-r)} \\ &= e^{-\lambda_{j-i}(1-1-\beta_{j,i}+\beta_{j,i}r)} \end{aligned}$$

e

$$\begin{aligned} P_{Z_{T-i+kT}}(1 - \beta_{j,j}\beta_{T,i} + \beta_{j,j}\beta_{T,i}r) &= e^{-\lambda_{T-i}\beta_{T,i}\beta_{j,j}(1-r)} \\ &= e^{-\lambda_{T-i}(1-1+\beta_{T,i}\beta_{j,j}-\beta_{T,i}\beta_{j,j}r)}, \end{aligned}$$

pelo que $P_{Z_{j-i+kT}}(s) = e^{-\lambda_{j-i}(1-s)}$ e $P_{Z_{T-i+kT}}(r) = e^{-\lambda_{T-i}(1-r)}$. Como são as funções geradoras de probabilidade de v.a.'s com distribuições de Poisson, i.e., $Z_{i+kT} \sim P(\lambda_i)$, para $i = 1, \dots, T$, fica concluída a demonstração. ■

2.2 Métodos de estimação

Seja (X_1, \dots, X_{NT}) uma sucessão de v.a.'s que satisfazem (2.1) e $\boldsymbol{\theta} = (\alpha_1, \lambda_1, \dots, \alpha_T, \lambda_T)$ o vector dos parâmetros desconhecidos que se pretende estimar. A notação utilizada pressupõe a existência de N ciclos completos de observações e, sem perda de generalidade é também assumido que $X_0 = x_0$. Nesta secção os métodos de estimação que vão estar em destaque são os usuais no domínio do tempo, nomeadamente o método baseado nos momentos, os métodos dos mínimos quadrados condicionais simples e ponderados e o método da máxima verosimilhança condicional.

2.2.1 Estimadores baseados nos momentos

Uma forma simples de estimar os parâmetros do modelo INAR periódico é usar o método de estimação de Yule-Walker que consiste na substituição da função autocovariância pela respectiva função de autocovariância amostral. Para determinar $\alpha_1, \dots, \alpha_T$ basta usar $\gamma_t(1)$ para os diversos períodos. A partir do Lema 2.3 tem-se que

$$\begin{cases} \gamma_{i+kT}(1) = \alpha_{i+1}\sigma_i^2, & i = 1, \dots, T; \\ \mu_i = \alpha_i\mu_{i-1} + \lambda_i \end{cases} \quad (2.10)$$

onde $\mu_0 = \mu_T$. Substituindo $\gamma_{i+kT}(1)$, μ_i , e σ_i^2 pelas correspondentes estimativas amostrais, para $i = 1, \dots, T$, obtêm-se os estimadores de Yule-Walker (EYW) $\hat{\alpha}_{i,YW}$ e $\hat{\lambda}_{i,YW}$, de α_i e λ_i , dados por

$$\hat{\lambda}_{i,YW} = \begin{cases} \bar{X}_{i,N} - \hat{\alpha}_i \bar{X}_{i-1,N}, & i = 2, \dots, T \\ \bar{X}_1 - \hat{\alpha}_1 \bar{X}_{T,N}, & i = 1 \end{cases}, \quad (2.11)$$

e

$$\hat{\alpha}_{i,YW} = \begin{cases} \frac{N \sum_{k=0}^{N-1} (X_{kT} - \bar{X}_{T,N})(X_{1+nT} - \bar{X}_{1,N})}{(N-1) \sum_{k=0}^{N-1} (X_{kT} - \bar{X}_{T,N})^2}, & i = 1 \\ \frac{\sum_{k=0}^{N-1} (X_{i-1+kT} - \bar{X}_{i-1,N})(X_{i+kT} - \bar{X}_{i,N})}{\sum_{k=0}^{N-1} (X_{i-1+kT} - \bar{X}_{i-1,N})^2}, & i = 2, \dots, T \end{cases} \quad (2.12)$$

com

$$\bar{X}_{i,N} = \begin{cases} \frac{1}{N} \sum_{k=0}^{N-1} X_{i+kT}, & i = 1, \dots, T-1 \\ \frac{1}{N+1} \sum_{k=0}^N X_{kT}, & i = T \end{cases}. \quad (2.13)$$

Posteriormente ir-se-á mostrar que os estimadores de Yule-Walker são assintoticamente equivalentes aos estimadores dos mínimos quadrados condicionais e portanto possuem as mesmas propriedades assintóticas que estes últimos, nomeadamente a mesma distribuição assintótica.

2.2.2 Estimadores dos mínimos quadrados condicionais

A aplicação do método de estimação proposto por Klimko e Nelson (1978) para estimar o vector de parâmetros $\boldsymbol{\theta}$ do modelo descrito em (2.1), permite obter os estimadores $\hat{\boldsymbol{\theta}}_{MQC} = (\hat{\alpha}_{1,MQC}, \hat{\lambda}_{1,MQC}, \dots, \hat{\alpha}_{T,MQC}, \hat{\lambda}_{T,MQC})$ que minimizam a expressão

$$Q(\boldsymbol{\theta}) = \sum_{k=0}^{N-1} \sum_{i=1}^T (X_{i+kT} - \alpha_i X_{i+kT-1} - \lambda_i)^2.$$

Estes estimadores são solução dos seguintes T sistemas

$$\begin{cases} \sum_{k=0}^{N-1} X_{i+kT-1} (X_{i+kT} - \alpha_i X_{i+kT-1} - \lambda_i) = 0 \\ \sum_{k=0}^{N-1} (X_{i+kT} - \alpha_i X_{i+kT-1} - \lambda_i) = 0 \end{cases}, \quad i = 1, \dots, T; \quad (2.14)$$

resultando em

$$\begin{cases} \hat{\alpha}_{i,MQC} = \frac{N \sum_{k=0}^{N-1} X_{i+kT} X_{i-1+kT} - \sum_{k=0}^{N-1} X_{i+kT} \sum_{k=0}^{N-1} X_{i-1+kT}}{N \sum_{k=0}^{N-1} X_{i-1+kT}^2 - \left(\sum_{k=0}^{N-1} X_{i-1+kT} \right)^2} \\ \hat{\lambda}_{i,MQC} = N^{-1} \left(\sum_{k=0}^{N-1} X_{i+kT} - \hat{\alpha}_{i,MQC} \sum_{k=0}^{N-1} X_{i-1+kT} \right) \end{cases}, \quad i = 1, \dots, T. \quad (2.15)$$

Teorema 2.5. *No modelo PINAR(1)_T verifica-se que:*

- (a) $\hat{\alpha}_{i,YW} = \hat{\alpha}_{i,MQC}$, $i = 1, \dots, T - 1$, e $\hat{\lambda}_{i,YW} = \hat{\lambda}_{i,MQC}$, $i = 2, \dots, T - 1$;
- (b) $\hat{\lambda}_{T,MQC} = \hat{\lambda}_{T,YW} + o_p(1)$, $\hat{\alpha}_{1,MQC} = \hat{\alpha}_{1,YW} + o_p(1)$ e $\hat{\lambda}_{1,MQC} = \hat{\lambda}_{1,YW} + o_p(1)$.

Demonstração Considere-se

$$S_{i,N}^2 = \begin{cases} \frac{1}{N-1} \sum_{n=0}^{N-1} (X_{i+nT} - \bar{X}_{i,N})^2 & i = 1, \dots, T-1 \\ \frac{1}{N} \sum_{n=0}^N (X_{nT} - \bar{X}_T)^2 & i = T \end{cases}.$$

Para mostrar (a) basta atender à definição da média amostral $\bar{X}_{i,N}$, $i = 1, 2, \dots, T$ usada em (2.13), para se concluir que os estimadores de Yule-Walker e dos mínimos quadrados

condicionais para λ_i , $i = 2, \dots, T-1$, coincidem se os respectivos estimadores dos parâmetros da operação *binomial thinning*, α_i , coincidirem. Com efeito, tendo em conta as expressões em (2.12) tem-se, para $i = 2, \dots, T$,

$$\begin{aligned}
& \hat{\alpha}_{i,YW} - \hat{\alpha}_{i,MQC} = \\
&= \frac{\sum_{n=0}^{N-1} (X_{i+nT} - \bar{X}_{i,N})(X_{i-1+nT} - \bar{X}_{i-1,N})}{\sum_{n=0}^{N-1} (X_{i-1+nT} - \bar{X}_{i-1,N})^2} - \frac{N \sum_{n=0}^{N-1} X_{i-1+nT} X_{nT} - \sum_{n=0}^{N-1} X_{i-1+nT} \sum_{n=0}^{N-1} X_{i+nT}}{N \sum_{n=0}^{N-1} X_{i-1+nT}^2 - \left(\sum_{n=0}^{N-1} X_{i-1+nT} \right)^2} \\
&= \frac{\sum_{n=0}^{N-1} X_{i+nT} X_{i-1+nT} - \sum_{n=0}^{N-1} X_{i+nT} \bar{X}_{i-1,N} - \sum_{n=0}^{N-1} \bar{X}_{i,N} X_{i-1+nT} + N \bar{X}_{i,N} \bar{X}_{i-1,N}}{NS_{i-1,N}^2} - \\
& \quad - \frac{\sum_{n=0}^{N-1} X_{i-1+nT} X_{i+nT} - \frac{1}{N} \sum_{n=0}^{N-1} X_{i-1+nT} \sum_{n=0}^{N-1} X_{i+nT}}{NS_{i-1,N}^2} \\
&= 0.
\end{aligned}$$

Para mostrar (b) comece-se por verificar que os estimadores em causa para λ_T diferem de uma quantidade assintoticamente desprezável. De facto

$$\begin{aligned}
& \hat{\lambda}_{T,MQC} - \hat{\lambda}_{T,YW} = \\
&= \frac{\sum_{n=0}^{N-1} X_{T+nT} - \hat{\alpha}_{T,MQC} \sum_{n=0}^{N-1} X_{T-1+nT}}{N} - \frac{\sum_{n=0}^N X_{nT} - \hat{\alpha}_{T,YW} N^{-1} \sum_{n=0}^{N-1} X_{T-1+nT}}{N+1} \\
&= N^{-1} \sum_{n=0}^{N-1} X_{T-1+nT} (\hat{\alpha}_{T,YW} - \hat{\alpha}_{T,MQC}) + N^{-1} \sum_{n=0}^{N-1} X_{T+nT} - (N+1)^{-1} \sum_{n=0}^N X_{nT} \\
&= (\hat{\alpha}_{T,YW} - \hat{\alpha}_{T,MQC}) \bar{X}_{T-1,N} + \bar{X}_{T,N-1} - \bar{X}_{T,N} \\
&= \bar{X}_{T,N-1} - \bar{X}_{T,N} \\
&\xrightarrow{P} 0.
\end{aligned}$$

No caso de $i = 1$, os estimadores de α_1 e de λ_1 não coincidem mas diferem de uma quantidade que é assintoticamente desprezável. De seguida verifica-se que tal acontece,

tendo por base o mesmo raciocínio usado por Freeland e McCabe (2005). Ora

$$\begin{aligned}
& \hat{\alpha}_{1;YW} - \hat{\alpha}_{1,MQC} = \\
& = \frac{N \sum_{n=0}^{N-1} (X_{nT} - \bar{X}_{T,N})(X_{1+nT} - \bar{X}_{1,N})}{(N-1) \sum_{n=0}^N (X_{nT} - \bar{X}_{T,N})^2} - \frac{N \sum_{n=0}^{N-1} X_{1+nT} X_{nT} - \sum_{n=0}^{N-1} X_{1+nT} \sum_{n=0}^{N-1} X_{nT}}{N \sum_{n=0}^{N-1} X_{nT}^2 - \left(\sum_{n=0}^{N-1} X_{nT} \right)^2} \\
& = \frac{\sum_{n=0}^{N-1} (X_{nT} - \bar{X}_{T,N})(X_{1+nT} - \bar{X}_{1,N})}{(N-1) S_{T,N}^2} - \frac{\sum_{n=0}^{N-1} X_{1+nT} X_{nT} - \frac{1}{N} \sum_{n=0}^{N-1} X_{1+nT} \left(\sum_{n=0}^N X_{nT} - X_{NT} \right)}{\sum_{n=0}^{N-1} X_{nT}^2 - \frac{1}{N} \left(\sum_{n=0}^{N-1} X_{nT} \right)^2} \\
& = \frac{\sum_{n=0}^{N-1} X_{nT} X_{1+nT} - \bar{X}_{1,N} \sum_{n=0}^{N-1} X_{nT} - N \bar{X}_{T,N} \bar{X}_{1,N} + N \bar{X}_{T,N} \bar{X}_{1,N}}{(N-1) S_{T,N}^2} - \\
& \quad - \frac{\sum_{n=0}^{N-1} X_{1+nT} X_{nT} - (N+1) \bar{X}_{1,N} \bar{X}_{T,N} + \bar{X}_{1,N} X_{NT}}{(N-1) S_{T,N-1}^2} \\
& = \frac{\sum_{n=0}^{N-1} X_{nT} X_{1+nT} (S_{T,N-1}^2 - S_{T,N}^2) - S_{T,N-1}^2 \bar{X}_{1,N} \left(\sum_{n=0}^N X_{nT} - X_{NT} \right)}{(N-1) S_{T,N}^2 S_{T,N-1}^2} + \\
& \quad + \frac{S_{T,N}^2 (-(N+1) \bar{X}_{1,N} \bar{X}_{T,N} - \bar{X}_{1,N} X_{NT})}{(N-1) S_{T,N}^2 S_{T,N-1}^2} \\
& = \frac{(S_{T,N-1}^2 - S_{T,N}^2) \sum_{n=0}^{N-1} X_{nT} X_{1+nT} - (S_{T,N-1}^2 - S_{T,N}^2) (N+1) \bar{X}_{1,N} \bar{X}_{T,N}}{(N-1) S_{T,N-1}^2 S_{T,N}^2} + \\
& \quad + \frac{(S_{T,N-1}^2 - S_{T,N}^2) \bar{X}_{1,N} X_{NT}}{(N-1) S_{T,N-1}^2 S_{T,N}^2} \\
& = \frac{(S_{T,N-1}^2 - S_{T,N}^2)}{S_{T,N-1}^2 S_{T,N}^2} \left(\frac{\sum_{n=0}^{N-1} X_{nT} X_{1+nT}}{N-1} - \frac{N+1}{N-1} \bar{X}_{T,N} \bar{X}_{1,N} + \frac{X_{NT} \bar{X}_{1,N}}{N-1} \right)
\end{aligned}$$

$\xrightarrow{P} 0$

Analogamente,

$$\begin{aligned}
& \hat{\lambda}_{1,MQC} - \hat{\lambda}_{1,YW} = \\
& = N^{-1} \left(\sum_{n=0}^{N-1} X_{1+nT} - \hat{\alpha}_{1,MQC} \sum_{n=0}^{N-1} X_{nT} \right) - \left(N^{-1} \sum_{n=0}^{N-1} X_{1+nT} - \hat{\alpha}_{1,YW} (N+1)^{-1} \sum_{n=0}^N X_{nT} \right) \\
& = -\hat{\alpha}_{1,MQC} N^{-1} \sum_{n=0}^{N-1} X_{nT} + \hat{\alpha}_{1,YW} (N+1)^{-1} \sum_{n=0}^N X_{nT} \\
& = (N(N+1))^{-1} \left(N\hat{\alpha}_{1,YW} \sum_{n=0}^N X_{nT} - (N+1)\hat{\alpha}_{1,MQC} \sum_{n=0}^{N-1} X_{nT} \right) \\
& = (N(N+1))^{-1} (N\hat{\alpha}_{1,YW} - (N+1)\hat{\alpha}_{1,MQC}) \sum_{n=0}^N X_{nT} - N^{-1}\hat{\alpha}_{1,MQC} X_{NT} \\
& = N^{-1} (N\hat{\alpha}_{1,YW} - (N+1)\hat{\alpha}_{1,MQC}) \bar{X}_{T,N} + N^{-1}\hat{\alpha}_{1,MQC} X_{NT} \\
& = (\hat{\alpha}_{1,YW} - \hat{\alpha}_{1,MQC}) \bar{X}_{T,N} - N^{-1}\hat{\alpha}_{1,MQC} \bar{X}_{T,N} + N^{-1}\hat{\alpha}_{1,MQC} X_{NT} \\
& = (\hat{\alpha}_{1,YW} - \hat{\alpha}_{1,MQC}) \bar{X}_{T,N} + N^{-1}\hat{\alpha}_{1,MQC} (X_{NT} - \bar{X}_{T,N}) \\
& \xrightarrow{P} 0.
\end{aligned}$$

Conclui-se assim que $\hat{\alpha}_{i,YW} = \hat{\alpha}_{i,MQC}$, para $i = 2, \dots, T$, e que $\hat{\lambda}_{i,YW} = \hat{\lambda}_{i,MQC}$, para $i = 2, \dots, T-1$. Para os restantes casos mostrou-se que $\hat{\alpha}_{1,MQC} = \hat{\alpha}_{1,YW} + o_p(1)$, $\hat{\lambda}_{1,MQC} = \hat{\lambda}_{1,YW} + o_p(1)$ e $\hat{\lambda}_{T,MQC} = \hat{\lambda}_{T,YW} + o_p(1)$. ■

Para facilitar a obtenção da distribuição assintótica de $\hat{\boldsymbol{\theta}}_{MQC}$ introduz-se o seguinte lema.

Lema 2.6. *Seja $S_N(\boldsymbol{\theta}) = \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta})$ com*

$$\xi_k(\boldsymbol{\theta}) = \sum_{i=1}^T U_{i+kT} \left(\frac{\partial U_{i+kT}}{\partial \boldsymbol{\theta}} \right) = (U_{1+kT} X_{nT}, U_{1+nT}, \dots, U_{T+kT} X_{T-1nT}, U_{T-1+nT})',$$

onde $U_{i+kT} = X_{i+kT} - \alpha_i X_{i+kT-1} - \lambda_i$. As seguintes convergências são verificadas quando

$N \rightarrow \infty$.

$$(i) \quad N^{-1} \frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \xrightarrow{P} A_1, \quad (2.16)$$

$$(ii) \quad N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})' \xrightarrow{P} A_2, \quad (2.17)$$

$$(iii) \quad N^{-1} S_N(\boldsymbol{\theta}) \xrightarrow{P} \vec{0}, \quad (2.18)$$

com

$$A_1 = \begin{bmatrix} \Psi_1 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & \Psi_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \Psi_3 & \dots & \mathbf{0}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \Psi_T \end{bmatrix} \quad e \quad A_2 = \begin{bmatrix} \Omega_1 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & \Omega_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \Omega_3 & \dots & \mathbf{0}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \Omega_T \end{bmatrix},$$

onde

$$\Psi_i = \begin{bmatrix} \mu_{i-1} + \mu_{i-1}^2 & \mu_{i-1} \\ \mu_{i-1} & 1 \end{bmatrix}, \quad \Omega_i = \begin{bmatrix} \alpha_i(1 - \alpha_i)m_{i-1,3} + \lambda_i m_{i-1,2} & \alpha_i(1 - \alpha_i)m_{i-1,2} + \lambda_i \mu_{i-1} \\ \alpha_i(1 - \alpha_i)m_{i-1,2} + \lambda_i \mu_{i-1} & \alpha_i(1 - \alpha_i)\mu_{i-1} + \lambda_i \end{bmatrix},$$

para $i=1, \dots, T$, com $\mu_0 = \mu_T$, e $m_{i,j} = E[X_{i+kT}^j]$, $m_{0,j} = E[X_{kT}^j]$, $\alpha_0 = \alpha_T$ e $\lambda_0 = \lambda_T$. $\vec{0} = [0 \dots 0]'$ um vector de dimensão $2T$.

Demonstração Para provar (2.16), basta provar que $B_{i,N} \xrightarrow{P} \Psi_i$, para $i = 1, \dots, T$.

Comece-se por notar que

$$N^{-1} \frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} B_{1,N} & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & B_{2,N} & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & B_{3,N} & \dots & \mathbf{0}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & B_{T,N} \end{bmatrix} \quad e \quad B_{i,N} = \begin{bmatrix} \frac{\sum_{n=0}^{N-1} X_{i-1+nT}^2}{N} & \frac{\sum_{n=0}^{N-1} X_{i-1+nT}}{N} \\ \frac{\sum_{n=0}^{N-1} X_{i-1+nT}}{N} & 1 \end{bmatrix}.$$

Em virtude dos elementos da diagonal secundária de $B_{i,N}$ serem iguais a \bar{X}_{i-1} (que tende em probabilidade para μ_{i-1}) e do elemento na posição (1,1) de $B_{i,N}$ ser $N^{-1} \sum_{n=0}^{N-1} X_{i-1+nT}^2$

que pode tomar a forma

$$N^{-1} \sum_{n=0}^{N-1} X_{i-1+nT}^2 = N^{-1} \sum_{n=0}^{N-1} (X_{i-1+nT} - \bar{X}_{i-1})^2 + \bar{X}_{i-1}^2 \xrightarrow{P} \sigma_{i-1}^2 + \mu_{i-1}^2 = \mu_{i-1} + \mu_{i-1}^2,$$

fica provado a convergência em probabilidade de $B_{i,N}$ para Ψ_i , para $i = 1, 2, \dots, T$.

Para provar (2.17), note-se que a matriz $N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})'$ por ser simétrica pode ser escrita na forma

$$N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})' = \begin{bmatrix} C_{N11} & C_{N12} & C_{N13} & \dots & C_{N1T} \\ C_{N12} & C_{N22} & C_{N23} & \dots & C_{N2T} \\ C_{N13} & C_{N23} & C_{N33} & \dots & C_{N3T} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_{N1T} & C_{N2T} & C_{N3T} & \dots & C_{NTT} \end{bmatrix}, \quad (2.19)$$

com

$$C_{Nij} = \begin{bmatrix} N^{-1} \sum_{k=0}^{N-1} U_{i+kT} X_{i-1+kT} U_{j+kT} X_{j-1+kT} & N^{-1} \sum_{k=0}^{N-1} U_{i+kT} U_{j+kT} X_{j-1+kT} \\ N^{-1} \sum_{k=0}^{N-1} U_{i+kT} X_{i-1+kT} U_{j+kT} & N^{-1} \sum_{k=0}^{N-1} U_{i+kT} U_{j+kT} \end{bmatrix}, \quad j \geq i.$$

O elemento (m, l) da matriz C_{Nij} , denotado por $(C_{Nij})_{ml}$, para $i = 1, \dots, T$, $j \geq i$, e $l, m = 1, 2$, pode ser expresso na forma $(C_{Nij})_{ml} = N^{-1} \sum_{k=0}^{N-1} U_{i+kT} X_{i-1+kT}^{2-l} U_{j+kT} X_{j-1+kT}^{2-m}$.

Consequentemente os valores médios

$$E[(C_{Nij})_{ml}] = N^{-1} \sum_{k=0}^{N-1} E[U_{i+kT} X_{i-1+kT}^{2-l} X_{j-1+kT}^{2-m} \underbrace{E[U_{j+kT} | X_{i-1+kT}, U_{i+kT}, X_{j-1+kT}]}_0] = 0, \quad i \neq j$$

e

$$\begin{aligned} E[(C_{Nii})_{ml}] &= N^{-1} \sum_{k=0}^{N-1} E[X_{i-1+kT}^{4-m-l} \underbrace{E[(X_{i+kT} - \alpha_i X_{i-1+kT} - \lambda_i)^2 | X_{i-1+kT}]}_{V[X_{i+kT} | X_{i-1+kT}]}] \\ &= N^{-1} \sum_{k=0}^{N-1} E[X_{i-1+kT}^{4-m-l} (\alpha_i (1 - \alpha_i) X_{i-1+kT} + \lambda_i)] \\ &= \alpha_i (1 - \alpha_i) m_{i-1, (5-m-l)} + \lambda_i m_{i-1, (4-m-l)} \\ &= (\Omega_i)_{ml}, \end{aligned}$$

garantem que $E[N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})'] = A_2$. Falta apenas mostrar que a variância de cada elemento das matrizes C_{Nij} , para $i = 1, \dots, T$, tende para zero; i.e., $V(C_{Nij})_{ml} \rightarrow 0$ para $m, l = 1, 2$. Para esse fim, note-se que, para $j = 1, \dots, T$ e para $a \leq 4$, $U_{j+nT}^a = \sum_{m=0}^a C_m^a (\alpha_j \circ X_{j-1+nT} - \alpha_j X_{j-1+nT})^m (Z_{j+nT} - \lambda_j)^{a-m}$ cujo valor médio se pode escrever como

$$\begin{aligned}
 & E[U_{j+nT}^a | X_{i-1+nT}, X_{i+nT}, X_{j-1+nT}] = \\
 & \begin{cases} 1 & a = 0 \\ 0 & a = 1 \\ V[\alpha_j \circ X_{j-1+nT} | X_{j-1+nT}] + V[Z_{j+nT}] & a = 2 \\ E[(Z_{j+nT} - \lambda_j)^3] + E[(\alpha_j \circ X_{j-1+nT} - \alpha_j X_{j-1+nT})^3 | X_{i-1+nT}, X_{i+nT}, X_{j-1+nT}] & a = 3 \\ E[(Z_{j+nT} - \lambda_j)^4] + 6V[\alpha_j \circ X_{j-1+nT} | X_{j-1+nT}] V[Z_{j+nT}] + \\ + E[(\alpha_j \circ X_{j-1+nT} - \alpha_j X_{j-1+nT})^4 | X_{j-1+nT}] & a = 4 \end{cases} \\
 & \begin{cases} 1 & a = 0 \\ 0 & a = 1 \\ \alpha_j(1 - \alpha_j)X_{j-1+nT} + \lambda_j & a = 2 \\ \lambda_j + \alpha_j(1 - 3\alpha_j + 2\alpha_j^2 + 1)X_{j-1+nT} & a = 3 \\ 3\lambda_j^2 + \lambda_j + 6\alpha_j(1 - \alpha_j)\lambda_j X_{j-1+nT} + 3X_{j-1+nT}^2(\alpha_j^4 + \alpha_j^2) & a = 4 \\ -X_{j-1+nT}(6\alpha_j^4 + 6\alpha_j^3 + 7\alpha_j^2 - \alpha_j) & \end{cases} \quad (2.20)
 \end{aligned}$$

assumindo que $C_m^a = 0$ para $m > a$.

Para $j > i$, tem-se

$$\begin{aligned}
 & V[(C_{Nij})_{ml}] = \\
 & = V \left[N^{-1} \sum_{n=0}^{N-1} U_{i+nT} X_{i-1+nT}^{2-l} U_{j+nT} X_{j-1+nT}^{2-m} \right] \\
 & = N^{-2} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} E[U_{i+nT} U_{j+nT} U_{i+kT} U_{j+kT} (X_{i-1+nT} X_{i-1+kT})^{2-l} (X_{j-1+kT} X_{j-1+nT})^{2-m}]
 \end{aligned}$$

que simplificada toma a forma

$$\begin{aligned}
V[(C_{Nij})_{ml}] &= \\
&= N^{-2} \sum_{n=0}^{N-1} E[U_{i+nT}^2 U_{j+nT}^2 X_{i-1+nT}^{4-2l} X_{j-1+nT}^{4-2m}]. \\
&= N^{-2} \sum_{n=0}^{N-1} (\alpha_j(1 - \alpha_j) E[U_{i+nT}^2 X_{i-1+nT}^{4-2l} X_{j-1+nT}^{5-2m}] + \lambda_j E[U_{i+nT}^2 X_{i-1+nT}^{4-2l} X_{j-1+nT}^{4-2m}]).
\end{aligned}$$

Após cálculos simples mas morosos (daí a sua omissão), que passam por escrever X_{j-1+nT} à custa de X_{i-1+nT} e conjugar com o desenvolvimento de U_{i+nT}^2 , tem-se que $V[(C_{Nij})_{ml}] = N^{-1} Const \rightarrow 0$, o que permite concluir que $C_{Nij} \xrightarrow{P} \mathbf{0}_2$, para $j > i$, $i = 1, 2, \dots, T$. Considerando agora o caso de $j = i$, vem

$$\begin{aligned}
V[(C_{Nii})_{ml}] &= \\
&= V \left[N^{-1} \sum_{n=0}^{N-1} U_{i+nT}^2 X_{i-1+nT}^{4-m-l} \right] \\
&= N^{-2} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} E[U_{i+nT}^2 U_{i+kT}^2 X_{i-1+nT}^{4-m-l} X_{i-1+kT}^{4-m-l}] - (\alpha_i(1 - \alpha_i)m_{i-1,(5-m-l)} + \lambda_i m_{i-1,(4-m-l)})^2 \\
&\text{(as parcelas que aparecem nos somatórios são exactamente iguais para } n \text{ e para } k) \\
&= 2N^{-2} \sum_{n=0}^{N-1} \sum_{k=0}^{n-1} E[U_{i+kT}^2 X_{i-1+nT}^{4-m-l} X_{i-1+kT}^{4-m-l} E[U_{i+nT}^2 | X_{i-1+nT}, X_{i+kT}, X_{i-1+kT}]] + \\
&\quad + N^{-2} \sum_{n=0}^{N-1} E[U_{i+nT}^4 X_{i-1+nT}^{8-2m-2l}] - (\alpha_i(1 - \alpha_i)m_{i-1,(5-m-l)} + \lambda_i m_{i-1,(4-m-l)})^2 \\
&= 2N^{-2} \sum_{n=0}^{N-1} \sum_{k=0}^{n-1} \alpha_i(1 - \alpha_i) E[U_{i+kT}^2 X_{i-1+nT}^{5-m-l} X_{i-1+kT}^{4-m-l}] + \lambda_i E[U_{i+kT}^2 X_{i-1+nT}^{4-m-l} X_{i-1+kT}^{4-m-l}] \\
&N^{-1} Const_1 - (\alpha_i(1 - \alpha_i)m_{i-1,(5-m-l)} + \lambda_i m_{i-1,(4-m-l)})^2.
\end{aligned}$$

Escrevendo X_{i-1+nT} à custa de X_{i-1+kT} , é possível mostrar, após alguns cálculos morosos, que

$$\begin{aligned}
V[(C_{Nii})_{ml}] &\longrightarrow [I_{\{m=l=2\}}m_{i-1,1} + I_{\{m+l=3\}}m_{i-1,2} + I_{\{m=l=1\}}m_{i-1,3}] \times \\
&\quad \times \{\alpha_i(1 - \alpha_i)m_{i-1,5-m-l} + m_{i-1,4-m-l}\lambda_i\} + \\
&\quad \lambda_i[I_{\{m=l=2\}} + I_{\{m+l=3\}}m_{i-1,1} + I_{\{m=l=1\}}m_{i-1,2}] \times \\
&\quad \times \{\alpha_i(1 - \alpha_i)m_{i-1,5-m-l} + m_{i-1,4-m-l}\lambda_i\} \\
&\quad - (\alpha_i(1 - \alpha_i)m_{i-1,(5-m-l)} + \lambda_i m_{i-1,(4-m-l)})^2 \\
&= 0,
\end{aligned}$$

o que permite concluir que $C_{Nii} \xrightarrow{P} \Omega_i$, para $i = 1, 2, \dots, T$ e portanto fica provado o resultado (2.17), isto é, que $E[N^{-1} \sum_{k=0}^{N-1} \xi_k(\boldsymbol{\theta}) \xi_k(\boldsymbol{\theta})'] \xrightarrow{P} A_2$.

Finalmente, para mostrar o resultado (2.18), comece-se por notar que a média de cada componente do vector $S_N(\boldsymbol{\theta})$ é zero, pois tendo em conta a expressão (2.20) tem-se

$$E[U_{i+kT} X_{i-1+kT}] = E[X_{i-1+kT} E[U_{i+kT} | X_{i-1+kT}]] = 0$$

e $E[U_{i+kT}] = 0$, para $i = 1, \dots, T$. Desta forma pode-se afirmar que

$$E[N^{-1} S_N(\boldsymbol{\theta})] = N^{-1} \sum_{k=0}^{N-1} E \begin{bmatrix} U_{1+kT} X_{kT} \\ U_{1+kT} \\ \vdots \\ U_{T(k+1)} X_{T(k+1)-1} \\ U_{T(k+1)} \end{bmatrix} = \vec{0}.$$

Para terminar a prova basta mostrar que a variância de cada componente tende para zero quando N tende para mais infinito, i.e. $V \left[N^{-1} \sum_{k=0}^{N-1} U_{i+kT} X_{i-1+kT}^m \right] \rightarrow 0$, para $i = 1, 2, \dots, T$ e $m = 0, 1$. Com efeito

$$\begin{aligned}
V \left[N^{-1} \sum_{k=0}^{N-1} U_{i+kT} X_{i-1+kT}^m \right] &= \\
&= N^{-2} E \left[\left(\sum_{k=0}^{N-1} U_{i+kT} X_{i-1+kT}^m \right)^2 \right] \\
&= N^{-2} \sum_{n=0}^{N-1} \sum_{\substack{k=0 \\ k \neq n}}^{N-1} E [U_{i+nT} X_{i-1+nT}^m U_{i+kT} X_{i-1+kT}^m] + N^{-2} \sum_{n=0}^{N-1} E [U_{i+nT}^2 X_{i-1+nT}^{2m}] \\
&= 2N^{-2} \sum_{n=0}^{N-1} \sum_{k=0}^{n-1} E [X_{i-1+nT}^m U_{i+kT} X_{i-1+kT}^m E[U_{i+nT} | X_{i-1+nT}, X_{i-1+kT}, X_{i+kT}]] + \\
&\quad + N^{-2} \sum_{n=0}^{N-1} E [X_{i-1+nT}^{2m} E[U_{i+nT}^2 | X_{i-1+nT}]].
\end{aligned}$$

Pela expressão (2.20) pode-se concluir que

$$\lim_{N \rightarrow +\infty} V \left[N^{-1} \sum_{k=0}^{N-1} U_{i+kT} X_{i-1+kT}^m \right] = \lim_{N \rightarrow +\infty} N^{-1} (\alpha_i(1 - \alpha_i)m_{i-1,2m+1} + \lambda_i m_{i-1,2m}) = 0,$$

o que conclui a prova do resultado (2.18) e consequentemente do Lema 2.6. \blacksquare

O próximo resultado estabelece as propriedades assintóticas dos estimadores dos mínimos quadrados condicionais.

Teorema 2.7. *No modelo PINAR(1)_T o vector de estimadores $\hat{\boldsymbol{\theta}}_{MQC}$ obtido pelo método dos mínimos quadrados condicionais é assintoticamente normal, isto é*

$$N^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_{MQC} - \boldsymbol{\theta}) \xrightarrow{d} N(\vec{0}, A_1^{-1} A_2 A_1^{-1}).$$

Demonstração Considere-se $\mathcal{F}_{k-1} = \sigma(X_{kT}, X_{kT-1}, \dots)$ a σ -álgebra gerada por $\{X_i, i \leq kT\}$. Uma vez que $E[\xi_k(\boldsymbol{\theta}) | \mathcal{F}_{k-1}] = \vec{0}$ conclui-se que $\xi_k(\boldsymbol{\theta})$ é uma martingala relativamente a \mathcal{F}_k com média zero. Por (2.17) o teorema limite central para martingalas (e.g. Hall e Heyde, 1980) aplica-se a $S_k(\boldsymbol{\theta})$, obtendo-se

$$N^{-\frac{1}{2}} S_N(\boldsymbol{\theta}) \xrightarrow{d} N(\vec{0}, A_2), \quad N \rightarrow \infty. \tag{2.21}$$

Considere-se agora a expansão de Taylor de ordem 1

$$S_N(\hat{\boldsymbol{\theta}}_{MQC}) = S_N(\boldsymbol{\theta}) + \left(\frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) (\hat{\boldsymbol{\theta}}_{MQC} - \boldsymbol{\theta}) + R_N, \tag{2.22}$$

sendo o resto $R_N = O_p(1)$ quando $N \rightarrow \infty$. Substituindo $S_N(\hat{\boldsymbol{\theta}}_{MQC}) = 0$ em (2.22) obtém-se

$$N^{1/2}(\hat{\boldsymbol{\theta}}_{MQC} - \boldsymbol{\theta}) = - \left(N^{-1} \frac{\partial S_N(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^{-1} N^{-1/2} S_N(\boldsymbol{\theta}) + o_p(1). \quad (2.23)$$

Usando (2.16) do Lema 2.6, a equação (2.21) e o Teorema de Slutsky na equação (2.23) fica provado o resultado do teorema. \blacksquare

2.2.3 Estimadores dos mínimos quadrados ponderados

Como a variância condicional de $X_t|X_{t-1}$ não é constante ao longo do tempo e pelo facto dos estimadores dos mínimos quadrados pressuporem a homocedasticidade do modelo, os estimadores de mínimos quadrados condicionais ponderados podem ser uma alternativa a ter em conta. O método dos mínimos quadrados condicional ponderado (MQCP) consiste na obtenção do vector $\boldsymbol{\theta}$ que minimiza a soma dos quadrados das diferenças entre cada observação e o seu valor médio condicional, ponderado pelo inverso da sua variância condicional $V[X_{i+nT}|X_{i-1+nT}] = V(\boldsymbol{\theta}, X_{i-1+nT})$, supondo-se que esta é conhecida. Como tal não acontece, a variância condicional tem de ser estimada, usando estimadores consistentes (ver Heyde e Lin, 1992) para os parâmetros α_i e λ_i , $i = 1, 2, \dots, T$. O vector de estimadores MQCP de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}}_{MQCP} = (\hat{\alpha}_{1,MQCP}, \hat{\lambda}_{1,MQCP}, \dots, \hat{\alpha}_{T,MQCP}, \hat{\lambda}_{T,MQCP})$, para o modelo PINAR(1) $_T$ resulta da minimização de

$$Q(\boldsymbol{\theta}) = \sum_{n=0}^{N-1} \sum_{i=1}^T \frac{(X_{i+nT} - E[X_{i+nT}|X_{i-1+nT}])^2}{\hat{V}(\boldsymbol{\theta}, X_{i-1+nT})} = \sum_{n=0}^{N-1} \sum_{i=1}^T \frac{(X_{i+nT} - \alpha_i X_{i-1+nT} - \lambda_i)^2}{\hat{\alpha}_i(1 - \hat{\alpha}_i)X_{i-1+nT} + \hat{\lambda}_i}.$$

A utilização de $\hat{V}(\hat{\boldsymbol{\theta}}, X_{i+kT-1}) = \hat{\alpha}_{i,MQC}(1 - \hat{\alpha}_{i,MQC})X_{i+kT-1} + \hat{\lambda}_{i,MQC}$, para $i = 1, \dots, T$,

como estimador de $V[\theta, X_{i-1+nT}]$, conduz às seguintes expressões

$$\hat{\alpha}_{i,MQCP} = \frac{\left(\sum_{k=0}^{N-1} \frac{1}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right) \left(\sum_{k=0}^{N-1} \frac{X_{i+kT} X_{i+kT-1}}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right)}{\left(\sum_{k=0}^{N-1} \frac{1}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right) \left(\sum_{k=0}^{N-1} \frac{X_{i+kT-1}^2}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right) - \left(\sum_{k=0}^{N-1} \frac{X_{i+kT-1}}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right)^2} - \frac{\left(\sum_{k=0}^{N-1} \frac{X_{i+kT}}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right) \left(\sum_{k=0}^{N-1} \frac{X_{i+kT-1}}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right)}{\left(\sum_{k=0}^{N-1} \frac{1}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right) \left(\sum_{k=0}^{N-1} \frac{X_{i+kT-1}^2}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right) - \left(\sum_{k=0}^{N-1} \frac{X_{i+kT-1}}{\hat{V}(\hat{\theta}, X_{i+kT-1})} \right)^2}$$

e

$$\hat{\lambda}_{i,MQCP} = \frac{\sum_{k=0}^{N-1} X_{i+kT} - \hat{\alpha}_{i,MQCP} \sum_{k=0}^{N-1} X_{i+kT-1}}{\sum_{k=0}^{N-1} \frac{1}{\hat{V}(\hat{\theta}, X_{i+kT-1})}}.$$

2.2.4 Estimadores de máxima verosimilhança condicional

Considerando x_0 fixo, a função de verosimilhança para o modelo PINAR(1) $_T$ é dada por

$$L(\theta) = \prod_{k=0}^{N-1} \prod_{i=1}^T p_i(x_{i+kT} | x_{i-1+kT}),$$

com

$$p_i(x_{i+kT} | x_{i-1+kT}) = e^{-\lambda_i} \sum_{m=0}^{\min(x_{i+kT}, x_{i-1+kT})} C_m^{x_{i-1+kT}} \alpha_i^m (1 - \alpha_i)^{x_{i-1+kT}-m} \frac{\lambda_i^{x_{i+kT}-m}}{(x_{i+kT} - m)!},$$

$i = 1, \dots, T$. O vector de estimadores MVC, $\hat{\theta}_{MVC} = (\hat{\alpha}_{1,MVC}, \hat{\lambda}_{1,MVC}, \dots, \hat{\alpha}_{T,MVC}, \hat{\lambda}_{T,MVC})$, é obtido através da maximização do logaritmo da função de verosimilhança

$$l(\theta) = \sum_{k=0}^{N-1} \sum_{i=1}^T \ln(p_i(x_{i+kT} | x_{i-1+kT})),$$

levando à resolução de T sistemas de equações da forma

$$\begin{cases} \sum_{k=0}^{N-1} (x_{i+kT} - \alpha_i x_{i-1+kT}) &= \lambda_i \sum_{k=0}^{N-1} \frac{p_i(x_{i+kT} - 1 | x_{i-1+kT})}{p_i(x_{i+kT} | x_{i-1+kT})} \\ \sum_{k=0}^{N-1} \frac{p_i(x_{i+kT} - 1 | x_{i-1+kT})}{p_i(x_{i+kT} | x_{i-1+kT})} &= N \end{cases}, \quad (2.24)$$

para $i = 1, \dots, T$. É de realçar que as estimativas para os λ_i 's são obtidas a partir das estimativas dos α_i 's através da seguinte expressão

$$\hat{\lambda}_{i,MVC} = \frac{1}{N} \sum_{k=0}^{N-1} (x_{i+kT} - \hat{\alpha}_{i,MVC} x_{i-1+kT}); \quad (2.25)$$

no entanto não é possível obter uma expressão analítica para os α_i 's e é necessário aplicar métodos numéricos para resolução de equações não lineares. A substituição de λ_i na primeira equação de (2.24) pela expressão obtida em (2.25) permite afirmar que as estimativas de máxima verosimilhança condicional dos α_i 's são os zeros das seguintes funções

$$f(\alpha_i) = \sum_{k=0}^{N-1} \left[(x_{i+kT} - \alpha_i x_{i-1+kT}) - \frac{1}{N} \sum_{n=0}^{N-1} (x_{i+nT} - \alpha_i x_{i-1+nT}) \times \frac{p_i(x_{i+kT} - 1 | x_{i-1+kT})}{p_i(x_{i+kT} | x_{i-1+kT})} \right], \\ i = 1, \dots, T.$$

Como as T funções acima referidas são funções de uma só variável, para as quais se verificou que possuíam um único zero no intervalo $(0,1)$, utilizou-se o método da bissecção por ser um método que converge sempre para o zero da função.

Franke e Seligmann (1993) mostraram que, sob certas condições de regularidade relativas à distribuição das inovações do modelo, que são verificadas pela distribuição de Poisson, as estimativas de máxima verosimilhança dos parâmetros do modelo INAR(1) são assintoticamente normais. O Teorema que se segue é uma adaptação desse Teorema ao modelo PINAR(1) $_T$.

Teorema 2.8. *Os estimadores de máxima verosimilhança condicional para o modelo PINAR(1)_T têm a seguinte distribuição assintótica,*

$$\sqrt{N} \begin{bmatrix} \hat{\alpha}_{1,MVC} - \alpha_1 \\ \hat{\lambda}_{1,MVC} - \lambda_1 \\ \vdots \\ \hat{\alpha}_{T,MVC} - \alpha_T \\ \hat{\lambda}_{T,MVC} - \lambda_T \end{bmatrix} \xrightarrow{d} N(\vec{0}, I^{-1}), \quad (2.26)$$

onde

$$I = \begin{bmatrix} M_1 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \mathbf{0}_2 & M_2 & \mathbf{0}_2 & \dots & \mathbf{0}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \dots & M_T \end{bmatrix}, \quad M_i = \begin{bmatrix} -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i^2}\right] & -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i \partial \lambda_i}\right] \\ -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i \partial \lambda_i}\right] & -E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda_i^2}\right] \end{bmatrix},$$

com

$$\begin{aligned} E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i^2}\right] &= \frac{N}{\alpha_i^2(1-\alpha_i)^2} \sum_{x_{i+T}}^{+\infty} \sum_{x_{i-1+T}}^{+\infty} P(X_{i-1+T} = x_{i-1+T}) \times \\ &\times \left\{ [(2\alpha_i - 1)x_{i+T} - \alpha_i^2 x_{i-1+T}] p_i(x_{i+T}|x_{i-1+T}) + \right. \\ &\left. + 2(1-\alpha_i)\lambda_i p_i(x_{i+T} - 1|x_{i-1+T}) + \lambda_i^2 p_i(x_{i+T} - 2|x_{i-1+T}) - \lambda_i^2 \frac{p_i^2(x_{i+T} - 1|x_{i-1+T})}{p_i(x_{i+T}|x_{i-1+T})} \right\}, \end{aligned}$$

$$\begin{aligned} E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i \partial \lambda_i}\right] &= \frac{N}{\alpha_i(1-\alpha_i)} \sum_{x_{i+T}}^{+\infty} \sum_{x_{i-1+T}}^{+\infty} P(X_{i-1+T} = x_{i-1+T}) \times \\ &\times \left\{ \lambda_i \frac{p_i^2(x_{i+T} - 1|x_{i-1+T})}{p_i(x_{i+T}|x_{i-1+T})} - p_i(x_{i+T} - 1|x_{i-1+T}) - \lambda_i p_i(x_{i+T} - 2|x_{i-1+T}) + \right\} \end{aligned}$$

e

$$E\left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda_i^2}\right] = N \sum_{x_{i+T}}^{+\infty} \sum_{x_{i-1+T}}^{+\infty} P(X_{i-1+T} = x_{i-1+T}) \times \left\{ p_i(x_{i+T} - 2|x_{i-1+T}) - \frac{p_i^2(x_{i+T} - 1|x_{i-1+T})}{p_i(x_{i+T}|x_{i-1+T})} \right\}.$$

Demonstração Pretende-se verificar as condições de regularidade do logaritmo da verossimilhança que é dado por

$$l(\boldsymbol{\theta}) = \sum_{k=0}^{N-1} \sum_{i=1}^T \ln(p_i(x_{i+kT}|x_{i-1+kT})),$$

onde $p_i(x_{i+kT}|x_{i-1+kT})$ é a convolução da distribuição binomial de parâmetros x_{i-1+kT} e α_i com a distribuição de Poisson de média λ_i ; para tal basta verificar que essas condições também são válidas para $p_i(x_{i+kT}|x_{i-1+kT})$, $i = 1, 2, \dots, T$. No modelo PINAR(1) $_T$ as inovações têm distribuição de Poisson e como tal verificam as condições provadas por Frank e Seligmann (1993), que são as seguintes para $i = 1, 2, \dots, T$:

(C1) O conjunto $\{a : P(Z_{i+kT} = a) = f(a, \lambda_i)\}$, não depende de λ_i ;

(C2) $E[Z_{i+kT}^3] < \infty$;

(C3) $P(Z_{i+kT} = k)$ é três vezes continuamente diferenciável relativamente a λ_i no espaço dos parâmetros B ;

(C4) Para qualquer $\lambda_i \in B$ existe uma vizinhança U de λ_i tal que

1. $\sum_{l=0}^{\infty} \sup_{\lambda_i \in U} f(l, \lambda_i) < \infty$,
2. $\sum_{l=0}^{\infty} \sup_{\lambda_i \in U} |f'(l, \lambda_i)| < \infty$,
3. $\sum_{l=0}^{\infty} \sup_{\lambda_i \in U} |f''(l, \lambda_i)| < \infty$;

(C5) Para qualquer $\lambda_i \in B$ existe uma vizinhança U de λ_i e sucessões crescentes $\psi_i(n)$, $\psi_{ii}(n)$ e $\psi_{iii}(n)$, $n \geq 0$ tal que $\forall \lambda_i \in B$ e $\forall m \leq n$ com $f(m, \lambda_i)$ não nulos

$$|f'(m, \lambda_i)| \leq \psi_i(n) f(m, \lambda_i)$$

$$|f''(m, \lambda_i)| \leq \psi_{ii}(n) f(m, \lambda_i)$$

$$|f^{(3)}(m, \lambda_i)| \leq \psi_{iii}(n) f(m, \lambda_i)$$

e relativamente à distribuição estacionária de $\{X_{i+nT}\}$

$$E[\psi_i^3(X_i)] < \infty$$

$$E[X_i \psi_{ij}(X_{i+1})] < \infty$$

$$E[\psi_i(X_i) \psi_{jk}(X_{i+1})] < \infty$$

$$E[\psi_{ijk}(X_i)] < \infty;$$

(C6) A matriz $\Sigma(\theta_i) = \sigma_i$ é não singular,

$$(\sigma_i)_{11} = E \left(\frac{\partial}{\partial \alpha_i} \log p_i(X_i | X_{i-1}) \right)^2$$

$$(\sigma_i)_{21} = (\sigma_i)_{12} = E \left(\frac{\partial}{\partial \alpha_i} \log p_i(X_i | X_{i-1}) \frac{\partial}{\partial \lambda_i} \log p_i(X_i | X_{i-1}) \right)$$

$$(\sigma_i)_{22} = E \left(\frac{\partial}{\partial \lambda_i} \log p_i(X_i | X_{i-1}) \right)^2.$$

Para a distribuição de Poisson os autores identificaram as funções $\psi_i(n)$, $\psi_{ii}(n)$ e $\psi_{iii}(n)$ que tornam válida a condição (C5); neste caso são $\psi_i(n) = \text{const}_i \cdot n$, $\psi_{ii}(n) = \text{const}_i \cdot n^2$ e $\psi_{iii}(n) = \text{const}_i \cdot n^3$, onde const_i designa uma constante apropriada. Mostraram ainda que as condições correspondentes a (C1), ..., (C6) implicam as condições dos resultados de Billingley (1961) sobre os estimadores de parâmetros de processos de Markov para o modelo INAR. Uma vez que, para cada um dos períodos $i = 1, \dots, T$ do modelo PINAR(1) $_T$, $p_i(\cdot | \cdot)$ é a função de transição do modelo INAR com parâmetros α_i e λ_i , fica garantido que também neste caso as condições (C1) a (C6) implicam as condições gerais do Teorema de Billingsley (1961). ■

2.3 Estudo de simulação

Nesta secção pretende-se estudar o desempenho dos vários estimadores dos parâmetros do modelo PINAR(1) $_T$ anteriormente apresentados, para amostras de pequena, média e grande dimensão, e comparar os seus comportamentos. Este estudo tem carácter indicativo, uma

vez que as conclusões aqui extraídas são válidas apenas para os cenários analisados.

Para tal simularam-se 1000 realizações do processo

$$X_t = \phi_t \circ X_{t-1} + Z_t,$$

onde, para $t = i + 4n$ e $i = 1, 2, 3, 4$, tem-se $\phi_t = \alpha_i$ e $Z_t \sim P(\lambda_i)$; tal significa que se assume $T = 4$. A dimensão considerada foi $4N$, onde $N = 20, 50, 100, 500$. Os valores de α e λ considerados na simulação encontram-se descritos na Tabela 2.1. A escolha dos valores dos parâmetros tem em linha de conta modelos cujos elementos do vector α são, todos não inferiores a 0.5 ou todos não superiores a 0.5 ou ainda a situação em que os que são superiores a 0.5 estão intercalados com os elementos que são inferiores a 0.5. Além disso, também se analisam as seguintes situações:

1. a elementos de maior valor de α corresponderem elementos de maior valor de λ ;
2. a maiores valores de α corresponderem elementos de λ de menor valor.

Com o objectivo de fazer comparação dos diversos métodos de estimação apresentados, apenas são admitidas realizações para as quais os valores estimados dos parâmetros se encontram dentro do suporte dos parâmetros, para todos os métodos de estimação, sendo as restantes realizações rejeitadas.

Em termos de implementação computacional, a simulação das realizações e a estimação dos parâmetros são obtidas através das respectivas funções PINAR_1.R e ESTIMACAO.R construídas no *software* R, que se encontram em apêndice.

Para cada caso analisado são avaliadas as seguintes medidas amostrais:

$$\bar{\theta}_j = \frac{1}{1000} \sum_{m=1}^{1000} \theta_j^{(m)} \text{ e } EQM(\theta_j) = \frac{1}{1000} \sum_{m=1}^{1000} (\theta_j^{(m)} - \theta_j)^2,$$

onde θ_j representa a j -ésima componente, $j = 1, \dots, 8$, do vector de parâmetros $\theta = (\alpha_1, \lambda_1, \dots, \alpha_4, \lambda_4)$.

Tabela 2.1: Valores de α e λ usados na simulação

α	(0.85,0.50,0.76,0.63)	(0.6,0.4,0.8,0.25)	(0.4,0.2,0.5,0,1)
λ	(4,1,3,2)	(3,2,4,1)	(4,1,3,2)
	(2,4,1,3)	(2,4,1,3)	(2,3,1,4)
	(6,1,5,3)	(5,3,6,1)	(6,1,5,3)
	(3,6,1,5)	(1,6,3,5)	(3,5,1,6)

As Tabelas 2.2-2.4 e as Figuras 2.3-2.5, em conjunto com as Tabelas e Figuras do apêndice B, resumem os resultados do estudo de simulação. As caixas de bigodes apresentadas nas Figuras 2.3-2.5 estão dispostas segundo a ordem de apresentação dos métodos de estimação na Secção 2.2; a referir, YW, MQC, MQCP, MVC. A partir das referidas tabelas e figuras conclui-se que o método de máxima verosimilhança condicional se destaca dos restantes em termos de viés e erro quadrático médio, sendo mais relevante nas realizações de pequena e média dimensão. Para realizações de grande dimensão os métodos de estimação apresentam resultados muito semelhantes e de acordo com as propriedades assintóticas que possuem, isto é simetria e consistência. Este estudo de simulação também revela que não existe vantagem em usar o método dos mínimos quadrados ponderados relativamente ao método dos mínimos quadrados tradicional, pois as diferenças entre as médias amostrais e os erros quadráticos médios são pouco significativas.

Em qualquer dos métodos de estimação apresentados as estimativas dos parâmetros para cada período são obtidas isoladamente, pelo que as estimativas dos parâmetros associados a cada período apenas se influenciam mutuamente. Verifica-se que em média, as estimativas de α_i são melhores para valores pequenos de λ_i , $i = 1, \dots, 4$, e o recíproco também se aplica.

Tabela 2.2: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.5, 0.76, 0.63)$ e $\lambda = (4, 1, 3, 2)$.

N		α				λ			
20	YW	0.737 (0.039)	0.423 (0.018)	0.706 (0.036)	0.597 (0.028)	4.717 (1.965)	1.771 (1.728)	3.291 (1.356)	2.264 (1.488)
	MQC	0.740 (0.037)	0.423 (0.018)	0.706 (0.036)	0.597 (0.028)	4.697 (1.882)	1.771 (1.728)	3.291 (1.356)	2.255 (1.488)
	MQCP	0.741 (0.037)	0.422 (0.018)	0.706 (0.036)	0.597 (0.027)	4.692 (1.886)	1.786 (1.685)	3.289 (1.356)	2.259 (1.452)
	MVC	0.838 (0.009)	0.493 (0.006)	0.769 (0.013)	0.649 (0.011)	3.995 (0.577)	1.078 (0.434)	2.885 (0.524)	1.846 (0.547)
50	YW	0.810 (0.014)	0.479 (0.006)	0.750 (0.013)	0.626 (0.011)	4.230 (0.641)	1.205 (0.543)	3.061 (0.498)	2.030 (0.627)
	MQC	0.812 (0.013)	0.479 (0.006)	0.750 (0.013)	0.626 (0.011)	4.223 (0.621)	1.205 (0.543)	3.061 (0.498)	2.031 (0.626)
	MQCP	0.812 (0.013)	0.477 (0.006)	0.749 (0.013)	0.625 (0.011)	4.219 (0.616)	1.229 (0.503)	3.062 (0.498)	2.042 (0.611)
	MVC	0.849 (0.004)	0.504 (0.003)	0.767 (0.005)	0.638 (0.004)	3.957 (0.236)	0.961 (0.220)	2.946 (0.214)	1.930 (0.230)
100	YW	0.837 (0.007)	0.492 (0.003)	0.754 (0.008)	0.628 (0.006)	4.080 (0.326)	1.079 (0.297)	3.040 (0.277)	2.018 (0.311)
	MQC	0.837 (0.007)	0.492 (0.003)	0.754 (0.008)	0.628 (0.006)	4.078 (0.323)	1.079 (0.297)	3.040 (0.277)	2.020 (0.310)
	MQCP	0.838 (0.007)	0.492 (0.003)	0.754 (0.008)	0.627 (0.006)	4.077 (0.321)	1.083 (0.276)	3.039 (0.275)	2.023 (0.305)
	MVC	0.854 (0.002)	0.502 (0.002)	0.764 (0.003)	0.635 (0.002)	3.960 (0.131)	0.975 (0.129)	2.975 (0.108)	1.963 (0.125)
500	YW	0.847 (0.002)	0.499 (0.001)	0.759 (0.002)	0.629 (0.001)	4.017 (0.074)	1.014 (0.071)	3.005 (0.056)	2.004 (0.065)
	MQC	0.847 (0.002)	0.499 (0.001)	0.759 (0.002)	0.629 (0.001)	4.016 (0.074)	1.014 (0.071)	3.005 (0.057)	2.005 (0.065)
	MQCP	0.847 (0.002)	0.499 (0.001)	0.760 (0.001)	0.629 (0.001)	4.015 (0.073)	1.017 (0.064)	3.003 (0.055)	2.004 (0.062)
	MVC	0.849 (0.001)	0.500 (0.0003)	0.761 (0.001)	0.631 (0.0004)	4.003 (0.028)	0.998 (0.026)	2.993 (0.024)	1.991 (0.025)

Tabela 2.3: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.6, 0.4, 0.8, 0.25)$ e $\lambda = (5, 3, 6, 1)$.

N		α				λ			
20	YW	0.546 (0.0570)	0.406 (0.032)	0.688 (0.051)	0.192 (0.011)	5.205 (1.022)	2.965 (1.635)	6.606 (2.015)	1.648 (1.184)
	MQC	0.545 (0.0552)	0.406 (0.032)	0.688 (0.051)	0.192 (0.011)	5.209 (1.004)	2.965 (1.635)	6.606 (2.015)	1.651 (1.198)
	MQCP	0.545 (0.0555)	0.406 (0.031)	0.689 (0.051)	0.190 (0.011)	5.209 (1.007)	2.970 (1.614)	6.598 (2.009)	1.677 (1.203)
	MVC	0.587 (0.0446)	0.411 (0.021)	0.748 (0.027)	0.211 (0.009)	5.057 (0.834)	2.932 (1.088)	6.255 (1.231)	1.444 (0.945)
50	YW	0.579 (0.035)	0.391 (0.014)	0.762 (0.020)	0.228 (0.005)	5.051 (0.565)	3.051 (0.737)	6.212 (0.756)	1.244 (0.538)
	MQC	0.578 (0.034)	0.391 (0.014)	0.762 (0.020)	0.228 (0.005)	5.053 (0.560)	3.051 (0.737)	6.212 (0.756)	1.243 (0.537)
	MQCP	0.578 (0.034)	0.389 (0.014)	0.762 (0.020)	0.227 (0.005)	5.053 (0.558)	3.061 (0.711)	6.210 (0.761)	1.250 (0.525)
	MVC	0.593 (0.025)	0.398 (0.010)	0.789 (0.010)	0.237 (0.004)	4.995 (0.416)	2.994 (0.544)	6.056 (0.437)	1.140 (0.389)
100	YW	0.603 (0.016)	0.401 (0.007)	0.796 (0.010)	0.241 (0.003)	4.991 (0.263)	2.989 (0.390)	6.006 (0.398)	1.095 (0.271)
	MQC	0.603 (0.016)	0.401 (0.007)	0.796 (0.010)	0.241 (0.003)	4.990 (0.262)	2.989 (0.390)	6.006 (0.398)	1.096 (0.272)
	MQCP	0.603 (0.016)	0.400 (0.007)	0.796 (0.010)	0.240 (0.002)	4.989 (0.260)	2.995 (0.372)	6.007 (0.400)	1.115 (0.264)
	MVC	0.605 (0.011)	0.401 (0.005)	0.803 (0.005)	0.244 (0.002)	4.982 (0.196)	2.993 (0.266)	5.967 (0.242)	1.062 (0.205)
500	YW	0.598 (0.003)	0.401 (0.002)	0.800 (0.003)	0.25 (0.001)	5.009 (0.052)	2.992 (0.081)	6.000 (0.090)	0.996 (0.062)
	MQC	0.598 (0.003)	0.401 (0.002)	0.800 (0.003)	0.25 (0.001)	5.009 (0.051)	2.992 (0.081)	6.000 (0.090)	0.996 (0.062)
	MQCP	0.598 (0.003)	0.401 (0.002)	0.800 (0.003)	0.25 (0.001)	5.009 (0.052)	2.992 (0.079)	6.000 (0.091)	1.001 (0.059)
	MVC	0.599 (0.002)	0.400 (0.001)	0.801 (0.001)	0.25 (0.000)	5.004 (0.039)	2.993 (0.051)	5.994 (0.046)	0.996 (0.041)

Tabela 2.4: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (2, 3, 1, 4)$.

N		α				λ			
20	YW	0.403 (0.033)	0.288 (0.043)	0.478 (0.025)	0.276 (0.0724)	2.011 (0.628)	2.715 (0.686)	1.114 (0.362)	3.488 (0.767)
	MQC	0.404 (0.033)	0.288 (0.043)	0.478 (0.025)	0.276 (0.072)	2.009 (0.6206)	2.715 (0.686)	1.114 (0.362)	3.492 (0.766)
	MQCP	0.400 (0.031)	0.287 (0.042)	0.475 (0.023)	0.276 (0.072)	2.025 (0.589)	2.720 (0.672)	1.125 (0.338)	3.490 (0.760)
	MVC	0.412 (0.024)	0.293 (0.037)	0.498 (0.015)	0.280 (0.070)	1.974 (0.444)	2.696 (0.614)	1.032 (0.212)	3.482 (0.731)
50	YW	0.409 (0.015)	0.221 (0.016)	0.498 (0.012)	0.177 (0.021)	1.967 (0.292)	2.932 (0.254)	1.011 (0.172)	3.780 (0.231)
	MQC	0.409 (0.015)	0.221 (0.016)	0.498 (0.012)	0.177 (0.021)	1.967 (0.289)	2.932 (0.254)	1.011 (0.172)	3.780 (0.235)
	MQCP	0.407 (0.014)	0.220 (0.015)	0.495 (0.021)	0.177 (0.005)	1.973 (0.276)	2.935 (0.248)	1.021 (0.156)	3.776 (0.236)
	MVC	0.410 (0.011)	0.225 (0.015)	0.500 (0.006)	0.183 (0.022)	1.963 (0.219)	2.915 (0.237)	1.002 (0.093)	3.760 (0.241)
100	YW	0.402 (0.007)	0.207 (0.009)	0.503 (0.006)	0.146 (0.012)	1.997 (0.153)	2.974 (0.143)	0.990 (0.083)	3.867 (0.141)
	MQC	0.402 (0.007)	0.207 (0.009)	0.503 (0.006)	0.146 (0.012)	1.996 (0.153)	2.974 (0.143)	0.990 (0.083)	3.867 (0.142)
	MQCP	0.402 (0.007)	0.206 (0.009)	0.502 (0.005)	0.145 (0.012)	1.997 (0.149)	2.976 (0.142)	0.994 (0.074)	3.867 (0.141)
	MVC	0.403 (0.005)	0.208 (0.008)	0.502 (0.003)	0.147 (0.012)	1.991 (0.108)	2.970 (0.131)	0.994 (0.043)	3.864 (0.139)
500	YW	0.398 (0.002)	0.199 (0.002)	0.502 (0.001)	0.104 (0.003)	2.009 (0.029)	3.000 (0.033)	0.994 (0.018)	3.989 (0.028)
	MQC	0.399 (0.002)	0.199 (0.002)	0.502 (0.001)	0.104 (0.003)	2.009 (0.029)	3.000 (0.033)	0.994 (0.018)	3.989 (0.028)
	MQCP	0.398 (0.001)	0.199 (0.002)	0.502 (0.001)	0.104 (0.003)	2.011 (0.028)	3.000 (0.033)	0.995 (0.016)	3.989 (0.028)
	MVC	0.399 (0.001)	0.199 (0.002)	0.501 (0.001)	0.105 (0.003)	2.005 (0.021)	3.001 (0.031)	0.997 (0.009)	3.988 (0.027)

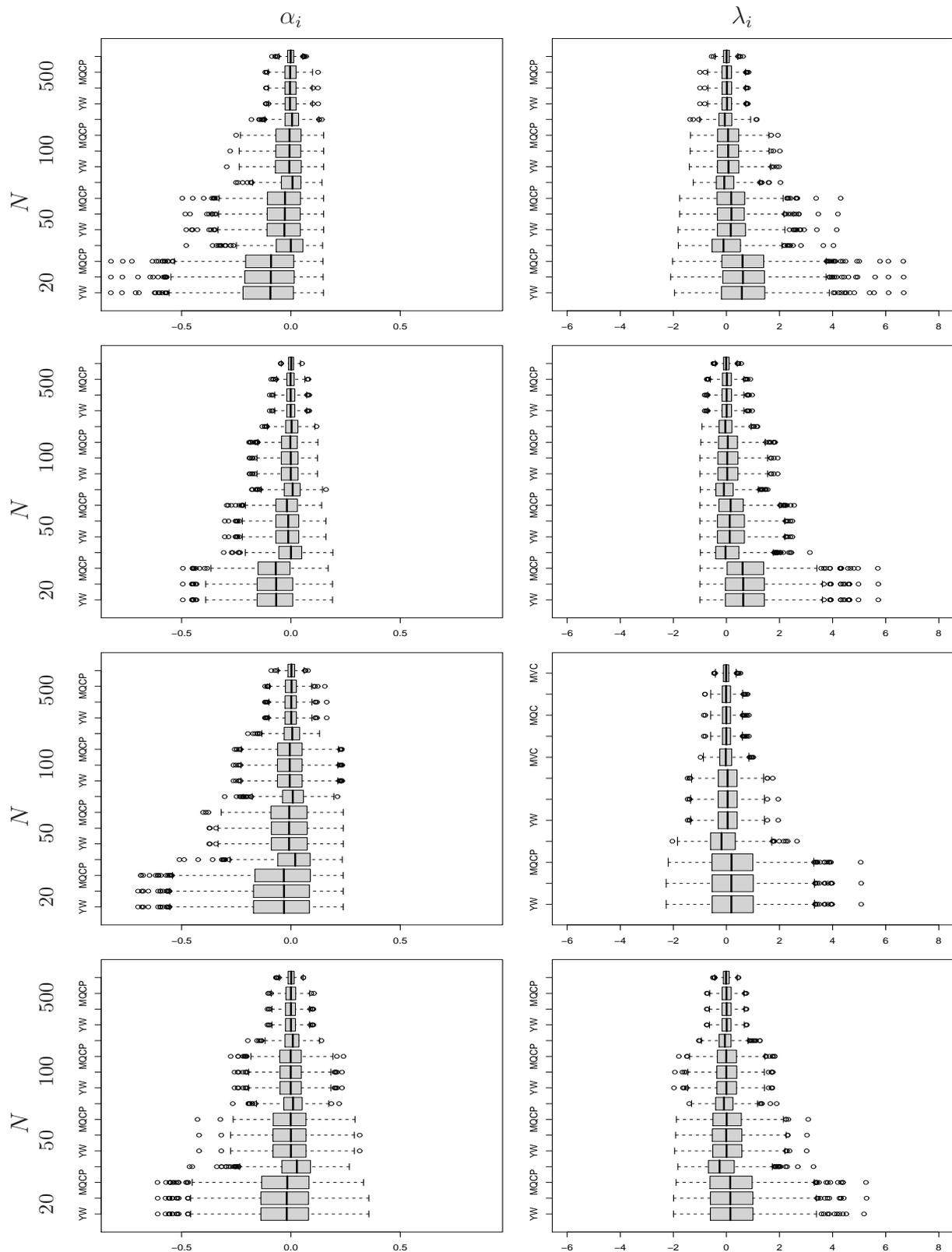


Figura 2.3: Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (4, 1, 3, 2)$.

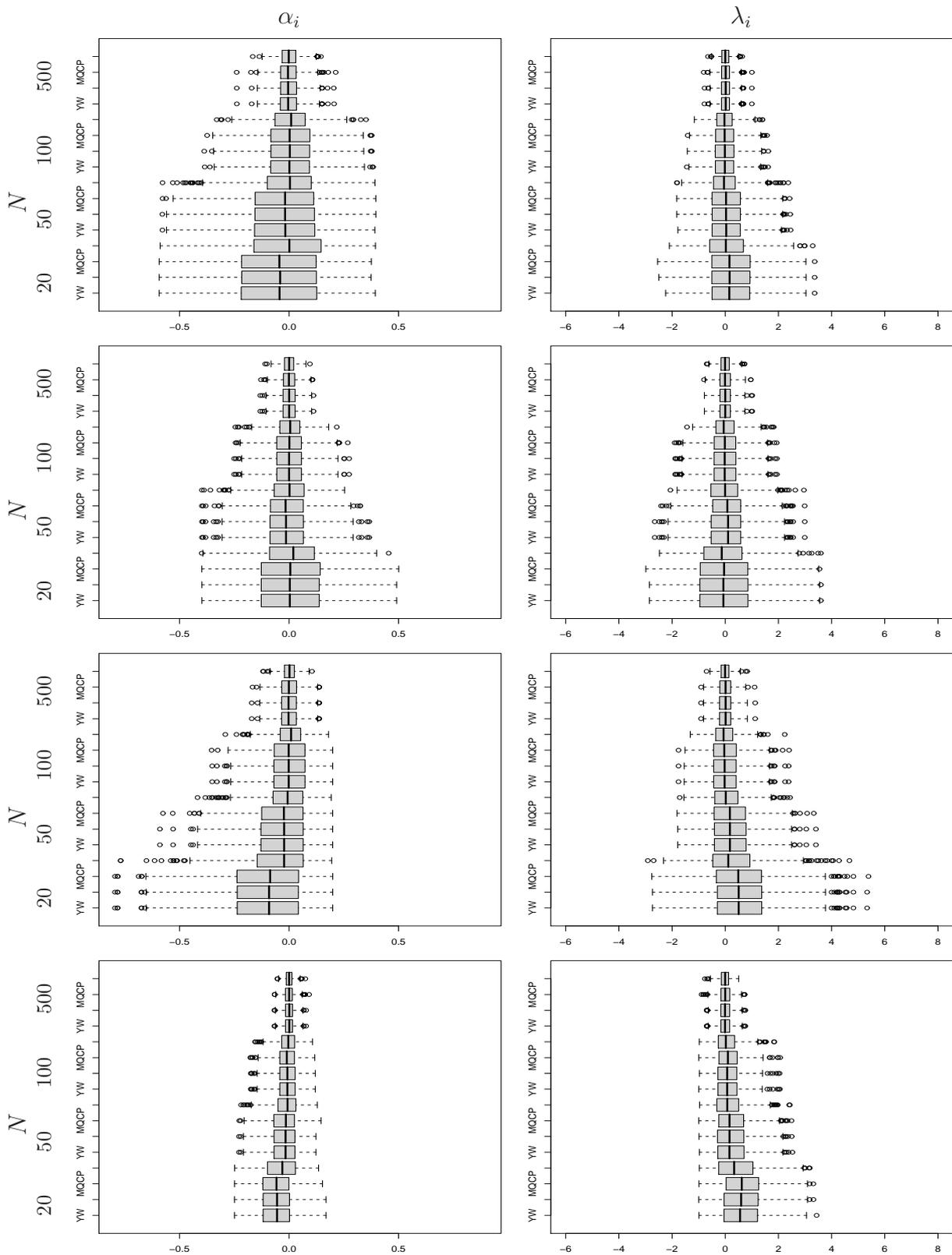


Figura 2.4: Viés das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (5, 3, 6, 1)$.

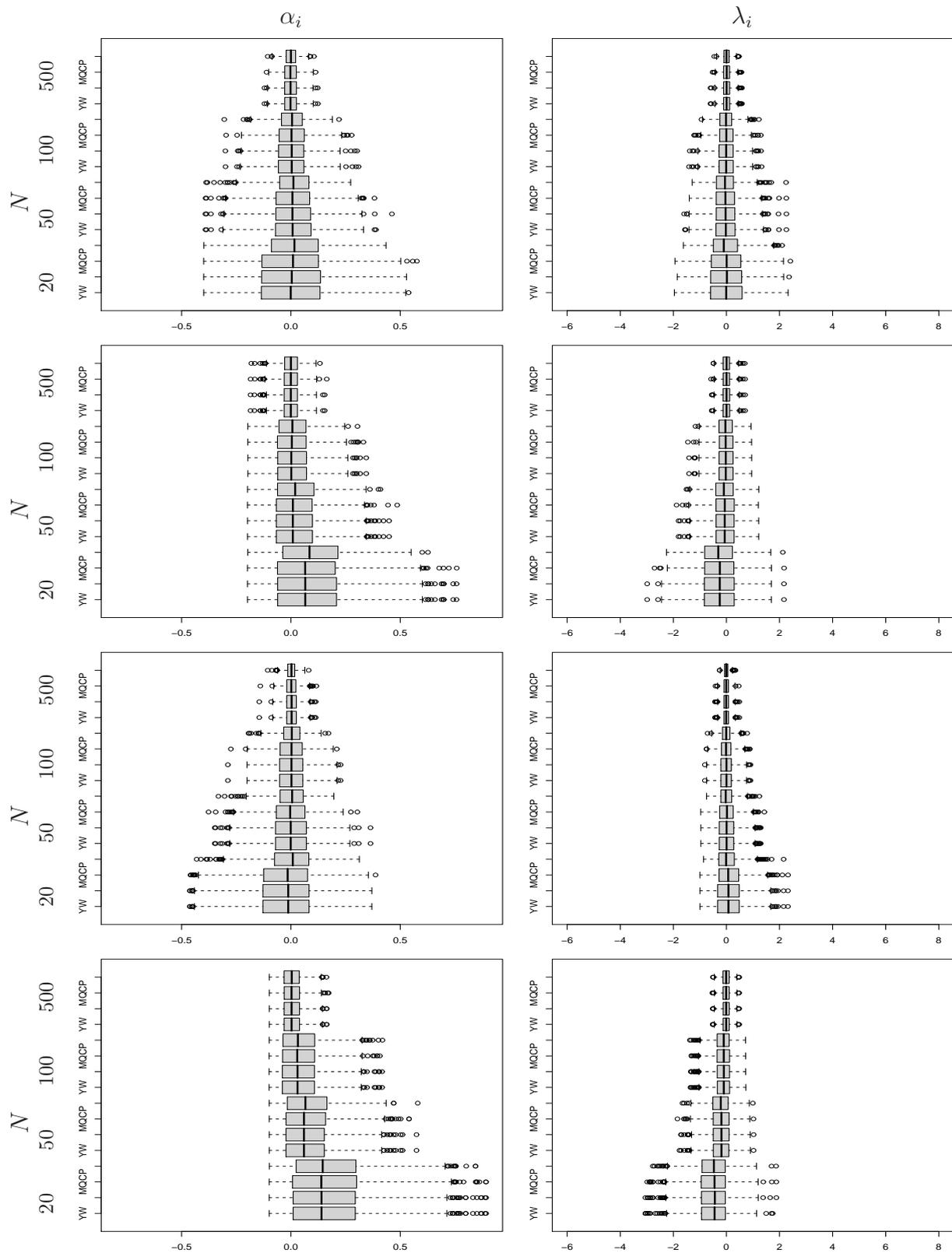


Figura 2.5: Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (2, 3, 1, 4)$.

Tabela 2.5: Percentagens de estimativas dos parâmetros (α, λ) que se encontram dentro do espaço dos parâmetros.

Parâmetros		N = 20			N = 50			N = 100			N = 500			
		YW	MQC	MQP MVC	YW	MQC	MQP MVC	YW	MQC	MQP MVC	YW	MQC	MQP MVC	
α	λ													
	(4,1,3,2)	49	50	47	67	77	75	95	92	92	91	100	100	100
	(2,4,1,3)	65	66	64	98	88	89	87	100	98	98	100	100	100
	(6,1,5,3)	45	45	42	63	69	67	89	86	86	84	97	100	100
	(3,6,1,5)	58	59	56	89	82	83	80	100	93	94	100	100	100
	(3,2,4,1)	53	53	51	59	83	83	82	92	95	95	99	100	100
	(2,4,1,3)	68	68	67	83	94	94	94	97	99	99	100	100	100
	(5,3,6,1)	47	48	45	55	78	76	89	92	92	90	98	100	100
	(1,6,3,5)	55	56	54	73	87	87	86	96	98	98	100	100	100
	(4,1,3,2)	42	42	40	43	75	75	75	77	90	90	90	100	100
	(2,3,1,4)	44	44	44	48	64	64	64	64	78	78	78	97	97
	(6,1,5,3)	39	39	37	41	73	73	72	75	89	89	90	100	100
	(3,5,1,6)	40	40	39	44	61	61	60	63	76	76	76	96	96

Considerem-se os casos em que as realizações são de pequena e média dimensão. Para cada período, as conclusões que se obtêm são semelhantes às referentes ao caso $\text{INAR}(1)$ com inovações de Poisson; isto é, para $i = 1, 2, 3, 4$ quando α_i é inferior a 0.5 as respectivas estimativas são em geral mais dispersas, possuem uma assimetria positiva e existe em média uma sobrestimação desse parâmetro enquanto que as estimativas de λ_i são mais simétricas. Quando α_i é superior a 0.5 as estimativas são em geral menos dispersas, possuem uma assimetria negativa e existe em média subestimação desse parâmetro e as estimativas de λ_i apresentam uma assimetria positiva. Para α_i próximo de 0.5 as suas estimativas são aproximadamente simétricas bem como as estimativas de λ_i .

Outro aspecto importante a analisar é o comportamento dos vários métodos de estimação quanto à taxa de estimativas admissíveis. As percentagens de admissibilidade do vector de parâmetros $\boldsymbol{\theta} = (\alpha_1, \lambda_1, \dots, \alpha_4, \lambda_4)$ encontram-se na Tabela 2.5 para os diferentes métodos de estimação apresentados. Para os casos em que $\boldsymbol{\alpha} = (0.40, 0.20, 0, 50, 0.10)$ as percentagens obtidas são as mais baixas, o que reforça um menor desempenho dos estimadores nestas situações. Quando se consideram realizações de grandes dimensões as percentagens de estimativas dentro do espaço de parâmetros são todas próximas de 100%, para todos os modelos analisados. Assim, para realizações cuja dimensão é desta ordem de grandeza a escolha do método de estimação fica facilitada, pois todos produzem estimativas válidas e próximas dos verdadeiros parâmetros. Na maioria das situações em que as realizações são de pequenas e médias dimensões, o método de máxima verosimilhança condicional destaca-se dos restantes reforçando a importância da sua aplicação nestes casos.

2.4 Predição no modelo $\text{PINAR}_T(1)$

Considere-se uma realização $\underline{X} = (X_1, \dots, X_T, \dots, X_{1+NT}, \dots, X_{i+NT})$, com $i \in \{1, \dots, T\}$, do modelo $\text{PINAR}_T(1)$ definido em (2.1). Pretende-se encontrar um previsor h passos à frente de X_{i+NT+h} . Ao longo desta secção vai-se assumir que $h = j + kT$, para algum

$j \in \{1, \dots, T\}$. Antes de serem definidos os previsores pontuais e intervalares determine-se a função de probabilidade de X_{i+NT+h} , condicional a X_{i+NT} . Atendendo à definição do modelo e aplicando recursivamente a equação (2.1), X_{i+NT+h} toma a forma

$$X_{i+NT+h} \stackrel{d}{=} (\beta_{j+i,j}\beta_{T,T}^k) \circ X_{i+NT} + V_{j+i+kT},$$

onde
$$V_{j+i+kT} = \sum_{m=0}^{j-1} \beta_{j+i,m} \circ Z_{j+i-m+NT} + \sum_{w=0}^{k-1} \sum_{m=0}^{T-1} \beta_{j+i+T(N+k),m+j+Tw} \circ Z_{i+T(N+k-w)-m}.$$

Pelas propriedades da operação *thinning* todas as parcelas de V_{j+i+kT} têm distribuição de Poisson. Como estas parcelas são independentes fica garantido que V_{j+i+kT} tem também distribuição de Poisson de parâmetro λ^* que será especificado posteriormente. Como X_{i+NT} é independente das inovações futuras e das respectivas operações de filtragem pode-se concluir que condicional a X_{i+NT} , X_{i+NT+h} é a convolução entre uma distribuição binomial de parâmetros $(X_{i+NT}, \beta_{j+i,j}\beta_{T,T}^k)$ e uma distribuição de Poisson de parâmetro λ^* permitindo escrever a função de probabilidade de X_{i+NT+h} condicional a X_{i+NT} na forma

$$P(X_{i+NT+h} = y | X_{i+NT} = x) = \sum_{m=0}^{\min(y,x)} C_m^x (\beta_{i+j,j}\beta_{T,T}^k)^m (1 - \beta_{i+j,j}\beta_{T,T}^k)^{x-m} e^{-\lambda^*} \frac{(\lambda^*)^{y-m}}{(y-m)!}.$$

Proposição 2.9. *No modelo PINAR(1)_T $X_{i+TN+h} | X_{i+TN}$, $h = j + Tk$ para algum $j \in \{1, \dots, T\}$ e $k \in \mathbb{N}_0$, tem as seguintes propriedades*

(a) $E[X_{i+TN+h} | X_{i+TN}] = \beta_{i+j,j}\beta_{T,T}^k X_{i+TN} + \mu_{\langle j+i \rangle} - \beta_{i+j,j}\beta_{T,T}^k \mu_i;$

(b) $V[X_{i+TN+h} | X_{i+TN}] = \beta_{i+j,j}\beta_{T,T}^k (1 - \beta_{i+j,j}\beta_{T,T}^k) X_{i+TN} + \mu_{\langle j+i \rangle} - \beta_{i+j,j}\beta_{T,T}^k \mu_i;$

(c) *quando $k \rightarrow +\infty$, $X_{i+TN+h} | X_{i+TN}$ tem distribuição de Poisson de parâmetro $\mu_{\langle j+i \rangle}$,*

$$\text{onde } \langle j+i \rangle = \begin{cases} j+i & j+i \leq T \\ j+i-T & j+i > T \end{cases}.$$

Demonstração Para mostrar (a) e (b) basta provar que a média e a variância de V_{j+i+kT} são iguais a $\lambda^* = \mu_{<j+i>} - \beta_{i+j,j}\beta_{T,T}^k\mu_i$, uma vez que $E[(\beta_{j+i,j}\beta_{T,T}^k) \circ X_{i+NT}|X_{i+NT}] = \beta_{j+i,j}\beta_{T,T}^k X_{i+NT}$ e $V[(\beta_{j+i,j}\beta_{T,T}^k) \circ X_{i+NT}|X_{i+NT}] = \beta_{j+i,j}\beta_{T,T}^k(1 - \beta_{j+i,j}\beta_{T,T}^k)X_{i+NT}$.

$$\begin{aligned}
E[V_{i+j+kT}] &= \\
&= \sum_{w=0}^{k-1} \sum_{m=0}^{T-1} \beta_{j+i+T(N+k),m+j+Tw} \nu_{i+T(N+k-w)-m} + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \sum_{w=0}^{k-1} \sum_{m=0}^{T-1} \beta_{j+i,j} \beta_{i+T(N+k),m} \beta_{T,T}^w \nu_{i+T-m} + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j} \frac{1 - \beta_{T,T}^k}{1 - \beta_{T,T}} \sum_{m=0}^{T-1} \beta_{i+T,m} \nu_{i+T-m} + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j} (1 - \beta_{T,T}^k) \frac{1}{1 - \beta_{T,T}} \left(\sum_{m=0}^{i-1} \beta_{i+T,m} \nu_{i+T-m} + \sum_{m=0}^{T-i-1} \beta_{i+T,m+i} \nu_{T-m} \right) + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j} (1 - \beta_{T,T}^k) \underbrace{\frac{1}{1 - \beta_{T,T}} \left(\sum_{m=0}^{i-1} \beta_{i,m} \lambda_{i-m} + \beta_{i,i} \sum_{m=0}^{T-i-1} \beta_{T,m} \lambda_{T-m} \right)}_{\mu_i} + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j} (1 - \beta_{T,T}^k) \mu_i + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j} \mu_i - \beta_{j+i,j} \beta_{T,T}^k \mu_i + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m}.
\end{aligned}$$

Sabe-se que $\mu_i = \beta_{i,i} \mu_T + \sum_{m=0}^{i-1} \beta_{i,m} \lambda_{i-m}$ e mostra-se que $\sum_{m=j}^{j+i-1} \beta_{j+i,m} \nu_{j+i-m} = \beta_{j+i,j} \sum_{m=0}^{i-1} \beta_{i,m} \lambda_{i-m}$,

portanto obtém-se

$$\begin{aligned}
E[V_{i+j+kT}] &= \beta_{j+i,j} \left(\beta_{i,i} \mu_T + \sum_{m=0}^{i-1} \beta_{i,m} \lambda_{i-m} \right) - \beta_{j+i,j} \beta_{T,T}^k \mu_i + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j+i} \mu_T + \beta_{j+i,j} \sum_{m=0}^{i-1} \beta_{i,m} \lambda_{i-m} - \beta_{j+i,j} \beta_{T,T}^k \mu_i + \sum_{m=0}^{j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j+i} \mu_T - \beta_{j+i,j} \beta_{T,T}^k \mu_i + \sum_{m=0}^{i+j-1} \beta_{j+i,m} \nu_{j+i-m} \\
&= \beta_{j+i,j+i} \mu_T + \sum_{m=0}^{i+j-1} \beta_{j+i,m} \nu_{j+i-m} - \beta_{j+i,j} \beta_{T,T}^k \mu_i \\
&= \mu_{<j+i>} - \beta_{T,T}^k \beta_{j+i,j} \mu_i.
\end{aligned}$$

Este último resultado decorre da análise dos casos $j + i > T$ e $j + i \leq T$.

Para mostrar o resultado em (c), note-se que a função geradora de probabilidades de $X_{i+TN+h}|X_{i+TN}$ é dada por

$$P_{X_{i+TN+h}|X_{i+TN}}(r) = (1 - \beta_{i+j,j}\beta_{T,T}^k + \beta_{i+j,j}\beta_{T,T}^k r)^{X_{i+NT}} e^{-(\mu_{<j+i>} - \beta_{T,T}^k \beta_{j+i,j}\mu_i)(1-r)}.$$

A função geradora de probabilidades que se obtém fazendo k tender para $+\infty$ é a de uma distribuição de Poisson de parâmetro $\mu_{<j+i>}$. ■

2.4.1 Predição pontual

Para previsores a $h = j + kT$ ($j \in \{1, \dots, T\}$) passos vão ser considerados a média, a mediana e a moda da distribuição de X_{i+NT+h} condicional a X_{i+NT} , os quais são denotados por $\hat{X}_{i+NT+h} = E[X_{i+NT+h}|X_{i+NT}]$, \hat{m}_{i+NT+h} e $\hat{m}o_{i+NT+h}$ respectivamente. O primeiro é o previsor que minimiza o erro quadrático médio, já o segundo é aquele que minimiza o erro absoluto médio e o último maximiza a própria função de probabilidade. É desejável que, no contexto de séries de contagens, as previsões pontuais sejam valores inteiros, i.e., sejam coerentes - consultar por exemplo Freeland e McCabe (2004). Assim, \hat{X}_{i+NT+h} é considerado um previsor não coerente enquanto que os restantes previsores são coerentes com o modelo.

2.4.2 Predição intervalar

A obtenção de intervalos de confiança para um valor predito baseia-se na função de probabilidade do erro de previsão h passos à frente condicional à realização do processo já observado, i.e., é feita com base em

$$e_{i+NT+h}|\underline{X} = X_{i+NT+h}|\underline{X} - \hat{X}_{i+NT+h}|\underline{X},$$

onde $h = j + kT$, $X_{i+NT} = x$ e \hat{X}_{i+NT+h} é a média da distribuição de X_{i+NT+h} condicional a X_{i+NT} . Observe-se que e_{i+NT+h} é uma v.a. discreta que assume valores no conjunto

$D_{i+NT+h} = \{e \in \mathbb{R} : e = y - \beta_{i+j,j}\beta_{T,T}^k x - \mu_{<j+i>} + \beta_{i+j,j}\beta_{T,T}^k \mu_i, y \in \mathbb{N}_0\}$ e conseqüentemente tem a seguinte função de probabilidade

$$P(e_{i+NT+h} = e | X_{i+NT} = x) = P(X_{i+NT+h} = y | X_{i+NT} = x) = \sum_{m=0}^{\min(y,x)} C_m^x (\beta_{i+j,j}\beta_{T,T}^k)^m (1 - \beta_{i+j,j}\beta_{T,T}^k)^{x-m} e^{-\lambda^*} \frac{(\lambda^*)^{y-m}}{(y-m)!}.$$

A partir da função de probabilidade do erro de previsão o intervalo a $\gamma \times 100\%$ de confiança para X_{i+NT+h} é dado por

$$(\hat{X}_{i+NT+h} + e_1, \hat{X}_{i+NT+h} + e_2), \quad (2.27)$$

onde e_1 é o maior valor de e_{i+NT+h} que verifica $P(e_{i+NT+h} \leq e_1) \leq (1 - \gamma)/2$ e e_2 é o menor valor de e_{i+NT+h} para o qual se tem $P(e_{i+NT+h} \leq e_2) \geq (1 + \gamma)/2$.

2.4.3 Estudo de simulação

Para analisar e comparar os diferentes previsores estudados anteriormente, utilizaram-se os seguintes modelos do estudo de simulação apresentado na Secção 2.3: (M1) com parâmetros $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (4, 1, 3, 2)$, (M2) cujos parâmetros são $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (5, 3, 6, 1)$ e por último (M3) $\alpha = (0.40, 0.20, 0.50, 0.10)$ e $\lambda = (2, 3, 1, 4)$. É importante não esquecer que, para cada um dos três modelos, foram simuladas realizações do modelo de dimensões $n = 4N = 80, 400, 2000$ onde para cada dimensão, se simularam 1000 réplicas independentes. Com as amostras resultantes efectuaram-se previsões pontuais e intervalares h -passos à frente para $h = j + 4k$ com $j \in \{1, 2, 3, 4\}$ e $k \in \{0, \dots, 5\}$, ou seja fizeram-se previsões com h a variar de 1 a 24. Para além da previsão a h passos também se efectuaram previsões passo a passo. Nos diferentes preditores utilizados consideraram-se as estimativas de máxima verosimilhança, porque foram estas que revelaram melhores resultados na Secção 2.3.

Predição pontual

A Tabela 2.6 mostra as predições pontuais para uma realização particular do modelo M1 com $n = 400$, onde h varia de 1 a 40. A Tabela 2.6 encontra-se dividida nos períodos correspondentes à periodicidade do modelo e pode-se verificar que à medida que o número de ciclos, k , aumenta as estimativas pontuais estão a convergir para a moda, mediana e média da respectiva distribuição de Poisson estimada. Os mesmos valores encontram-se representados na Figura 2.6.

Para comparar previsores entre modelos e entre dimensões utilizaram-se o erro quadrático médio (EQM) para a média condicional, o erro absoluto médio (EAM) para a mediana e a função perda tudo ou nada (FPTN) para a moda. Esta última é definida por

$$FPTN = \frac{1}{m} \sum_{k=1}^m I(x_{t+h}^{(k)}),$$

onde m representa o número de réplicas e

$$I(x_{t+h}^{(k)}) = \begin{cases} 1 & \text{se } |\hat{m}o_{t+h}^k - x_{t+h}^k| > 1 \\ 0 & \text{se } |\hat{m}o_{t+h}^k - x_{t+h}^k| \leq 1 \end{cases}.$$

De modo a comparar os previsores dados por \hat{X}_{t+h} , \hat{m}_{t+h} e $\hat{m}o_{t+h}$ utilizou-se a medida erro absoluto percentual médio (EAPM) definido por

$$EAPM = \frac{1}{H} \sum_{h=1}^H |X_{t+h}^* - X_{t+h}|/X_t,$$

onde H representa o número de previsões realizadas e X_{t+h}^* representa um dos previsores utilizados, de acordo com a metodologia. Na Tabela 2.7 apresentam-se as medidas EQM para a média condicional, EAM para a mediana e a FPTN para a moda a partir de 24 previsões h passos à frente, para todos os modelos e todas as dimensões consideradas. Na Tabela 2.8 apresentam-se as mesmas medidas para o mesmo número de previsões mas agora apenas a um passo à frente.

Tabela 2.6: Predições pontuais para uma amostra, de dimensão $n = 400$ com $x_{400} = 15$, do modelo M1 para $h = j + 4k$, $j = \{1, 2, 3, 4\}$ $k = \{0, \dots, 9\}$.

k	j							
	1	2	3	4				
0	x_{n+h} 17	\hat{x}_{n+h} 17.41	x_{n+h} 6	\hat{x}_{n+h} 9.69	x_{n+h} 8	\hat{x}_{n+h} 9.81	x_{n+h} 8	\hat{x}_{n+h} 8.59
1	$\hat{\nu}_{n+h}$ 17	$\hat{\nu}_{n+h}$ 11	$\hat{\nu}_{n+h}$ 6	$\hat{\nu}_{n+h}$ 7	$\hat{\nu}_{n+h}$ 6.72	$\hat{\nu}_{n+h}$ 8	$\hat{\nu}_{n+h}$ 7.75	$\hat{\nu}_{n+h}$ 7
2	$\hat{\omega}_{n+h}$ 10	$\hat{\omega}_{n+h}$ 10	$\hat{\omega}_{n+h}$ 5	$\hat{\omega}_{n+h}$ 6	$\hat{\omega}_{n+h}$ 6.05	$\hat{\omega}_{n+h}$ 6	$\hat{\omega}_{n+h}$ 7.28	$\hat{\omega}_{n+h}$ 6
3	$\hat{\rho}_{n+h}$ 9	$\hat{\rho}_{n+h}$ 10	$\hat{\rho}_{n+h}$ 2	$\hat{\rho}_{n+h}$ 6	$\hat{\rho}_{n+h}$ 5.89	$\hat{\rho}_{n+h}$ 4	$\hat{\rho}_{n+h}$ 7.17	$\hat{\rho}_{n+h}$ 6
4	$\hat{\sigma}_{n+h}$ 9	$\hat{\sigma}_{n+h}$ 10	$\hat{\sigma}_{n+h}$ 7	$\hat{\sigma}_{n+h}$ 6	$\hat{\sigma}_{n+h}$ 5.86	$\hat{\sigma}_{n+h}$ 8	$\hat{\sigma}_{n+h}$ 7.15	$\hat{\sigma}_{n+h}$ 6
5	$\hat{\tau}_{n+h}$ 9	$\hat{\tau}_{n+h}$ 10	$\hat{\tau}_{n+h}$ 5	$\hat{\tau}_{n+h}$ 6	$\hat{\tau}_{n+h}$ 5.85	$\hat{\tau}_{n+h}$ 7	$\hat{\tau}_{n+h}$ 7.14	$\hat{\tau}_{n+h}$ 6
6	$\hat{\eta}_{n+h}$ 9	$\hat{\eta}_{n+h}$ 10	$\hat{\eta}_{n+h}$ 6	$\hat{\eta}_{n+h}$ 6	$\hat{\eta}_{n+h}$ 5.85	$\hat{\eta}_{n+h}$ 6	$\hat{\eta}_{n+h}$ 7.14	$\hat{\eta}_{n+h}$ 6
7	$\hat{\theta}_{n+h}$ 9	$\hat{\theta}_{n+h}$ 10	$\hat{\theta}_{n+h}$ 1	$\hat{\theta}_{n+h}$ 6	$\hat{\theta}_{n+h}$ 5.85	$\hat{\theta}_{n+h}$ 3	$\hat{\theta}_{n+h}$ 7.14	$\hat{\theta}_{n+h}$ 6
8	$\hat{\phi}_{n+h}$ 9	$\hat{\phi}_{n+h}$ 10	$\hat{\phi}_{n+h}$ 3	$\hat{\phi}_{n+h}$ 6	$\hat{\phi}_{n+h}$ 5.85	$\hat{\phi}_{n+h}$ 4	$\hat{\phi}_{n+h}$ 7.14	$\hat{\phi}_{n+h}$ 6
9	$\hat{\psi}_{n+h}$ 9	$\hat{\psi}_{n+h}$ 10	$\hat{\psi}_{n+h}$ 5	$\hat{\psi}_{n+h}$ 6	$\hat{\psi}_{n+h}$ 5.85	$\hat{\psi}_{n+h}$ 7	$\hat{\psi}_{n+h}$ 7.14	$\hat{\psi}_{n+h}$ 6

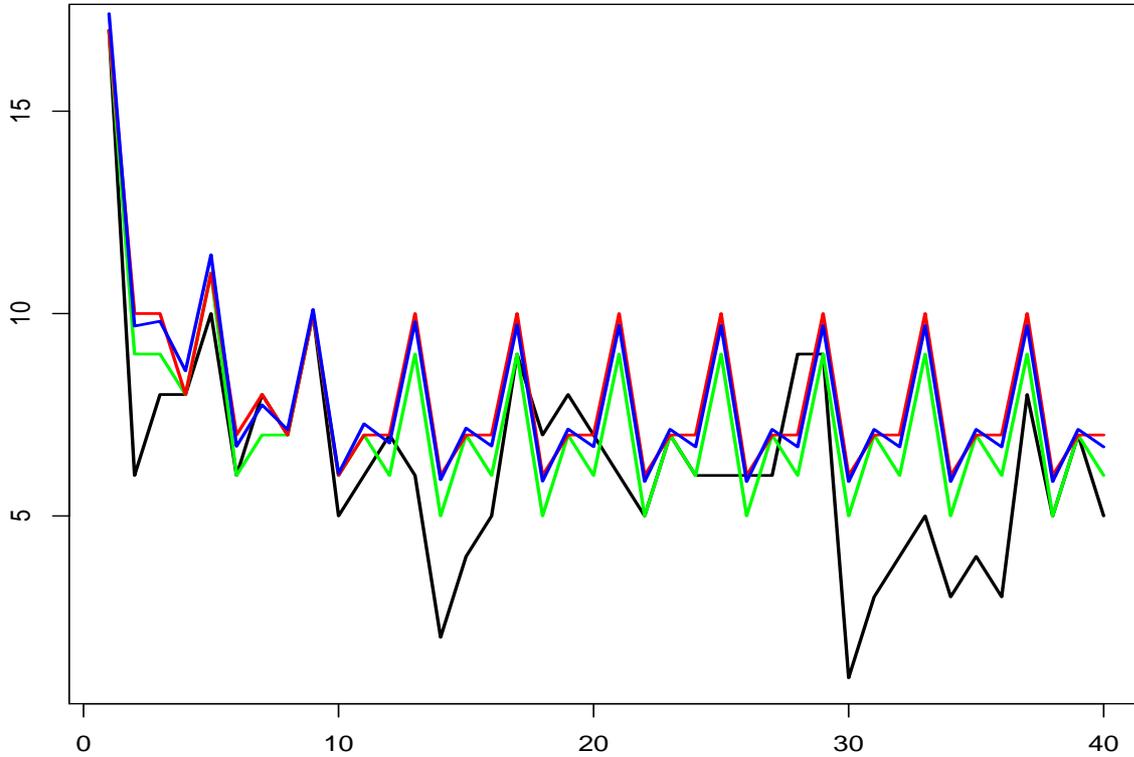


Figura 2.6: Valores de x_{t+h} (preto), \hat{x}_{t+h} (azul), \hat{m}_{t+h} (vermelho) e $\hat{m}_{o_{t+h}}$ (verde) de uma realização de M1.

Na Tabela 2.7 pode ver-se que, na maioria dos casos, um aumento da dimensão da amostra traduz-se num ligeiro decréscimo nas medidas EQM, EAM e FPTN. É de notar que os menores valores das medidas consideradas foram obtidos com o modelo M3, que corresponde ao modelo que tem todos os elementos de α não superiores a 0.5 enquanto que os valores mais elevados das medidas consideradas foram atingidos com o modelo M1 que corresponde a ter todos os elementos de α não inferiores a 0.5. Olhando para o erro absoluto percentual médio, pode-se concluir que, para os modelos analisados, a moda da distribuição condicional produz melhores resultados do que os restantes previsores, verificando-se esta conclusão para todos os modelos e todas as dimensões consideradas. Tal como seria de esperar, nas previsões um passo à frente (ver Tabela 2.8) as medidas em consideração são mais pequenas. Também nesta situação, a moda é o melhor dos previsores em análise.

Tabela 2.7: EQM, EAM, FPTN e EAPM de 24 predições h passos à frente para M1, M2 e M3.

		\hat{X}_{t+h}		\hat{m}_{t+h}		$\hat{m}o_{t+h}$	
		EQM	EAPM	EAM	EAPM	FPTN	EAPM
M1	$N = 20$	7.774	0.405	2.187	0.394	0.560	0.379
	$N = 50$	7.253	0.397	2.117	0.385	0.546	0.368
	$N = 100$	7.067	0.394	2.101	0.383	0.545	0.366
	$N = 500$	7.125	0.411	2.107	0.400	0.541	0.374
M2	$N = 20$	7.224	0.440	2.083	0.425	0.529	0.404
	$N = 50$	6.979	0.447	2.058	0.432	0.526	0.405
	$N = 100$	6.876	0.433	2.042	0.418	0.526	0.391
	$N = 500$	6.801	0.438	2.034	0.429	0.518	0.390
M3	$N = 20$	3.839	0.576	1.534	0.545	0.416	0.504
	$N = 50$	3.706	0.568	1.514	0.542	0.415	0.492
	$N = 100$	3.657	0.567	1.506	0.546	0.413	0.487
	$N = 500$	3.687	0.570	1.513	0.561	0.406	0.478

Tabela 2.8: EQM, EAM, FPTN e EAPM de 24 predições um passo à frente para M1, M2 e M3.

		\hat{X}_{t+h}		\hat{m}_{t+h}		$\hat{m}o_{t+h}$	
		EQM	EAPM	EAM	EAPM	FPTN	EAPM
M1	$N = 20$	4.358	0.289	1.637	0.283	0.441	0.276
	$N = 50$	4.068	0.283	1.581	0.276	0.428	0.270
	$N = 100$	4.046	0.280	1.578	0.273	0.427	0.266
	$N = 500$	3.957	0.288	1.558	0.280	0.419	0.272
M2	$N = 20$	5.592	0.388	1.848	0.376	0.489	0.361
	$N = 50$	5.309	0.387	1.796	0.374	0.475	0.359
	$N = 100$	5.250	0.376	1.779	0.363	0.472	0.349
	$N = 500$	5.157	0.377	1.766	0.364	0.469	0.350
M3	$N = 20$	3.479	0.531	1.444	0.507	0.382	0.481
	$N = 50$	3.300	0.520	1.408	0.493	0.373	0.467
	$N = 100$	3.644	0.536	1.469	0.510	0.384	0.474
	$N = 500$	3.249	0.519	1.394	0.486	0.366	0.466

Predição intervalar

Para ilustrar as previsões intervalares considerou-se uma amostra do modelo M1 de dimensão 400 e calcularam-se os intervalos de confiança para as 24 observações seguintes h passos à frente que se apresentam na Figura 2.7.

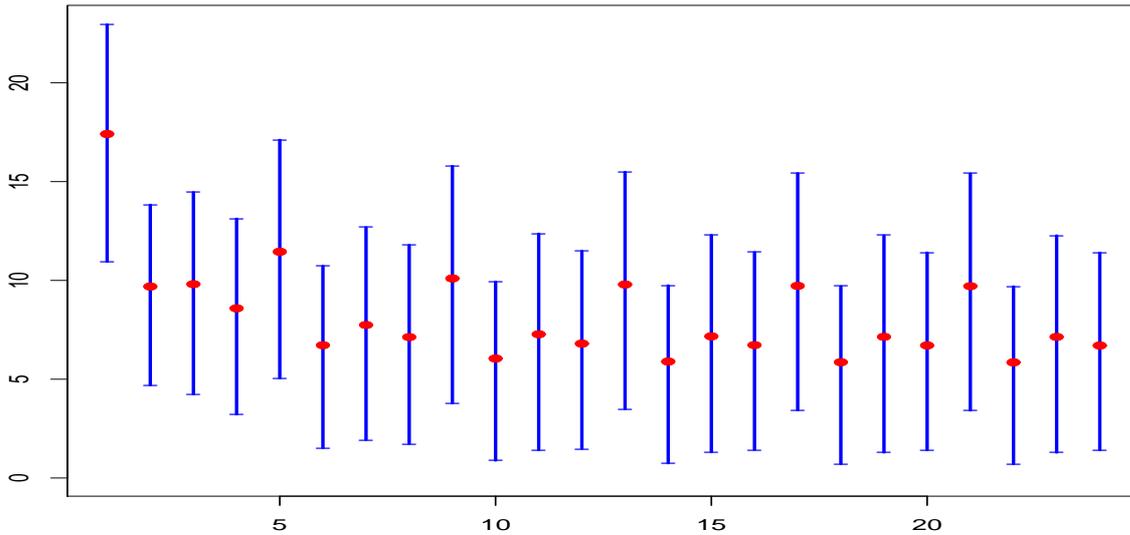


Figura 2.7: Preditores intervalares, a 95% confiança, de x_{t+h} (• representa \hat{x}_{t+h}), $h = 1, \dots, 24$.

Tabela 2.9: Probabilidades de cobertura a partir das previsões intervalares (95% confiança), h passos à frente, $h = 1, \dots, 24$.

N	20	50	100	500
M1	0.948	0.965	0.962	0.973
M2	0.965	0.976	0.977	0.977
M3	0.967	0.962	0.958	0.970

É de notar que a partir de $h = 13$ os intervalos são iguais dentro de cada período. A Tabela 2.9 mostra as probabilidades de cobertura a partir dos previsores intervalares, a 95% de confiança, $h = 1, \dots, 24$ passos à frente enquanto que a Tabela 2.10 contém as

Tabela 2.10: Probabilidades de cobertura a partir das 24 previsões intervalares (95% confiança), 1 passo à frente.

N	20	50	100	500
M1	0.952	0.951	0.958	0.957
M2	0.959	0.968	0.968	0.970
M3	0.962	0.964	0.944	0.972

probabilidades de cobertura a partir dos previsores intervalares um passo à frente para as 24 previsões. Da análise das Tabelas 2.9 e 2.10 pode-se concluir que, tanto para as previsões intervalares h passos à frente como para um passo à frente, as probabilidades de cobertura das previsões intervalares a 95% de confiança estão muito próximas do valor teórico. Esta conclusão é válida para todos os modelos e para todas as dimensões em estudo. É de salientar que na maioria dos casos esse valor é excedido; isto deve-se à definição dos limites dos intervalos de previsão para valores inteiros.

2.5 Exemplo de aplicação

Nesta secção usa-se o conjunto de dados reais correspondentes ao desemprego no Concelho de Penamacor, gentilmente cedidos pelo Instituto do Emprego e Formação Profissional, para ilustrar o modelo introduzido em (2.1). O conjunto de dados analisados é a série mensal do desemprego de curta duração no referido Concelho entre o período de Janeiro de 1997 a Dezembro de 2007 e encontra-se representado na Figura 2.1. A função de autocorrelação estimada da série encontra-se representada na Figura 2.2 (c).

Em termos de interpretação, o modelo PINAR(1)₁₂, parece adequar-se a este tipo de dados, uma vez que o número de desempregados no fim de cada mês é a soma dos desempregados do mês anterior que não arranjam trabalho durante esse mês com os que, durante esse mês,

ficaram desempregados. A probabilidade de um desempregado continuar desempregado no final do mês depende desse mês bem como a média dos novos desempregados depende do mês em que se está a fazer a contabilização.

Note-se que apenas se dispõe de 11 anos completos e como tal a aplicação a estes dados é meramente ilustrativa, uma vez que se tem de estimar 24 parâmetros com pouca informação. Além disso, repare-se que na Figura 2.2 os gráficos das médias mensais e das variâncias mensais não são muito semelhantes. Aliás as variâncias mensais são aproximadamente o dobro das médias mensais, o que poderá ser um indício de que este modelo não seja o mais adequado para modelar esta série. Por outro lado é de salientar também que as médias e variâncias amostrais foram obtidas com apenas 11 observações. Contudo e apesar destas duas limitações, vão-se estimar os parâmetros do modelo.

Tabela 2.11: Estimativas dos parâmetros do modelo PINAR(1)₁₂.

Mês	α	λ
Janeiro	0.823	23.953
Fevereiro	0.633	36.962
Março	0.538	48.821
Abril	0.571	44.812
Mai	0.724	26.406
Junho	0.395	50.487
Julho	0.449	40.433
Agosto	0.680	24.413
Setembro	0.787	19.711
Outubro	0.590	35.981
Novembro	0.828	21.292
Dezembro	0.587	51.483

Como o método MVC produziu melhores resultados para amostras de menores dimensões apenas se usou este método para estimar os parâmetros do modelo. As estimativas dos 24 parâmetros do modelo apresentam-se na Tabela 2.11. Na Figura 2.8 estão representados a série do número de desempregados de curta duração no Concelho de Penamacor e a vermelho os valores estimados pelo modelo.

Para analisar a qualidade do ajustamento do modelo $\text{PINAR}(1)_{12}$ ao conjunto de observações, representaram-se as funções de autocorrelação (FAC) e de autocorrelação parcial (FACP) estimadas para os resíduos do modelo, as quais se encontram na Figura 2.9. Os valores absolutos das estimativas das funções FAC e FACP são todos consideravelmente pequenos, inferiores a 0.2, e assemelham-se às FAC e FACP de uma sucessão de v.a.'s não correlacionadas. A aplicação dos testes não paramétricos usuais de aleatoriedade (teste dos *runs* e o teste da média dos quadrados das diferenças sucessivas) aos resíduos levaram, ao nível de significância de 5%, à não rejeição da hipótese de ausência de correlação entre os resíduos.

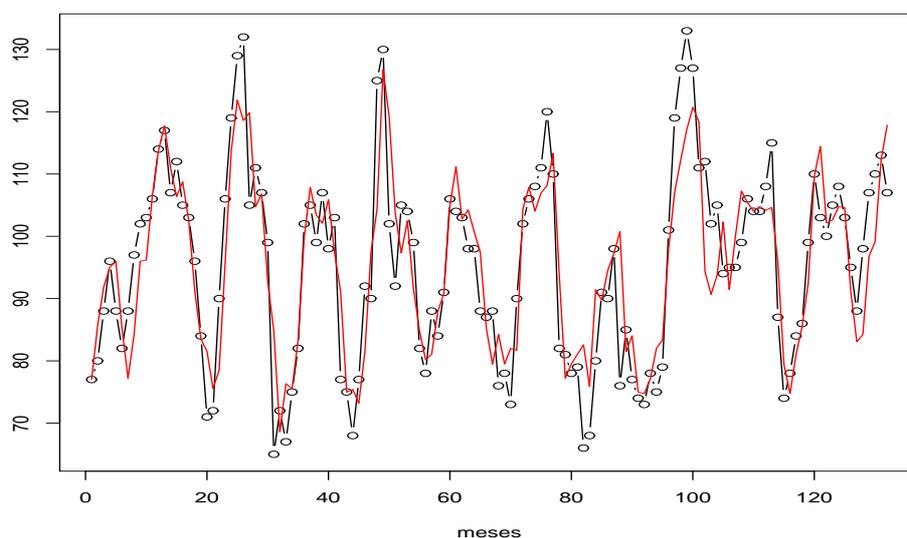


Figura 2.8: Valores estimados pelo modelo $\text{PINAR}(1)_{12}$ ajustado.

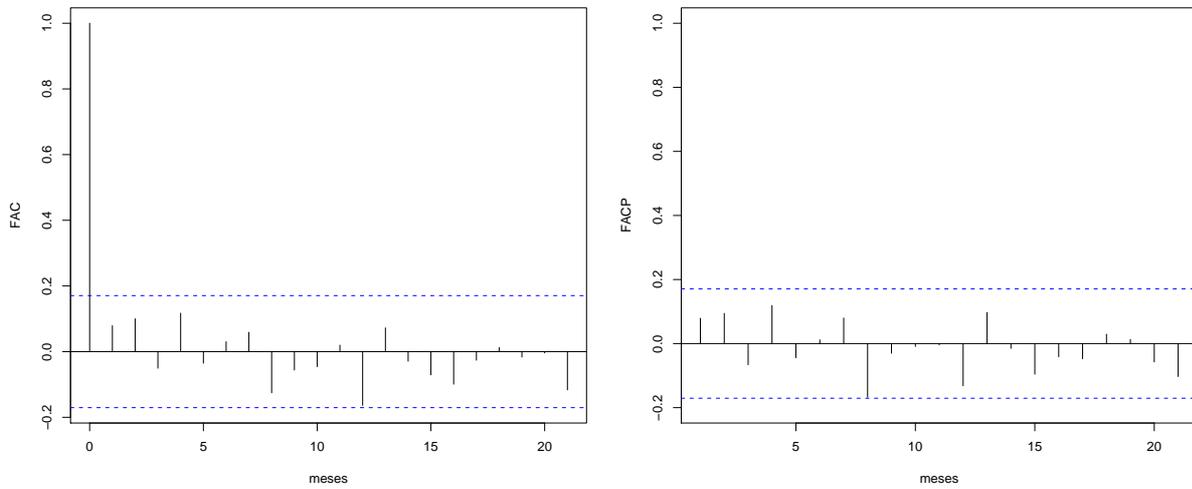


Figura 2.9: FAC e FACP amostrais dos resíduos do modelo $\text{PINAR}(1)_{12}$ ajustado.

2.6 Conclusões

Neste capítulo apresentou-se o modelo periódico auto-regressivo de valores inteiros de ordem 1 e período T , $\text{PINAR}_T(1)$ cujas inovações são uma sucessão periódica de v.a.'s independentes com distribuição de Poisson. Foi estabelecida a ergodicidade do processo e determinada a distribuição ciclo-estacionária do mesmo. Foram apresentados os métodos de estimação de Yule-walker, mínimos quadrados condicionais simples e ponderados e máxima verosimilhança condicional, e estudadas as suas propriedades assintóticas. A título ilustrativo apresentou-se um estudo de simulação onde ficou claro a superioridade do método da máxima verosimilhança, principalmente para amostras de pequena e média dimensão. Apresentaram-se previsores pontuais e intervalares para o modelo em estudo e verificou-se que, para o estudo de simulação efectuado, a moda da distribuição condicional produz em média melhores resultados. Em Monteiro *et al.* (2010) são apresentados os resultados mais relevantes que foram introduzidos neste capítulo.

O modelo estudado é auto-regressivo de primeira ordem, o que significa que dentro da

classe dos modelos periódicos existe muito trabalho ainda por efectuar, que passará pela generalização do modelo $\text{PINAR}_T(1)$ para a ordem p . Este tópico será alvo de futuros trabalhos de investigação. Um dos problemas dos modelos PARMA está relacionado com a parcimónia. Mesmo para o modelo mais simples poder-se-á ter um número muito elevado de parâmetros. Este problema também se verifica para os modelos PINAR. Pelo que, o desenvolvimento de procedimentos que permitam a redução da dimensão de parâmetros é um problema iminente, que será oportunamente alvo de investigação.

Capítulo 3

Modelos auto-regressivos de valores inteiros com limiares auto-induzidos

A classe dos modelos auto-regressivos de valores inteiros estacionários com distribuição de Poisson tem desempenhado um papel fundamental na análise de séries temporais estacionárias de valores inteiros. No entanto esta classe de modelos não é suficientemente ampla que permita captar características que surgem muitas vezes em séries de contagem, tais como saltos bruscos, mudanças da volatilidade no tempo e padrões do tipo linear que se repetem por troços. O objectivo deste capítulo é estudar modelos que permitam caracterizar séries de contagem que exibam esta última característica.

Neste capítulo apresenta-se a classe de modelos auto-regressivos de valores inteiros com limiares auto-induzidos, SETINAR (do inglês *Self-Exciting Threshold INteger AutoRegressive*), cujas inovações são independentes e identicamente distribuídas com distribuição de Poisson. Apenas será estudado o modelo com dois regimes, e para ambos os regimes a componente auto-regressiva é de ordem um e as inovações são v.a.'s independentes com a mesma distribuição de Poisson, que se designará de SETINAR(2;1). Na Secção 3.1 é definido o modelo e são apresentadas as suas propriedades probabilísticas. Na Secção 3.2

é estudado o problema da estimação dos parâmetros, sendo analisados os métodos dos mínimos quadrados condicionais e o de máxima verosimilhança condicional. A comparação destes métodos quanto ao seu desempenho para pequenas, médias e grandes dimensões é feito através de um estudo de simulação apresentado na Secção 3.3.

3.1 O modelo SETINAR(2;1)

O modelo SETINAR(2;1) é definido pela equação recursiva

$$X_t = \phi_t \circ X_{t-1} + Z_t, \quad t \in \mathbb{Z}, \quad (3.1)$$

com $\phi_t = \alpha_1 I_{t-1,1} + \alpha_2 I_{t-1,2}$, onde

$$I_{t-1,1} = \begin{cases} 1 & X_{t-1} \leq R \\ 0 & X_{t-1} > R \end{cases}$$

e $I_{t-1,2} = 1 - I_{t-1,1}$. O operador *thinning* \circ é definido por

$$\phi_t \circ X_{t-1} \stackrel{d}{=} I_{t-1,1} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_1) + I_{t-1,2} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_2),$$

com $(U_{i,t}(\alpha_1))$ e $(U_{i,t}(\alpha_2))$, $i = 1, 2, \dots$, sucessões de variáveis aleatórias independentes com distribuição de Bernoulli com probabilidades de sucesso dadas respectivamente por $P(U_{i,t}(\alpha_1) = 1) = \alpha_1$ e $P(U_{i,t}(\alpha_2) = 1) = \alpha_2$.

Para este modelo é assumido que $(Z_t)_{t \in \mathbb{Z}}$ é uma sucessão de v.a.'s independentes com distribuição de Poisson de parâmetro λ , onde para cada t , Z_t é independente de X_{t-1} e de $\phi_t \circ X_{t-1}$. O modelo em análise contempla apenas dois regimes, que se distinguem na parte auto-regressiva. Na Figura 3.1 apresentam-se algumas realizações do processo SETINAR(2;1). Nas figuras (a) e (b) as realizações tendem a mover-se bastantes vezes entre regimes, reflectindo o facto do parâmetro α_1 ser superior ao parâmetro α_2 . Nos casos (c) a (f) a probabilidade de permanência no sistema no primeiro regime é inferior à do

segundo, o que significa que a mudança entre regimes é menos frequente. Neste segunda situação a proximidade entre os parâmetros do operador *binomial thinning* dos dois regimes também não permite distinguir com muita clareza a existência dos dois regimes.

Proposição 3.1. X_t é uma cadeia de Markov irredutível, aperiódica e recorrente positiva, e portanto ergódica, o que garante a existência de um processo estritamente estacionário que satisfaz (3.1).

Demonstração Tendo em conta a definição do modelo (3.1), X_t é uma cadeia de Markov com espaço de estados \mathbb{N}_0 , cujas probabilidades de transição são dadas por

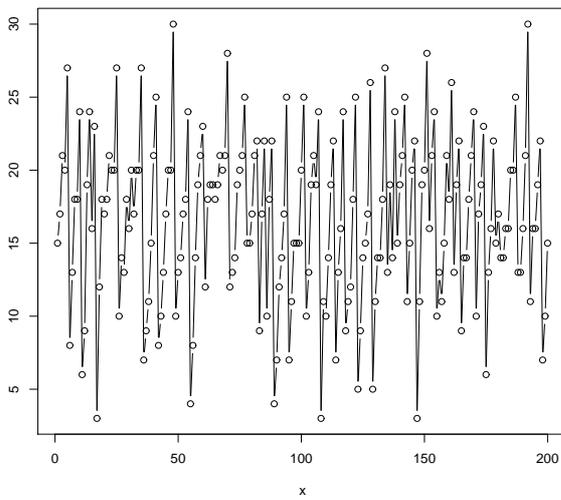
$$\begin{aligned} P_{i,j} &= P(X_t = j | X_{t-1} = i) \\ &= p(i, j, \alpha_1, \lambda) I_{\{i \leq R\}} + p(i, j, \alpha_2, \lambda) I_{\{i > R\}} \\ &= p(i, j, \alpha_1 I_{\{i \leq R\}} + \alpha_2 I_{\{i > R\}}, \lambda), \end{aligned}$$

onde

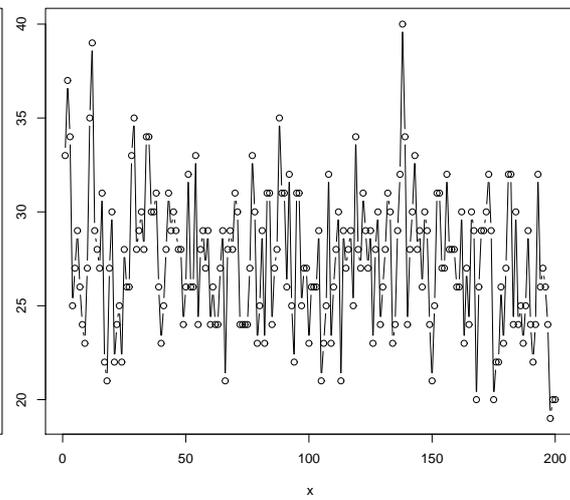
$$p(i, j, \alpha_k, \lambda) = \sum_{m=0}^{\min(i,j)} C_m^i \alpha_k^m (1 - \alpha_k)^{i-m} e^{-\lambda} \frac{\lambda^{j-m}}{(j-m)!} > 0; \quad k = 1, 2, \quad (3.2)$$

o que assegura que a cadeia é aperiódica e irredutível. Para mostrar que a cadeia é recorrente positiva basta provar que $\sum_{t=1}^{+\infty} P^t(0, 0) = +\infty$ (pois X_t é irredutível) com $P^t(x, y) = P(X_t = y | X_0 = x)$. Aplicando sucessivamente (3.1) obtém-se

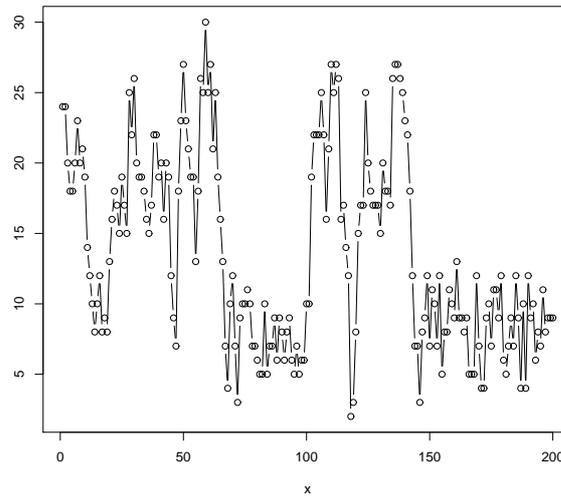
$$X_t = \phi_t \circ \phi_{t-1} \circ \cdots \circ \phi_1 \circ X_0 + \sum_{i=1}^{t-1} \phi_{t-1} \circ \phi_{t-2} \circ \cdots \circ \phi_{t-i} \circ Z_{t-i} + Z_t,$$



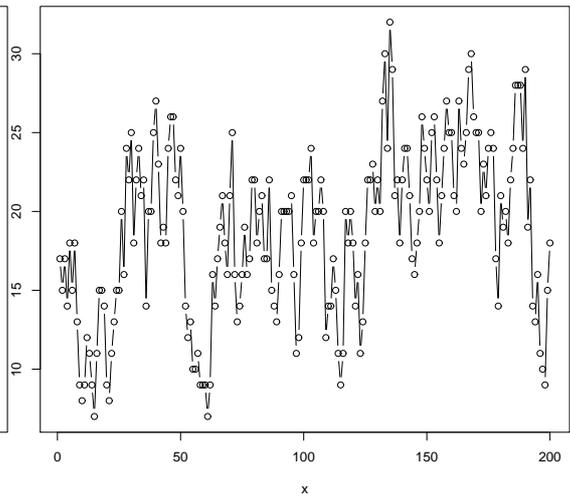
(a) $\lambda = 7; \alpha_1 = 0.8; \alpha_2 = 0.1; R = 21$



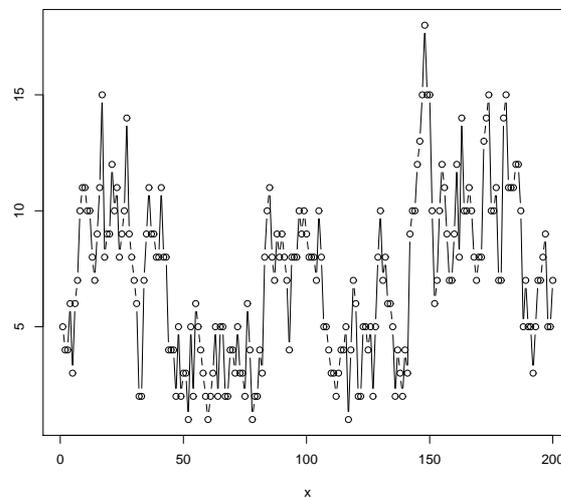
(b) $\lambda = 7; \alpha_1 = 0.8; \alpha_2 = 0.65; R = 28$



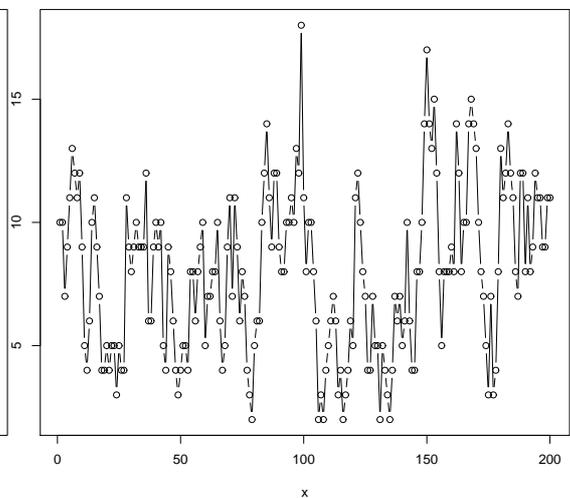
(c) $\lambda = 7; \alpha_1 = 0.2; \alpha_2 = 0.65; R = 14$



(d) $\lambda = 7; \alpha_1 = 0.5; \alpha_2 = 0.65; R = 17$



(e) $\lambda = 3; \alpha_1 = 0.2; \alpha_2 = 0.65; R = 6$



(f) $\lambda = 3; \alpha_1 = 0.5; \alpha_2 = 0.65; R = 7$

Figura 3.1: Gráficos de algumas realizações do modelo SETINAR(2;1).

permitindo escrever

$$\begin{aligned}
P^t(0, 0) &= \\
&= P\left(\sum_{i=1}^{t-1} \phi_{t-1} \circ \phi_{t-2} \circ \cdots \circ \phi_{t-i} \circ Z_{t-i} + Z_t = 0 \mid X_0 = 0\right) \\
&= P(Z_t = 0, \phi_{t-1} \circ Z_{t-1} = 0, \dots, \phi_{t-1} \circ \phi_{t-2} \circ \cdots \circ \phi_1 \circ Z_1 = 0 \mid X_0 = 0) \\
&= \sum_{i_2=1}^2 \cdots \sum_{i_{t-2}=1}^2 \sum_{i_{t-1}=1}^2 P(\phi_2 = \alpha_{i_2}, \dots, \phi_{t-2} = \alpha_{i_{t-2}}, \phi_{t-1} = \alpha_{i_{t-1}} \mid X_0 = 0) \times \\
&\times P(Z_t = 0, \alpha_{i_{t-1}} \circ Z_{t-1} = 0, \dots, \alpha_{i_{t-1}} \circ \alpha_{i_{t-2}} \circ \cdots \circ \alpha_{i_2} \circ Z_1 = 0 \mid X_0 = 0, \phi_2 = \alpha_{i_2}, \\
&\quad \phi_3 = \alpha_{i_3}, \dots, \phi_{t-1} = \alpha_{i_{t-1}}) \\
&= \sum_{i_2=1}^2 \cdots \sum_{i_{t-2}=1}^2 \sum_{i_{t-1}=1}^2 P(\phi_2 = \alpha_{i_2}, \dots, \phi_{t-2} = \alpha_{i_{t-2}}, \phi_{t-1} = \alpha_{i_{t-1}} \mid X_0 = 0) \times \\
&\quad \times e^{-\lambda(1+\alpha_{i_{t-1}}+\alpha_{i_{t-1}}\alpha_{i_{t-2}}+\cdots+\alpha_{i_{t-1}}\alpha_{i_{t-2}}\cdots\alpha_{i_2})}.
\end{aligned}$$

A última expressão implica que

$$e^{-\lambda \frac{1-\max(\alpha_1, \alpha_2)^t}{1-\max(\alpha_1, \alpha_2)}} \leq P^t(0, 0) \leq e^{-\lambda \frac{1-\min(\alpha_1, \alpha_2)^t}{1-\min(\alpha_1, \alpha_2)}}.$$

Como $P^t(0, 0) > 0$, e usando o critério de comparação para séries divergentes, conclui-se que $\sum_{i=1}^{+\infty} P^t(0, 0) = +\infty$. Isto prova que X_t é uma cadeia de Markov recorrente positiva e portanto ergódica, o que garante a existência da distribuição estritamente estacionária de (3.1). ■

Nota 3.2. À semelhança do que acontece para a maioria dos modelos de valores contínuos definidos por limiares, não foi possível obter uma fórmula analítica explícita para a distribuição marginal estacionária do modelo SETINAR. A partir da função geradora de probabilidades de X_t não se conseguiu identificar a sua distribuição marginal. Considerando $f_1(s) = \sum s^x P(X_t = x \mid X_t \leq R)$ e $f_2(s) = \sum s^x P(X_t = x \mid X_t > R)$ as funções geradoras de probabilidades de X_t condicional a X_t pertencer respectivamente, ao ramo inferior e ao ramo superior, a função geradora de probabilidades de X_t pode ser escrita na forma $P_{X_t}(s) = p f_1(s) + (1 - p) f_2(s)$. Por outro lado, tendo em conta o modelo definido em (3.1) e as propriedades da operação *thinning*, a função geradora de probabilidades de

X_t pode ser escrita através da expressão $P_{X_t}(s) = [pf_1(1 - \alpha_1 + \alpha_1 s) + (1 - p)f_2(1 - \alpha_2 + \alpha_2 s)]e^{-\lambda(1-s)}$. Igualando ambas as expressões a $P_{X_t}(s)$ obtém-se a seguinte equação $p(f_1(s) - f_1(1 - \alpha_1 + \alpha_1 s)e^{-\lambda(1-s)}) + (1 - p)(f_2(s) - f_2(1 - \alpha_2 + \alpha_2 s)e^{-\lambda(1-s)}) = 0$, que poderá não ter solução única. Quando se considera igual a zero cada uma das parcelas do primeiro membro da equação obtém-se $f_1(s) = e^{-\frac{\lambda}{1-\alpha_1}(1-s)}$ e $f_2(s) = e^{-\frac{\lambda}{1-\alpha_2}(1-s)}$; que não são mais do que funções geradoras de probabilidade de distribuições de Poisson de parâmetros $\frac{\lambda}{1-\alpha_1}$ e $\frac{\lambda}{1-\alpha_2}$, respectivamente. No entanto não se consegue mostrar que esta solução é única, uma vez que $f_i(s) \leq f_i(1 - \alpha_i + \alpha_i s)$, $i = 1, 2$; mas já não se conhece a relação entre $f_i(s)$ e $f_i(1 - \alpha_i + \alpha_i s)e^{-\lambda(1-s)}$ uma vez que $e^{-\lambda(1-s)} \leq 1$. Por outro lado quando $\alpha_1 > \alpha_2$ tem-se que $\frac{\lambda}{1-\alpha_1} > \frac{\lambda}{1-\alpha_2}$, o que significaria que o ramo inferior teria uma maior média que o ramo superior, indicando que nestes casos estas funções geradoras de probabilidade deixam de fazer sentido.

Lema 3.3. *Seja X_t o processo definido pela equação (3.1). Se $E[X_0^k] < \infty$ então, para $k \in \mathbb{N}$, $E[X_t^k] < \infty$.*

Demonstração Apenas se mostra o resultado para $k \leq 3$; no entanto como a distribuição de Poisson tem momentos finitos de qualquer ordem $k \in \mathbb{N}$, para $k > 3$ é possível usar os mesmos argumentos que serão aqui utilizados para $k \leq 3$. Para o caso $k = 1$ o objectivo é mostrar que

$$E[X_t] \leq \alpha_{max}^t E[X_0] + \mu_z \sum_{i=0}^{t-1} \alpha_{max}^i,$$

onde $\alpha_{max} = \max(\alpha_1, \alpha_2)$. Facilmente se verifica que a desigualdade é válida para $E[X_1]$. Considere-se a desigualdade válida para $E[X_{t-1}]$, então

$$\begin{aligned} E[X_t] &\leq \alpha_{max} E[X_{t-1}] + \mu_z \\ &\leq \alpha_{max} \left(\alpha_{max}^{t-1} E[X_0] + \mu_z \sum_{i=0}^{t-2} \alpha_{max}^i \right) + \mu_z \\ &\leq \alpha_{max}^t E[X_0] + \mu_z \sum_{i=0}^{t-1} \alpha_{max}^i < \infty, \end{aligned}$$

para $E[X_0] < \infty$. A partir da última desigualdade é possível afirmar que

$$E[X_t] \leq \alpha_{max}^t E[X_0] + \frac{\mu_z}{1 - \alpha_{max}}. \quad (3.3)$$

Para $k = 2$ é necessário desenvolver o segundo momento de X_t , para encontrar a expressão que se pretende majorar.

$$\begin{aligned} E[X_t^2] &= \\ &= E[(\phi_t \circ X_{t-1})^2] + E[Z_{t-1}^2] + 2\mu_z E[\phi_1 \circ X_{t-1}] \\ &= E[E[(\phi_1 \circ X_{t-1})^2 | X_{t-1}]] + (\sigma_Z^2 + \mu_z^2) + 2\mu_z(\alpha_1 E[X_{t-1} I_{0,1}] + \alpha_2 E[X_{t-1} I_{0,2}]) \\ &= E[(\alpha_1 \circ X_{t-1})^2 I_{0,1} + (\alpha_2 \circ X_{t-1})^2 I_{0,2}] + (\sigma_Z^2 + \mu_z^2) + \\ &\quad + 2\mu_z(\alpha_1 E[X_{t-1} I_{0,1}] + \alpha_2 E[X_{t-1} I_{0,2}]) \\ &= \alpha_1(1 - \alpha_1)E[X_{t-1} I_{0,1}] + \alpha_1^2 E[X_{t-1}^2 I_{0,1}] + \alpha_2(1 - \alpha_2)E[X_{t-1} I_{0,2}] + \alpha_2^2 E[X_{t-1}^2 I_{0,2}] + \\ &\quad + (\sigma_Z^2 + \mu_z^2) + 2\mu_z(\alpha_1 E[X_{t-1} I_{0,1}] + \alpha_2 E[X_{t-1} I_{0,2}]) \\ &\leq \alpha_{max}^2 E[X_{t-1}^2] + \frac{1}{4} E[X_{t-1}] + \sigma_Z^2 + \mu_z^2 + 2\mu_z \alpha_{max} E[X_{t-1}] \\ &\quad \text{usando (3.3)} \\ &\leq \alpha_{max}^2 E[X_{t-1}^2] + \left(\frac{1}{4} \alpha_{max}^{t-1} + 2\mu_z \max(\alpha_1, \alpha_2)^t \right) E[X_0] + \sigma_Z^2 + \\ &\quad + \frac{\mu_z}{4(1 - \alpha_{max})} + \mu_z^2 + 2\mu_z \alpha_{max} \frac{\mu_z}{1 - \alpha_{max}}. \end{aligned}$$

Pretende-se mostrar que

$$\begin{aligned} E[X_t^2] &\leq \alpha_{max}^{2t} E[X_0^2] + \left(\frac{1}{4} \alpha_{max}^{t-1} E[X_0] + 2\mu_z E[X_0] \alpha_{max}^t \right) \sum_{i=0}^{t-1} \alpha_{max}^i + \\ &\quad + \left(\sigma_Z^2 + \mu_z^2 + \frac{\mu_z}{4(1 - \alpha_{max})} + 2 \frac{\mu_z^2 \alpha_{max}}{1 - \alpha_{max}} \right) \sum_{i=0}^{t-1} \alpha_{max}^{2i}. \end{aligned}$$

Para $t = 1$ a expressão anterior é válida, uma vez que

$$\begin{aligned} E[X_1^2] &\leq \alpha_{max}^2 E[X_0^2] + \frac{1}{4} E[X_0] + \sigma_Z^2 + \mu_z^2 + 2\mu_z \alpha_{max} E[X_0] \\ &\leq \alpha_{max}^2 E[X_0^2] + \frac{1}{4} E[X_0] + 2\mu_z \alpha_{max} E[X_0] + \sigma_Z^2 + \mu_z^2 + \\ &\quad + \frac{\mu_z}{4(1 - \alpha_{max})} + 2 \frac{\mu_z^2 \alpha_{max}}{1 - \alpha_{max}}. \end{aligned}$$

Considere-se válida a desigualdade para $E(X_{t-1}^2)$. Então

$$\begin{aligned}
 E[X_t^2] &= E[(\phi_t \circ X_{t-1} + Z_t)^2] \\
 &\leq \alpha_{max}^2 E[X_{t-1}^2] + \frac{1}{4} E[X_{t-1}] + \sigma_Z^2 + \mu_z^2 + 2\mu_z \alpha_{max} E[X_{t-1}] \\
 &\leq \alpha_{max}^{2t} E[X_0^2] + \left(\frac{1}{4} \alpha_{max}^t E[X_0] + 2\mu_z E[X_0] \alpha_{max}^{t+1} \right) \times \\
 &\quad \times \sum_{i=0}^{t-2} \alpha_{max}^i + \left(\sigma_Z^2 + \mu_z^2 + \frac{\mu_z}{4(1-\alpha_{max})} + 2\frac{\mu_z^2 \alpha_{max}}{1-\alpha_{max}} \right) \times \\
 &\quad \times \alpha_{max}^2 \sum_{i=0}^{t-2} \alpha_{max}^{2i} + \frac{1}{4} \alpha_{max}^{t-1} E[X_0] + \frac{\mu_z}{4(1-\alpha_{max})} + \\
 &\quad + \sigma_Z^2 + \mu_z^2 + 2\mu_z \alpha_{max}^t E[X_0] + 2\mu_z \alpha_{max} \frac{\mu_z}{1-\alpha_{max}} \\
 &= \alpha_{max}^{2t} E[X_0^2] + \left(\frac{1}{4} \alpha_{max}^{t-1} E[X_0] + 2\mu_z E[X_0] \alpha_{max}^t \right) \times \\
 &\quad \times \sum_{i=0}^{t-1} \alpha_{max}^i + \left(\sigma_Z^2 + \mu_z^2 + \frac{\mu_z}{4(1-\alpha_{max})} + 2\frac{\mu_z^2 \alpha_{max}}{1-\alpha_{max}} \right) \times \sum_{i=0}^{t-1} \alpha_{max}^{2i} < \infty,
 \end{aligned}$$

uma vez que $E[X_0^j] < \infty$, $j = 1, 2$. Com vista à determinação da expressão que majora o terceiro momento de X_t , e atendendo a que $\sum_{i=0}^{t-1} \alpha_{max}^{mi} \leq \frac{1}{1-\alpha_{max}^m}$ ($m > 0$), pode-se escrever

$$E[X_t^2] \leq \alpha_{max}^{2t} E[X_0^2] + const_1 \alpha_{max}^t E[X_0] + const_2 \quad (3.4)$$

$$\text{com } const_1 = \frac{1}{4\alpha_{max}} + 2\mu_z \text{ e } const_2 = \left(\sigma_Z^2 + \mu_z^2 + \frac{\mu_z}{4(1-\alpha_{max})} + 2\frac{\mu_z^2 \alpha_{max}}{1-\alpha_{max}} \right) \frac{1}{1-\alpha_{max}^2}.$$

Desenvolvendo o momento de ordem 3 de X_t obtém-se

$$\begin{aligned}
E[X_t^3] &= \sum_{i=0}^3 C_i^3 E[(\phi_t \circ X_{t-1})^i] E[Z_t^{3-i}] \\
&= E[(\phi_t \circ X_{t-1})^3] + 3E[(\phi_t \circ X_{t-1})^2] E[Z_t] + 3E[\phi_t \circ X_{t-1}] E[Z_t^2] + E[Z_t^3] \\
&\leq \alpha_{max}^3 E[X_{t-1}^3] + E[Z_t^3] + \frac{3}{2} E[X_{t-1}^2] + \frac{1}{2} E[X_{t-1}] + \\
&\quad + 3 \left(\alpha_{max}^2 E[X_{t-1}^2] + \frac{1}{4} E[X_{t-1}] \right) \mu_Z + 3\alpha_{max} E[X_{t-1}] (\sigma_Z^2 + \mu_z) \\
&\leq \alpha_{max}^3 E[X_{t-1}^3] + E[Z_t^3] + \left(\frac{3}{2} + 3\alpha_{max}^2 \mu_z \right) E[X_{t-1}^2] + \\
&\quad + \left(\frac{1}{2} + \frac{3}{4} \mu_z + 3\alpha_{max} (\sigma_Z^2 + \mu_z) \right) E[X_{t-1}] \\
&\quad \text{usando (3.3) e (3.4)} \\
&\leq \alpha_{max}^{3t} E[X_0^3] + E[Z_t^3] + \left(\frac{3}{2} + 3\alpha_{max}^2 \mu_z \right) \alpha_{max}^{2t-2} E[X_0^2] + \\
&\quad + \alpha_{max}^{t-1} E[X_0] \left[\left(\frac{3}{2} + 3\alpha_{max}^2 \mu_z \right) const_1 + \frac{1}{2} + \frac{3}{4} \mu_z + 3\alpha_{max} (\sigma_Z^2 + \mu_z) \right] + \\
&\quad + \left(\frac{3}{2} + 3\alpha_{max}^2 \mu_z \right) const_2 + \left(\frac{1}{2} + \frac{3}{4} \mu_z + 3\alpha_{max} (\sigma_Z^2 + \mu_z) \right) \frac{\mu_z}{1 - \alpha_{max}}.
\end{aligned}$$

A partir desta desigualdade mostra-se, usando os mesmos argumentos que foram utilizados para $k = 2$, que

$$\begin{aligned}
E[X_t^3] &\leq \alpha_{max}^3 E[X_0^3] + \left(\frac{3}{2} + 3\alpha_{max}^2 \mu_z \right) \alpha_{max}^{2t-2} E[X_0^2] \sum_{i=0}^{t-1} \alpha_{max}^i \\
&\quad + E[X_0] \alpha_{max}^{t-1} \sum_{i=0}^{t-1} \alpha_{max}^{2i} \left(\left(\frac{3}{2} + 3\alpha_{max}^2 \mu_z \right) const_1 + \frac{1}{2} + \frac{3}{4} \mu_z + 3\alpha_{max} (\sigma_Z^2 + \mu_z) \right) \\
&\quad + \sum_{i=0}^{t-1} \alpha_{max}^{3i} \left(\left(\frac{3}{2} + 3\alpha_{max}^2 \mu_z \right) const_2 + E[Z_t^3] + \left(\frac{1}{2} + \right. \right. \\
&\quad \left. \left. + \frac{3}{4} \mu_z + 3\alpha_{max} (\sigma_Z^2 + \mu_z) \right) \frac{\mu_z}{1 - \alpha_{max}} \right) < \infty,
\end{aligned}$$

para $E[X_0^j] < \infty$, $j = 1, 2, 3$. ■

Nota 3.4. Note-se que uma vez que $E[X_t^k] < \infty$ então $m_k^{(1)} = E[X_t^k | I_{t,1}] < \infty$ e $m_k^{(2)} = E[X_t^k | I_{t,2}] < \infty$.

Para a determinação da média, variância e a função de autocovariância do processo defina-se $p = P(X_t \leq R)$, $u_1 = E[X_t | X_t \leq R]$, $u_2 = E[X_t | X_t > R]$, $\sigma_1^2 = V[X_t | X_t \leq R]$, $\sigma_2^2 = V[X_t | X_t > R]$, e $\gamma_k^{(1)} = Cov(X_t, X_{t+k} | X_{t+k} \leq R)$ e $\gamma_k^{(2)} = Cov(X_t, X_{t+k} | X_{t+k} > R)$.

Lema 3.5. Para o modelo $SETINAR(2;1)$, com $E[X_0^2] < \infty$, a média é dada por

$$\mu = E[X_t] = p\alpha_1 u_1 + (1-p)\alpha_2 u_2 + \lambda,$$

enquanto que a variância toma a forma

$$\begin{aligned} \sigma^2 = V[X_t] &= p(\alpha_1^2 \sigma_1^2 + \alpha_1(1-\alpha_1)u_1) + (1-p)(\alpha_2^2 \sigma_2^2 + \alpha_2(1-\alpha_2)u_2) + \lambda + \\ &+ p(1-p)(\alpha_1 u_1 - \alpha_2 u_2)^2. \end{aligned}$$

A função de autocovariância $\gamma(k) = Cov(X_t, X_{t+k})$ é dada por

$$\gamma(k) = \begin{cases} p\alpha_1 \sigma_1^2 + (1-p)\alpha_2 \sigma_2^2 + p\alpha_1 u_1(u_1 - \mu) + (1-p)\alpha_2 u_2(u_2 - \mu) & k = 1 \\ p\alpha_1 \gamma_{k-1}^{(1)} + (1-p)\alpha_2 \gamma_{k-1}^{(2)} + p\alpha_1 u_1(E[X_t | X_{t+k-1} \leq R] - \mu) \\ \quad + (1-p)\alpha_2 u_2(E[X_t | X_{t+k-1} > R] - \mu) & k \neq 1. \end{cases}$$

Demonstração

$$\begin{aligned} E[X_t] &= E[\phi_{t-1} \circ X_{t-1}] + \lambda \\ &= E[I_{t-1,1} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_1)] + E[I_{t-1,2} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_2)] + \lambda \\ &= p\alpha_1 E[X_{t-1} | X_{t-1} \leq r] + (1-p)\alpha_2 E[X_{t-1} | X_{t-1} > r] + \lambda \\ &= p\alpha_1 u_1 + (1-p)\alpha_2 u_2 + \lambda. \end{aligned}$$

No caso da variância de X_t tem-se

$$\begin{aligned}
V[X_t] &= V[\phi_{t-1} \circ X_{t-1}] + \lambda \\
&= V\left[I_{t-1,1} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_1) + I_{t-1,2} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_2)\right] + \lambda \\
&= V\left[I_{t-1,1} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_1)\right] + V\left[I_{t-1,2} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_2)\right] + \\
&\quad + 2Cov\left(I_{t-1,1} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_1), I_{t-1,2} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_2)\right) + \lambda \\
&= I + II + III + \lambda.
\end{aligned} \tag{3.5}$$

O primeiro termo do segundo membro da igualdade em (3.5) é

$$\begin{aligned}
I &= V\left[E\left[I_{t-1,1} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_1) \mid X_{t-1}\right]\right] + E\left[V\left[I_{t-1,1} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_1) \mid X_{t-1}\right]\right] \\
&= V\left[I_{t-1,1} \alpha_1 X_{t-1}\right] + E\left[I_{t-1,1} \alpha_1 (1 - \alpha_1) X_{t-1}\right] \\
&= \alpha_1^2 V\left[I_{t-1,1} X_{t-1}\right] + p \alpha_1 (1 - \alpha_1) u_1 \\
&= \alpha_1^2 E\left[I_{t-1,1} X_{t-1}^2\right] - \alpha_1^2 p^2 u_1^2 + p \alpha_1 (1 - \alpha_1) u_1 \\
&= \alpha_1^2 p (\sigma_1^2 + u_1^2) - \alpha_1^2 p^2 u_1^2 + p \alpha_1 (1 - \alpha_1) u_1 \\
&= p (\alpha_1^2 \sigma_1^2 + \alpha_1 (1 - \alpha_1) u_1) + p (1 - p) \alpha_1^2 u_1^2.
\end{aligned} \tag{3.6}$$

Pelos mesmos argumentos utilizados para I segue que

$$II = (1 - p) (\alpha_2^2 \sigma_2^2 + \alpha_2 (1 - \alpha_2) u_2) + p (1 - p) \alpha_2^2 u_2^2. \tag{3.7}$$

Finalmente III assume a forma

$$III = -2 \prod_{j=1}^2 E\left(I_{t-1,j} \sum_{i=1}^{X_{t-1}} U_{i,t}(\alpha_j)\right) = -2p(1-p) \prod_{j=1}^2 \alpha_j u_j.$$

A expressão da variância de X_t apresentada no Lema 3.5 resulta da substituição de (3.6), (3.7), e (3.8) na equação (3.5). Para determinar a função de autocovariância comece-se

por calcular esta função para $k > 0$,

$$\begin{aligned}
 Cov(X_t, X_{t+k}) &= Cov(X_t, \phi_{t+k} \circ X_{t+k-1} + Z_{t+k}) \\
 &= Cov(X_t, \phi_{t+k} \circ X_{t+k-1}) \\
 &= E[X_t \phi_{t+k} \circ X_{t+k-1}] - E[X_t]E[\phi_{t+k} \circ X_{t+k-1}] \\
 &= p\alpha_1 E[X_t X_{t+k-1} | X_{t+k-1} \leq R] + (1-p)\alpha_2 E[X_t X_{t+k-1} | X_{t+k-1} > R] - \\
 &\quad - E[X_t](p\alpha_1 u_1 + (1-p)\alpha_2 u_2) \\
 &= p\alpha_1 Cov(X_t, X_{t+k-1} | X_{t+k-1} \leq R) + p\alpha_1 u_1 E[X_t | X_{t+k-1} \leq R] + \\
 &\quad + \alpha_2 u_2 E[X_t | X_{t+k-1} > R] + (1-p)\alpha_2 Cov(X_t, X_{t+k-1} | X_{t+k-1} > R) - \\
 &\quad - E[X_t](p\alpha_1 u_1 + (1-p)\alpha_2 u_2) \\
 &= p\alpha_1 \gamma_{k-1}^{(1)} + (1-p)\alpha_2 \gamma_{k-1}^{(2)} + p\alpha_1 u_1 (E[X_t | X_{t+k-1} \leq R] - \mu) + \\
 &\quad + (1-p)\alpha_2 u_2 (E[X_t | X_{t+k-1} > R] - \mu)
 \end{aligned}$$

e

$$\begin{aligned}
 Cov(X_t, X_{t-k}) &= Cov(X_{t-k}, \phi_t \circ X_{t-1}) \\
 &= E[X_{t-k} \phi_t \circ X_{t-1}] - E[X_{t-k}]E[\phi_t \circ X_{t-1}] \\
 &= p\alpha_1 E[X_{t-k} X_{t-1} | X_{t-1} \leq R] + (1-p)\alpha_2 E[X_{t-k} X_{t-1} | X_{t-1} > R] - \\
 &\quad - E[X_{t-k}]E[\phi_t \circ X_{t-1}] \\
 &= p\alpha_1 Cov(X_{t-k}, X_{t-1} | X_{t-1} \leq R) + p\alpha_1 u_1 E[X_{t-k} | X_{t-1} \leq R] + \\
 &\quad + \alpha_2 u_2 E[X_{t-k} | X_{t-1} > R] + (1-p)\alpha_2 Cov(X_{t-k}, X_{t-1} | X_{t-1} > R) + \\
 &\quad + (1-p) - E[X_{t-k}](p\alpha_1 u_1 + (1-p)\alpha_2 u_2) \\
 &= p\alpha_1 Cov(X_t, X_{t+k-1} | X_{t+k-1} \leq R) + p\alpha_1 u_1 E[X_t | X_{t+k-1} \leq R] + \\
 &\quad + \alpha_2 u_2 E[X_t | X_{t+k-1} > R] + (1-p)\alpha_2 Cov(X_t, X_{t+k-1} | X_{t+k-1} > R) - \\
 &\quad - E[X_t](p\alpha_1 u_1 + (1-p)\alpha_2 u_2) \\
 &= p\alpha_1 \gamma_{k-1}^{(1)} + (1-p)\alpha_2 \gamma_{k-1}^{(2)} + p\alpha_1 u_1 (E[X_t | X_{t+k-1} \leq R] - \mu) + \\
 &\quad + (1-p)\alpha_2 u_2 (E[X_t | X_{t+k-1} > R] - \mu).
 \end{aligned}$$

Para $k = 1$ basta fazer a substituição na equação anterior obtendo-se

$$\gamma_1 = Cov(X_t, X_{t+1}) = p\alpha_1\sigma_1^2 + (1-p)\alpha_2\sigma_2^2 + p\alpha_1u_1(u_1 - \mu) + (1-p)\alpha_2u_2(u_2 - \mu).$$

■

Nota 3.6. Relativamente ao Lema 3.5 note-se que a média do processo pode ser também escrita na forma $E[X_t] = pu_1 + (1-p)u_2$, o que igualando ambas as expressões para a média resulta na seguinte equação $p(1-\alpha_1)u_1 + (1-p)(1-\alpha_2)u_2 - \lambda = 0$. Esta equação não possui uma solução analítica única em termos de u_1 e u_2 , esta pode ser resolvida em ordem a u_2 , ficando em função de u_1 . O objectivo seria encontrar u_1 e u_2 , para se conseguir escrever $E[X_t]$ de forma explícita.

Uma solução da equação anterior em que u_2 não depende explicitamente de u_1 obtém-se fazendo $u_1 = \frac{\lambda}{1-\alpha_1}$ e $u_2 = \frac{\lambda}{1-\alpha_2}$, que podem ser encarados como médias de processos INAR(1) Poisson com parâmetros associados à operação *thinning* α_1 e α_2 , respectivamente. Várias experiências de simulação indicaram que, para $\alpha_1 < \alpha_2$ as estimativas das médias dos ramos estão próximas deste valor; o mesmo não aconteceu no caso em que $\alpha_1 > \alpha_2$. Nesta última situação tem-se que $\frac{\lambda}{1-\alpha_1} > \frac{\lambda}{1-\alpha_2}$, o que violaria o facto de $u_2 > u_1$. Em termos práticos poder-se-á usar as expressões obtidas no Lema anterior para encontrar estimativas para a média e variância do processo. Basta, para cada um dos regimes, usar as respectivas médias e variâncias amostrais e usar a interpretação frequentista da probabilidade para estimar p .

Nos modelos auto-regressivos, inteiros ou contínuos, é possível escrever a função de autocovariância através de uma forma explícita em função dos seus parâmetros auto-regressivos e do *lag* k ou através de uma expressão recursiva. Para os modelos auto-regressivos com limiares tal não acontece e portanto optou-se por escrever a função de autocovariância de *lag* k em função da autocovariância condicional no *lag* imediatamente anterior. Note-se que a expressão que se obteve não pode ser usada de uma forma recursiva pois, para cada aplicação que se faça dessa expressão ter-se-á de adaptar e condicionar às variáveis que vão surgindo.

3.2 Métodos de estimação

Nesta secção abordar-se-ão os métodos dos mínimos quadrados condicionais e da máxima verosimilhança condicional. O método baseado nos momentos não é utilizado para estimar os parâmetros do modelo pois a percentagem de estimativas válidas que produz é muito reduzida, e como tal não é aqui analisado.

Considere-se (X_1, \dots, X_n) uma sucessão de v.a.'s que satisfazem (3.1) e $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) = (\alpha_1, \alpha_2, \lambda)$ o vector dos parâmetros desconhecidos. Ao longo desta secção assume-se que o nível do limiar R é conhecido.

3.2.1 Estimadores dos mínimos quadrados condicionais

Os estimadores MQC, $\hat{\boldsymbol{\theta}}_{MQC} = (\hat{\alpha}_{1,MQC}, \hat{\alpha}_{2,MQC}, \hat{\lambda}_{MQC})$, de $\boldsymbol{\theta}$ formam o vector que minimiza a expressão

$$Q(\boldsymbol{\theta}) = \sum_{t=2}^n (X_t - g(\boldsymbol{\theta}, X_{t-1}))^2 = \sum_{t=2}^n (X_t - \alpha_1 X_{t-1} I_{t-1,1} - \alpha_2 X_{t-1} I_{t-1,2} - \lambda)^2 = \sum_{t=2}^n U_t^2$$

e que se obtêm através da resolução do sistema

$$\begin{bmatrix} \sum_{t=2}^n X_{t-1}^2 I_{t-1,1} & 0 & \sum_{t=2}^n X_{t-1} I_{t-1,1} \\ 0 & \sum_{t=2}^n X_{t-1}^2 I_{t-1,2} & \sum_{t=2}^n X_{t-1} I_{t-1,2} \\ \sum_{t=2}^n X_{t-1} I_{t-1,1} & \sum_{t=2}^n X_{t-1} I_{t-1,2} & n-1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} \sum_{t=2}^n X_t X_{t-1} I_{t-1,1} \\ \sum_{t=2}^n X_t X_{t-1} I_{t-1,2} \\ \sum_{t=2}^n X_t \end{bmatrix}.$$

O resultado seguinte estabelece a distribuição assintótica dos estimadores MQC.

Teorema 3.7. *Os estimadores MQC são fortemente consistentes e a sua distribuição é assintoticamente normal, i.e.,*

$$n^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}) \xrightarrow{d} N(0, V^{-1}WV^{-1}),$$

onde V e W são matrizes quadradas de ordem 3, cujos elementos são

$$V_{ij} = E \left[\frac{\partial}{\partial \theta_i} g(\boldsymbol{\theta}, X_{t-1}) \frac{\partial}{\partial \theta_j} g(\boldsymbol{\theta}, X_{t-1}) \right] e$$

$$W_{ij} = E \left[U_t^2 \frac{\partial}{\partial \theta_i} g(\boldsymbol{\theta}, X_{t-1}) \frac{\partial}{\partial \theta_j} g(\boldsymbol{\theta}, X_{t-1}) \right],$$

respectivamente.

Demonstração Para demonstrar o resultado anterior basta aplicar os Teoremas 3.1 e 3.2 de Klimko e Nelson (1978). Facilmente se verifica que $g(\boldsymbol{\theta}, X_{t-1})$, $\frac{\partial g(\boldsymbol{\theta}, X_{t-1})}{\partial \theta_i}$, $\frac{\partial^2 g(\boldsymbol{\theta}, X_{t-1})}{\partial \theta_i \partial \theta_j}$ e $\frac{\partial^3 g(\boldsymbol{\theta}, X_{t-1})}{\partial \theta_i \partial \theta_j \partial \theta_k}$, $i, j, k \in \{1, 2, 3\}$, satisfazem as condições de regularidade do Teorema 3.1, por conseguinte os estimadores MQC são fortemente consistentes. Tendo em conta que $m_k^{(1)} < \infty$ e $m_k^{(2)} < \infty$, para $k = 1, 2, 3$, a aplicação do Teorema 3.2 garante que $\hat{\boldsymbol{\theta}}_{MQC}$ tem distribuição normal com média nula e matriz de variâncias e covariâncias dada por $V^{-1}WV^{-1}$, onde

$$V = \begin{bmatrix} pm_2^{(1)} & 0 & pu_1 \\ 0 & (1-p)m_2^{(2)} & (1-p)u_2 \\ pu_1 & (1-p)u_2 & 1 \end{bmatrix}$$

e

$$W = \begin{bmatrix} p[\alpha_1(1-\alpha_1)m_3^{(1)} + \lambda m_2^{(1)}] & 0 & p[\alpha_1(1-\alpha_1)m_2^{(1)} + \lambda u_1] \\ 0 & (1-p)[\alpha_2(1-\alpha_2)m_3^{(2)} + \lambda m_2^{(2)}] & (1-p)[\alpha_2(1-\alpha_2)m_2^{(2)} + \lambda u_2] \\ p[\alpha_1(1-\alpha_1)m_2^{(1)} + \lambda u_1] & (1-p)[\alpha_2(1-\alpha_2)m_2^{(2)} + \lambda u_2] & \mu \end{bmatrix}.$$

Para finalizar a prova basta verificar que a matriz V é invertível. Esta matriz tem o seguinte determinante

$$\begin{aligned} |V| &= p(1-p) \left(m_2^{(1)} m_2^{(2)} - pu_1^2 m_2^{(2)} - (1-p)u_2^2 m_2^{(1)} \right) \\ &= p(1-p) \left(\sigma_1^2 \sigma_2^2 + p\sigma_1^2 u_2^2 + (1-p)u_1^2 \sigma_2^2 \right) > 0, \end{aligned}$$

o que permite concluir que V possui inversa. ■

3.2.2 Estimadores de máxima verosimilhança condicional

Considere-se x_0 fixo, então a função de verosimilhança condicional para o modelo SETINAR(2; 1) toma a forma

$$\begin{aligned} L(\boldsymbol{\theta}) &= P(X_1 = x_1, \dots, X_n = x_n | x_0) \\ &= \prod_{t=1}^n P(X_t = x_t | X_{t-1} = x_{t-1}) \\ &= \prod_{t=1}^n p(x_{t-1}, x_t, \alpha_1 I_{t-1,1} + \alpha_2 I_{t-1,2}, \lambda). \end{aligned}$$

Os estimadores MVC $\hat{\boldsymbol{\theta}}_{MVC} = (\hat{\alpha}_{1,MVC}, \hat{\alpha}_{2,MVC}, \hat{\lambda}_{MVC})$ são os maximizantes do logaritmo da função de verosimilhança condicional

$$l(\boldsymbol{\theta}) = \sum_{t=1}^n \log(p(x_{t-1}, x_t, \alpha_1 I_{t-1,1} + \alpha_2 I_{t-1,2}, \lambda)).$$

A partir da derivação parcial de primeira ordem, relativamente aos parâmetros, obtém-se o sistema não linear

$$\left\{ \begin{array}{l} \frac{1}{\alpha_i(1-\alpha_i)} \sum_{t=1}^n I_{t-1,i}(x_t - \alpha_i x_{t-1}) - \lambda \frac{p(x_{t-1}, x_t - 1, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} I_{t-1,i} = 0, \quad i = 1, 2 \\ \sum_{t=1}^n \left(\frac{p(x_{t-1}, x_t - 1, \alpha_1, \lambda)}{p(x_{t-1}, x_t, \alpha_1, \lambda)} I_{t-1,1} + \frac{p(x_{t-1}, x_t - 1, \alpha_2, \lambda)}{p(x_{t-1}, x_t, \alpha_2, \lambda)} I_{t-1,2} \right) - n = 0 \end{array} \right. ,$$

que não permite determinar expressões analíticas para os seus estimadores, sendo portanto necessário aplicar métodos numéricos para resolver este sistema. Na implementação deste método de estimação foi utilizado o método de Newton-Raphson, que se encontra na função `threshold_mv.R` descrita no apêndice A.

A distribuição assintótica dos estimadores de máxima verosimilhança condicional é apresentada no próximo resultado.

Teorema 3.8. *Para um valor fixo R , os estimadores MVC têm distribuição assintoticamente normal, i.e.*

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_{MVC} - \boldsymbol{\theta}) = \sqrt{n} \begin{bmatrix} \hat{\alpha}_{1,MVC} - \alpha_1 \\ \hat{\alpha}_{2,MVC} - \alpha_2 \\ \hat{\lambda}_{MVC} - \lambda \end{bmatrix} \xrightarrow{d} N(0, I(\boldsymbol{\theta})^{-1}), \quad (3.8)$$

onde $I(\boldsymbol{\theta})$ é a matriz de informação de Fisher.

Demonstração Franke e Seligman (1993) mostraram que as seguintes condições são válidas para o modelo INAR 1 cujas inovações têm distribuição de Poisson de parâmetro λ .

(C1) O conjunto $\{k : P(Z_t = k) = f(k, \lambda) = e^{-\lambda} \frac{\lambda^k}{k!}\}$ não depende de λ ;

(C2) $E[Z_t^3] = \lambda^3 + 3\lambda^2 + \lambda < \infty$;

(C3) $P(Z_t = k)$ é três vezes diferenciável relativamente a λ ;

(C4) Para todo o $\lambda \in B$ existe uma vizinhança U de λ tal que

1. $\sum_{k=0}^{\infty} \sup_{\lambda \in U} f(k, \lambda) < \infty$
2. $\sum_{k=0}^{\infty} \sup_{\lambda \in U} \left| \frac{\partial f(k, \lambda)}{\partial \lambda} \right| < \infty$,
3. $\sum_{k=0}^{\infty} \sup_{\lambda \in U} \left| \frac{\partial^2 f(k, \lambda)}{\partial \lambda^2} \right| < \infty$;

(C5) Para todo o $\lambda \in B$ existe uma vizinhança U de λ e sucessões $\psi_1(n) = \text{const}1.n$, $\psi_{11}(n) = \text{const}2.n^2$ e $\psi_{111}(n) = \text{const}3.n^3$, com $\text{const}1, \text{const}2, \text{const}3$ constantes adequadas, $n \geq 0$ tais que $\forall \lambda \in B$ e $\forall m \leq n$, com $P(Z_t)$ não nula,

$$\begin{aligned} \left| \frac{\partial f(m, \lambda)}{\partial \lambda} \right| &\leq \psi_1(n) f(m, \lambda) \\ \left| \frac{\partial^2 f(m, \lambda)}{\partial \lambda^2} \right| &\leq \psi_{11}(n) f(m, \lambda) \\ \left| \frac{\partial^3 f(m, \lambda)}{\partial \lambda^3} \right| &\leq \psi_{111}(n) f(m, \lambda) \end{aligned}$$

e relativamente à distribuição estacionária de (X_t)

$$\begin{aligned} E[\psi_1^3(X_1)] &< \infty \\ E[X_1\psi_{11}(X_2)] &< \infty \\ E[\psi_i(X_1)\psi_{11}(X_2)] &< \infty \\ E[\psi_{111}(X_1)] &< \infty; \end{aligned}$$

(C6) A matriz de informação de Fisher, $I(\boldsymbol{\theta})$, é não singular, onde

$$I(\boldsymbol{\theta})_{ij} = E \left[\frac{\partial}{\partial \theta_i} \log P(X_1, X_2) \frac{\partial}{\partial \theta_j} \log P(X_1, X_2) \right], \quad i, j = 1, 2, 3.$$

Para o modelo SETINAR(2; 1) é necessário verificar que as condições acima descritas ainda são válidas e que estas, em conjunto com a Proposição 3.1, implicam as condições gerais dos Teoremas 2.1 e 2.2 em Billingsley (1961) sobre a distribuição assintótica dos estimadores dos parâmetros de processos de Markov. As condições (C1)-(C5) dizem respeito apenas à distribuição das inovações e como tal são também válidas para o modelo SETINAR(2; 1). A condição (C2) implica que $E[X_t^3] < \infty$, o que garante as condições sobre os momentos em (C5).

Para verificar a condição (C6), que garante a não redundância dos parâmetros, comece-se por determinar as derivadas de primeira ordem da função de transição relativamente aos parâmetros,

$$\begin{aligned} &\frac{\partial p(x_{t-1}, x_t)}{\partial \lambda} \\ &= I_{t-1,1} \frac{\partial p(x_{t-1}, x_t, \alpha_1, \lambda)}{\partial \lambda} + I_{t-1,2} \frac{\partial p(x_{t-1}, x_t, \alpha_2, \lambda)}{\partial \lambda} \\ &= \sum_{i=1}^2 I_{t-1,i} \left[\left(\frac{x_t}{\lambda} - 1 \right) p(x_{t-1}, x_t, \alpha_i, \lambda) - \left(\frac{x_{t-1}\alpha_i}{\lambda} - 1 \right) p(x_{t-1} - 1, x_t - 1, \alpha_i, \lambda) \right] \end{aligned} \tag{3.9}$$

e para $i = 1, 2$,

$$\frac{\partial p(x_{t-1}, x_t)}{\partial \alpha_i} = I_{t-1,i} \frac{x_{t-1}}{1 - \alpha_i} \{ p(x_{t-1} - 1, x_t - 1, \alpha_i, \lambda) - p(x_{t-1}, x_t, \alpha_i, \lambda) \}. \tag{3.10}$$

A partir das expressões (3.9) e (3.10) as derivadas de primeira ordem do logaritmo da função de transição são

$$\frac{\partial \log p(x_{t-1}, x_t)}{\partial \lambda} = \sum_{i=1}^2 I_{t-1,i} \frac{\partial}{\partial \lambda} \log p(x_{t-1}, x_t, \alpha_i, \lambda) \quad (3.11)$$

$$\frac{\partial \log p(x_{t-1}, x_t)}{\partial \alpha_i} = I_{t-1,i} \frac{\partial}{\partial \alpha_i} \log p(x_{t-1}, x_t, \alpha_i, \lambda), \quad i = 1, 2. \quad (3.12)$$

Mostra-se que o determinante da matriz de informação de Fisher $I(\boldsymbol{\theta})$ é dado por

$$|I(\boldsymbol{\theta})| = \sum_{i=1}^2 E\left[\left(\frac{\partial}{\partial \alpha_i} \log p(X_1, X_2, \alpha_i, \lambda)\right)^2 I_{X_1,i}\right] \times |A_i|,$$

onde, para $i = 1, 2$,

$$(A_i)_{11} = E\left[\left(\frac{\partial}{\partial \alpha_{3-i}} \log p(X_1, X_2, \alpha_{3-i}, \lambda)\right)^2 I_{X_1,3-i}\right];$$

$$(A_i)_{12} = (A_i)_{21} = E\left[\frac{\partial}{\partial \alpha_{3-i}} \log p(X_1, X_2, \alpha_{3-i}, \lambda) \frac{\partial}{\partial \lambda} \log p(X_1, X_2, \alpha_{3-i}, \lambda) I_{X_1,3-i}\right]; \quad e$$

$$(A_i)_{22} = E\left[\left(\frac{\partial}{\partial \lambda} \log p(X_1, X_2, \alpha_{3-i}, \lambda)\right)^2 I_{X_1,3-i}\right],$$

$$I_{X_1,1} = \begin{cases} 1, & X_1 \leq R \\ 0, & X_1 > R \end{cases} \quad e \quad I_{X_1,2} = 1 - I_{X_1,1}.$$

Cada uma das matrizes A_i , $i = 1, 2$, tem a mesma estrutura que a matriz de informação de Fisher analisado por Frank e Seligman (1993), portanto os mesmos argumentos podem ser usados para mostrar que A_i tem determinante positivo, permitindo concluir que a matriz de informação de Fisher é não singular, ou seja a condição (C6) também é válida para este modelo.

As equações (3.9)-(3.12) permitem concluir que, para cada um dos regimes do modelo, se está dentro da estrutura INAR considerada por Franke e Seligman (1993). Eles provaram que as condições (C1)-(C6) em conjunto com um resultado semelhante ao da Proposição 3.1 implicam as condições gerais dos Teoremas 2.1 e 2.2 de Billingsley (1961) sobre os estimadores dos parâmetros de processos de Markov.

Finalmente a matriz de informação de Fisher pode ser obtida através do simétrico do valor médio das derivadas parciais de segunda ordem de $l(\boldsymbol{\theta})$ que são

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i \partial \lambda} &= \sum_{t=1}^n -\frac{I_{t-1,i}}{\alpha_i(1-\alpha_i)} \left\{ \frac{p(x_{t-1}, x_t - 1, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} + \lambda \frac{p(x_{t-1}, x_t - 2, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} - \right. \\ &\quad \left. - \lambda \left(\frac{p(x_{t-1}, x_t - 1, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} \right)^2 \right\} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_i^2} &= \frac{1}{\alpha_i^2(1-\alpha_i)^2} \sum_{t=1}^n \left\{ (2\alpha_i - 1)x_t - \alpha_i^2 x_{t-1} + 2(1-\alpha_i)\lambda \frac{p(x_{t-1}, x_t - 1, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} + \right. \\ &\quad \left. + \lambda^2 \frac{p(x_{t-1}, x_t - 2, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} - \lambda^2 \left(\frac{p(x_{t-1}, x_t - 1, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} \right)^2 \right\} I_{t-1,i} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \lambda^2} &= \sum_{t=1}^n \sum_{i=1}^2 \left(\frac{p(x_{t-1}, x_t - 2, \alpha_i, \lambda)}{p(x_{t-1}, x_t, \alpha_i, \lambda)} - \frac{p(x_{t-1}, x_t - 1, \alpha_i, \lambda)^2}{p(x_{t-1}, x_t, \alpha_i, \lambda)^2} \right) I_{t-1,i} \\ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \alpha_1 \partial \alpha_2} &= 0, \quad i = 1, 2. \end{aligned}$$

■

3.3 Estudo de simulação

O objectivo desta secção é ilustrar os resultados teóricos obtidos na Secção 3.2 e analisar o comportamento dos estimadores MQC e MVC para amostras de pequena, média e grande dimensão. O estudo de simulação contempla as seguintes combinações dos α 's e λ : $\alpha_1 = \{0.2, 0.8\}$, $\alpha_2 = \{0.1, 0.65\}$ e $\lambda = \{3, 7\}$. Para cada combinação dos parâmetros o valor de R foi escolhido de modo a que pelo menos 50% das observações estivessem no primeiro regime. Desta forma foram considerados 8 modelos distintos SETINAR(2; 1) que se apresentam na Tabela 3.1.

Na prática o valor do limiar R é desconhecido e tem também de ser estimado. Para cada realização dos modelos criou-se uma grelha que começa no percentil 10 e termina no

percentil 90 e para cada valor de R na grelha aplicou-se cada um dos métodos analisados na secção anterior para estimar os restantes parâmetros. O conjunto das estimativas dos parâmetros para cada um dos métodos é aquele que otimiza a respectiva função que permite obter as estimativas.

Tabela 3.1: Parâmetros dos modelos SETINAR(2;1).

	$\lambda = 3$			$\lambda = 7$			
	α_1	α_2	R	α_1	α_2	R	
A1	0.2	0.1	4	B1	0.2	0.1	8
A2	0.2	0.65	6	B2	0.2	0.65	14
A3	0.8	0.1	9	B3	0.8	0.1	21
A4	0.8	0.65	11	B4	0.8	0.65	27

Para cada um dos modelos foram simuladas realizações de tamanhos $n = 50, 100, 200, 500$ e consideraram-se 1000 réplicas independentes. Com o objectivo de fazer comparação dos diversos métodos de estimação apresentados, apenas são admitidas realizações para as quais os valores estimados dos parâmetros se encontram dentro suporte dos parâmetros, para todos os métodos de estimação, sendo as restantes realizações rejeitadas. Os resultados estão resumidos nas Tabelas 3.2 e 3.3. Uma análise detalhada das tabelas revela a superioridade do método da máxima verosimilhança condicional em termos do enviesamento das estimativas bem como relativamente ao erro quadrático médio (EQM), com especial relevância quando as realizações são de pequena e média dimensão. As Figuras 3.2 e 3.3 contêm as caixas de bigodes do viés das estimativas MQC e MVC para cada um dos modelos.

É de referir que, para $\alpha_1 = 0.2$ o enviesamento das estimativas dos parâmetros é mais disperso do que para $\alpha_1 = 0.8$, independentemente do valor de λ . O enviesamento e a assimetria das estimativas dos parâmetros diminuem à medida que a dimensão das realizações aumenta, o que está de acordo com as propriedades dos estimadores de serem assintoticamente centrados e consistentes.

Tabela 3.2: Média amostral e EQM (entre parêntesis) para os modelos A1, A2, B1 e B2, com $\theta = \{(0.2, 0.1, 3), (0.2, 0.65, 3), (0.2, 0.1, 7), (0.2, 0.65, 7)\}$, respectivamente.

Modelo	Par	N							
		50		100		200		500	
		CLS	CML	CLS	CML	CLS	CML	CLS	CML
A1	α_1	0.413 (0.095)	0.388 (0.067)	0.336 (0.051)	0.322 (0.042)	0.280 (0.031)	0.276 (0.025)	0.222 (0.011)	0.221 (0.010)
	α_2	0.176 (0.020)	0.177 (0.021)	0.141 (0.010)	0.141 (0.011)	0.120 (0.005)	0.121 (0.005)	0.101 (0.002)	0.101 (0.002)
	λ	2.556 (0.464)	2.565 (0.422)	2.762 (0.215)	2.765 (0.201)	2.880 (0.113)	2.873 (0.107)	2.991 (0.049)	2.986 (0.045)
	R	4.261 (1.719)	4.277 (1.527)	4.218 (1.894)	4.233 (1.759)	4.197 (1.839)	4.149 (1.609)	4.222 (1.396)	4.210 (1.258)
	p	0.717		0.717		0.718		0.718	
A2	α_1	0.243 (0.049)	0.249 (0.024)	0.201 (0.018)	0.219 (0.012)	0.193 (0.008)	0.208 (0.006)	0.192 (0.004)	0.198 (0.003)
	α_2	0.585 (0.018)	0.609 (0.014)	0.622 (0.007)	0.635 (0.005)	0.635 (0.003)	0.643 (0.002)	0.643 (0.001)	0.646 (0.001)
	λ	2.905 (0.530)	2.849 (0.288)	3.028 (0.237)	2.946 (0.161)	3.042 (0.123)	2.981 (0.089)	3.038 (0.055)	3.012 (0.040)
	R	6.215 (2.331)	5.902 (0.906)	6.093 (0.459)	6.032 (0.228)	6.009 (0.017)	6.006 (0.010)	6 (0)	6 (0)
	p	0.698		0.673		0.661		0.655	
B1	α_1	0.470 (0.120)	0.428 (0.085)	0.352 (0.052)	0.341 (0.043)	0.263 (0.019)	0.266 (0.018)	0.212 (0.008)	0.215 (0.007)
	α_2	0.212 (0.027)	0.197 (0.020)	0.162 (0.013)	0.160 (0.011)	0.124 (0.005)	0.126 (0.005)	0.102 (0.003)	0.104 (0.002)
	λ	5.630 (3.431)	5.831 (2.462)	6.252 (1.442)	6.293 (1.197)	6.700 (0.579)	6.677 (0.563)	6.965 (0.286)	6.943 (0.254)
	R	8.324 (3.936)	8.319 (3.601)	8.222 (3.396)	8.270 (3.406)	8.315 (2.867)	8.254 (2.752)	8.251 (1.901)	8.192 (1.620)
	p	0.567		0.568		0.569		0.569	
B2	α_1	0.219 (0.033)	0.237 (0.020)	0.191 (0.014)	0.209 (0.010)	0.186 (0.007)	0.200 (0.006)	0.192 (0.003)	0.198 (0.002)
	α_2	0.599 (0.014)	0.616 (0.011)	0.617 (0.007)	0.628 (0.006)	0.632 (0.003)	0.640 (0.002)	0.643 (0.001)	0.646 (0.001)
	λ	7.025 (2.116)	6.821 (1.241)	7.176 (1.079)	6.995 (0.738)	7.142 (0.575)	7.011 (0.426)	7.074 (0.249)	7.024 (0.181)
	R	13.849 (4.013)	13.579 (2.211)	13.916 (0.742)	13.881 (0.659)	14.004 (0.044)	14.008 (0.052)	14.002 (0.004)	14.003 (0.005)
	p	0.673		0.684		0.688		0.688	

Tabela 3.3: Média amostral e EQM (entre parêntesis) das estimativas de θ para os modelos A3, A4, B3 e B4, com $\theta = \{(0.8, 0.1, 3), (0.8, 0.65, 3), (0.8, 0.1, 7), (0.8, 0.65, 7)\}$, respectivamente.

Modelo	Par	N							
		50		100		200		500	
		CLS	CML	CLS	CML	CLS	CML	CLS	CML
A3	α_1	0.820 (0.010)	0.804 (0.005)	0.813 (0.007)	0.805 (0.003)	0.805 (0.005)	0.803 (0.002)	0.800 (0.002)	0.799 (0.001)
	α_2	0.117 (0.005)	0.110 (0.004)	0.106 (0.003)	0.103 (0.002)	0.101 (0.002)	0.101 (0.001)	0.099 (0.001)	0.099 (0.001)
	λ	2.869 (0.388)	2.953 (0.196)	2.933 (0.264)	2.974 (0.116)	2.977 (0.175)	2.985 (0.067)	3.005 (0.074)	3.010 (0.027)
	R	8.998 (0.010)	8.996 (0.008)	9 (0)	9 (0)	9 (0)	9 (0)	9 (0)	9 (0)
	p	0.796		0.796		0.795		0.795	
	A4	α_1	0.804 (0.014)	0.784 (0.006)	0.813 (0.010)	0.799 (0.002)	0.807 (0.006)	0.801 (0.001)	0.799 (0.003)
α_2		0.396 (0.010)	0.387 (0.006)	0.403 (0.006)	0.396 (0.003)	0.402 (0.003)	0.400 (0.001)	0.398 (0.001)	0.400 (0.001)
λ		2.941 (0.878)	3.079 (0.383)	2.900 (0.592)	2.997 (0.164)	2.949 (0.363)	2.987 (0.089)	3.014 (0.172)	2.993 (0.035)
R		10.057 (0.129)	10.026 (0.092)	10.009 (0.009)	10.004 (0.004)	10 (0)	10 (0)	10 (0)	10 (0)
p		0.749		0.747		0.747		0.745	
B3		α_1	0.821 (0.008)	0.814 (0.004)	0.804 (0.006)	0.807 (0.002)	0.799 (0.003)	0.801 (0.001)	0.800 (0.001)
	α_2	0.112 (0.004)	0.109 (0.002)	0.103 (0.003)	0.105 (0.001)	0.098 (0.001)	0.100 (0.001)	0.100 (0.001)	0.100 (3.4e-04)
	λ	6.733 (1.684)	6.808 (0.935)	6.947 (1.213)	6.897 (0.549)	7.025 (0.549)	6.986 (0.269)	6.998 (0.233)	6.988 (0.114)
	R	20.939 (0.071)	20.939 (0.071)	20.999 (0.001)	20.999 (0.001)	21 (0)	21 (0)	21 (0)	21 (0)
	p	0.783		0.783		0.783		0.783	
	B4	α_1	0.797 (0.033)	0.802 (0.004)	0.807 (0.016)	0.799 (0.002)	0.809 (0.006)	0.799 (0.001)	0.801 (0.003)
α_2		0.645 (0.015)	0.643 (0.004)	0.655 (0.008)	0.647 (0.002)	0.655 (0.004)	0.648 (0.001)	0.650 (0.002)	0.650 (3.1e-04)
λ		7.017 (17.078)	6.988 (2.241)	6.765 (8.067)	7.023 (1.047)	6.782 (3.570)	7.019 (0.581)	6.975 (2.003)	6.984 (0.234)
R		28.108 (1.890)	28.091 (0.805)	28.003 (0.681)	28.028 (0.148)	28.008 (0.020)	28.005 (0.013)	28 (0)	28 (0)
p		0.654		0.657		0.657		0.656	

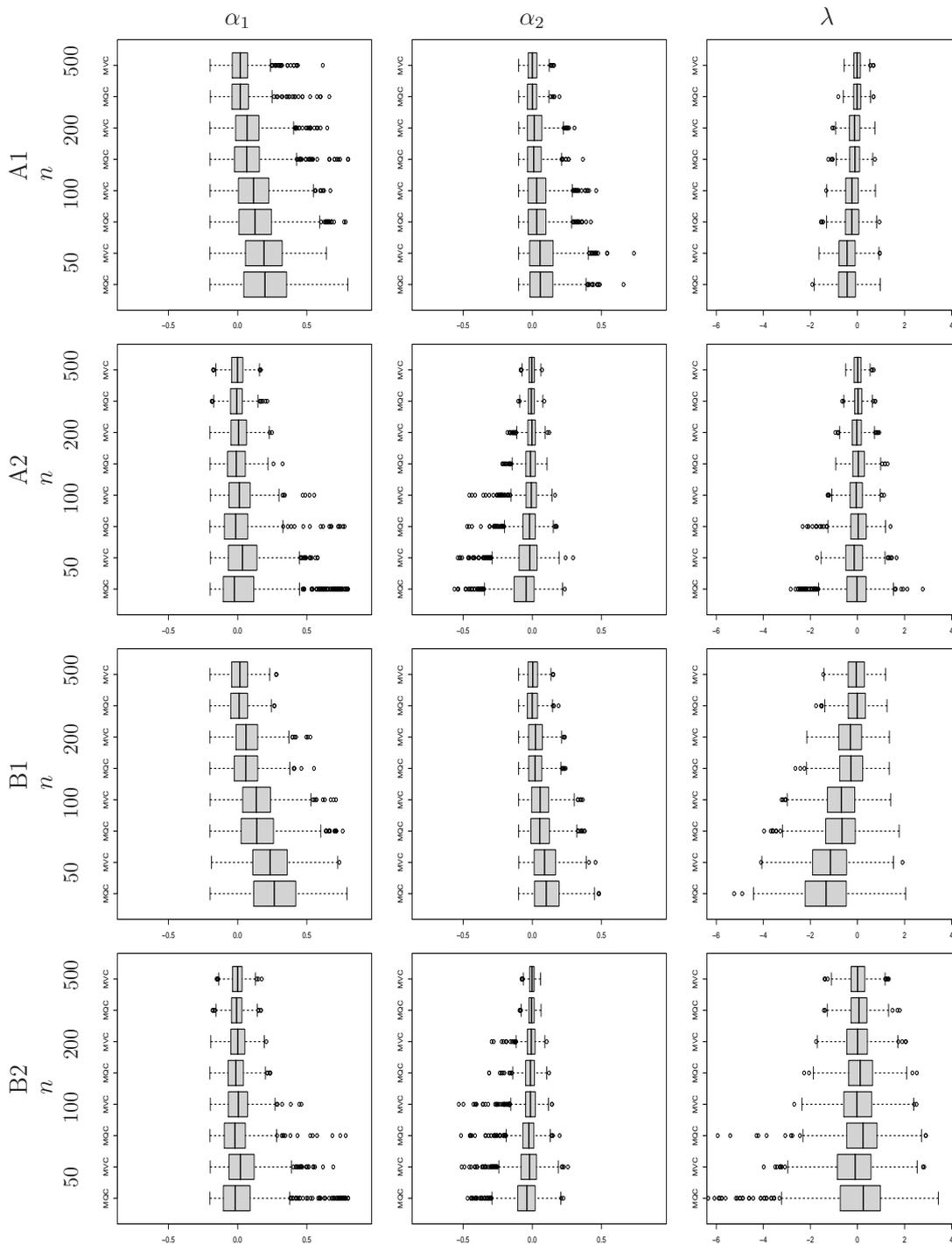


Figura 3.2: Viés das estimativas de θ para A1, A2, B1 e B2, com $\theta = (0.2, 0.1, 3)$, $\theta = (0.2, 0.65, 3)$, $\theta = (0.2, 0.1, 7)$ e $\theta = (0.2, 0.65, 7)$ respectivamente.

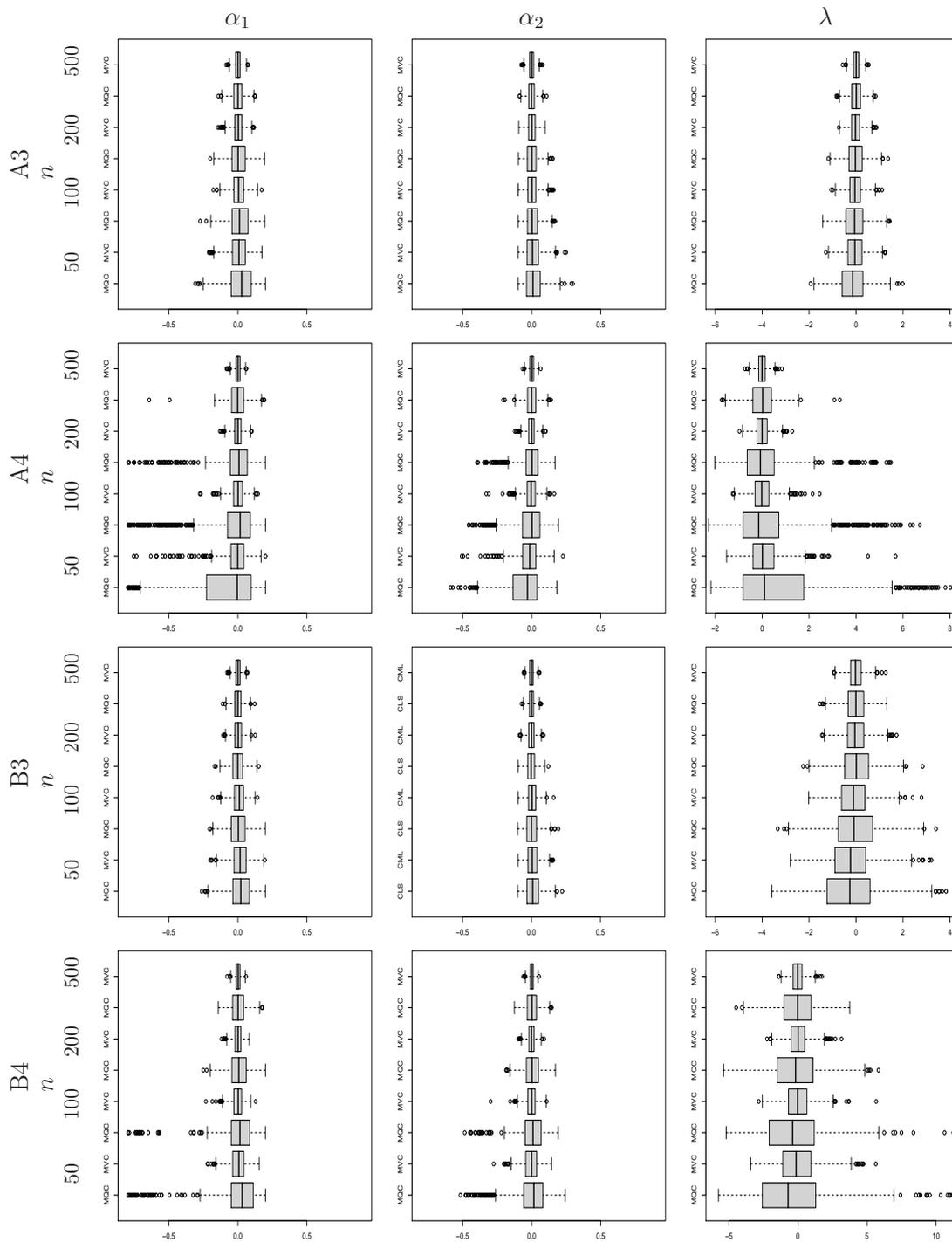


Figura 3.3: Viés das estimativas de θ para A3, A4, B3 e B4, com $\theta = (0.8, 0.1, 3)$, $\theta = (0.8, 0.65, 3)$, $\theta = (0.8, 0.1, 7)$ e $\theta = (0.8, 0.65, 7)$ respectivamente.

O estudo de simulação realizado foi mais vasto do que o apresentado nesta secção e permitiu concluir que, de uma forma geral, quanto maior é a diferença, em valor absoluto, entre os parâmetros α_1 e α_2 melhor é o desempenho dos métodos de estimação. Para os casos em que os valores de α_1 e α_2 estão muito próximos, os métodos de estimação analisados têm um menor desempenho, as estimativas dos parâmetros são mais dispersas e mais assimétricas, uma vez que nestas situações parece mais difícil para os métodos MQC e MVC distinguirem os dois regimes do modelo.

Na Tabela 3.4 encontram-se as percentagens de estimativas do vector de parâmetros que se encontram dentro do espaço de parâmetros para cada um dos métodos analisados e para todos os modelos em consideração. Para realizações de pequena e média dimensão, à excepção dos modelos A4 e B4, as percentagens obtidas para o método de máxima verosimilhança condicional são superiores aos do método dos mínimos quadrados, o que reforça a importância deste método nestas situações.

Tabela 3.4: Percentagens de estimativas do vector de parâmetros $\theta = (\alpha_1, \alpha_2, \lambda)$ que se encontram dentro do espaço dos parâmetros

Modelo	Parâmetros	$N = 50$		$N = 100$		$N = 200$		$N = 500$	
		MQC	MVC	MQC	MVC	MQC	MVC	MQC	MVC
A1	(0.2, 0.1, 3)	61	87	74	92	86	92	98	99
A2	(0.2, 0.65, 3)	73	92	87	96	96	98	100	100
B1	(0.2, 0.1, 7)	63	84	72	85	84	86	96	97
B2	(0.2, 0.65, 7)	62	85	84	99	96	99	100	100
A3	(0.8, 0.1, 3)	75	87	89	94	98	98	100	100
A4	(0.8, 0.65, 3)	88	78	94	83	98	83	100	94
B3	(0.8, 0.1, 7)	87	99	97	100	100	100	100	100
B4	(0.8, 0.65, 7)	89	73	95	83	99	84	100	94

3.4 Conclusões

Neste capítulo foi introduzido a classe de modelos auto-regressivos, de ordem 1, de valores inteiros definido por limiares auto-induzidos cujas inovações são v.a.'s independentes e identicamente distribuídas com distribuição de Poisson. Estabeleceu-se a estacionariedade e a ergodicidade do processo e determinaram-se os estimadores de mínimos quadrados condicionais e de máxima verosimilhança condicional, para os quais se obtiveram as suas propriedades assintóticas.

Muito ainda se encontra por explorar dentro da classe dos modelos auto-regressivos definidos por limiares auto-induzidos. Para o modelo estudado considerou-se que as inovações eram independentes e identicamente distribuídas. A primeira generalização a ser feita será o uso de dois processos de inovações, um para cada um dos troços do modelo. A classe mais geral pode ser definida por SETINARMA($k; p_1, \dots, p_k; q_1, \dots, q_k$), onde k representa o número de regimes e p_i e q_i representam respectivamente as ordens do modelo INARMA no regime i , $i = 1, \dots, k$. Estes tópicos serão abordado em futuros trabalhos de pesquisa.

Capítulo 4

Considerações finais e trabalho futuro

Neste trabalho foram estudados modelos de séries temporais de valores inteiros não negativos que pertencem à classe de modelos baseados no operador *binomial thinning* e suas extensões. O capítulo 1 incidiu na construção de sistemas de alarme *online* ótimos para processos de contagem nas abordagens clássica e bayesiana. O modelo para o qual se desenvolveu tais sistemas foi o modelo DSINAR(1), que é caracterizado por ser um modelo auto-regressivo de valores inteiros não negativos de primeira ordem, em que o parâmetro associado ao operador *binomial thinning* é função de covariáveis de interesse. Foi necessário fazer as devidas adaptações para o caso de valores inteiros não negativos, uma vez que os resultados conhecidos eram todos para processos de valores contínuos. Os critérios de escolha ótima também foram adaptados dadas as especificidades destas séries de contagem. Salienta-se que nos sistemas de alarmes ótimos analisados, em contexto de simulação ou no exemplo apresentado, o número de falsos alarmes produzido foi sempre elevado. Uma possibilidade de diminuir o número de falsos alarmes poderá passar pela inclusão de informação externa adicional no modelo ou colocar o nível da catástrofe como função do tempo. Este tópico será analisado em trabalho futuro.

No Capítulo 2 apresentou-se o modelo periódico auto-regressivo de valores inteiros de or-

dem 1 e período T , $\text{PINAR}(1)_T$, cujas inovações formam uma sucessão periódica de v.a.'s independentes com distribuição de Poisson. Estabeleceu-se a ergodicidade do processo e determinou-se a distribuição ciclo-estacionária do mesmo. Apresentaram-se os métodos de estimação de Yule-walker, mínimos quadrados condicionais simples e ponderados e máxima verosimilhança condicional para os parâmetros do modelo em estudo, e foram estudadas as suas propriedades assintóticas. A realização de um estudo de simulação revelou a superioridade do método da máxima verosimilhança especialmente para amostras de pequena e média dimensão. A predição pontual e intervalar de observações futuras foi também alvo de estudo. Efectuou-se um estudo de simulação onde a moda da distribuição condicional revelou produzir em média melhores resultados que os restantes previsores analisados.

O estudo efectuado apenas se centrou no modelo periódico auto-regressivo de primeira ordem. Muitos são os modelos que pertencem à classe dos modelos periódicos de valores inteiros e que ainda não estão estudados. O primeiro passo será a generalização do modelo $\text{PINAR}(1)_T$ para a ordem p e posteriormente a análise dos processos PINARMA . Estes serão alvos de futuros trabalhos de investigação.

A parcimónia tem sido um dos problemas relacionados com os modelos PARMA , mesmo para o modelo mais simples poder-se-á ter um número muito elevado de parâmetros. No caso dos modelos PINAR este problema também existe. Portanto é necessário desenvolver mecanismos que permitam reduzir a dimensão dos parâmetros nestes modelos. Este assunto poderá ser abordado numa fase posterior.

O Capítulo 3 abordou a classe de modelos auto-regressivos de ordem 1 de valores inteiros definido por limiares auto-induzidos, cujas inovações são v.a.'s independentes e identicamente distribuídas com distribuição de Poisson. Neste capítulo foi estabelecida a ergodicidade do processo e foram apresentados os estimadores, de mínimos quadrados condicionais e de máxima verosimilhança condicional, para os parâmetros do modelo em consideração. Para os estimadores anteriores obtiveram-se as suas propriedades assintóticas. Um estudo de simulação revelou que, em média, os estimadores de máxima verosimilhança condicional

produzem melhores resultados principalmente para amostras de pequena e média dimensão.

A classe dos modelos auto-regressivos definidos por limiares auto-induzidos ainda está por explorar. No modelo em análise, a sucessão das inovações era única, o que permite generalizar o modelo estudado passando a ter duas sucessões de inovações, uma para cada um dos regimes do modelo. O modelo analisado pode ser generalizado não só em termos de números de regimes mas também em termos de tipo de processo em cada regime. Esta classe pode ser definida por $\text{SETINARMA}(k; p_1, \dots, p_k; q_1, \dots, q_k)$, onde k representa o número de regimes e p_i e q_i denotam, respectivamente, as ordens do modelo INARMA no regime i , $i = 1, \dots, k$. Estes tópicos serão abordados em futuros trabalhos de pesquisa.

Um outro trabalho futuro que é necessário efectuar, e que é transversal a todos os capítulos, é desenvolver mecanismos que permitam fazer a identificação dos modelos apresentados e também procedimentos que permitam avaliar a qualidade do ajustamento efectuado. Para os modelos que pertencem à classe dos modelos baseados no operador *binomial thinning* estes assuntos estão pouco explorados na literatura e permanecem em aberto para modelos que pertencem a esta classe, nomeadamente para os que foram aqui apresentados.

Apêndices

Apêndice A

Funções implementadas no programa R

A.1 Funções utilizadas no Capítulo 1

A.1.1 Simulação dos modelos

```
#simula um processo ar1 e um processo desinar 1 pois com inovações poisson
ar1<-function(n,mu,sigma,phi){
n1<-n+500;
y<-0:(n1);
y[1]<-mu;
eps<-rnorm(n1,0,sigma);
for(i in 1:n1){
y[i+1]<-mu+phi*(y[i]-mu)+eps[i]
};
return(y[501:n1])
}
alpha<-function(w,y){
```

```

z<-exp(w*y)/(exp(w*y)+1);
return(z)
}
desinar_pois<-function(n,alfa,lambda){
  n1<-n+500;
  x<-1:n1;
  x[1]<-lambda*sum_media_beta(alfa[501:n1]);
  for (i in 1:(n1-1)){
    x[i+1]<-rpois(1,alfa[i+1]*x[i]+lambda)}
  return(x[501:n1])}

```

A.1.2 Estimação dos parâmetros

método da máxima versosimilhança condicional para o modelo DSNAR

```

mmv<-function(x,cov,a){
  n<-length(x);
  y<-cov[1:n];
  X_t<-x[2:n];
  X_t_1<-x[1:(n-1)];
  p_mq<-mqnl(x,cov,a);
  w<-p_mq[2];
  lambda<-p_mq[1];
  repeat{
    alfa<-exp(w*y)/(1+exp(w*y));
    dalfa_dw<-y*exp(w*y)/(1+exp(w*y))^ 2;
    dalfa_dw2<-y^ 2*exp(w*y)*(1-exp(w*y))/(1+exp(w*y))^ 3;
    s1<-rep(0,times=n);

```

```

s2<-rep(0,times=n);
ds2_dlb<-rep(0,times=n);
s3<-rep(0,times=n);
ds3_dlb<-rep(0,times=n);
s4<-rep(0,times=n);
J<-array(0,dim=c(2,2));
L<-array(0,dim=c(2,1));
for (t in 2:n){
  for(m in 0:min(x[t-1],x[t])){
    s1[t]<-s1[t]+alfa[t]^ m*(1-alfa[t])^ (x[t-1]-m)*lambda^ (x[t]-m)/
(factorial(m)*factorial(x[t-1]-m)*factorial(x[t]-m));
    s2[t]<-s2[t]+lambda^ (x[t]-m)/(factorial(m)*factorial(x[t-1]-m)*
factorial(x[t]-m))*(m*alfa[t]^ (m-1)*(1-alfa[t])^ (x[t-1]-m)-alfa[t]^ m*(x[t-1]-m)*(1-alfa[t])^
(x[t-1]-m-1));
    s3[t]<-s3[t]+alfa[t]^ m*(1-alfa[t])^ (x[t-1]-m)*(x[t]-m)*
lambda^ (x[t]-m-1)/(factorial(m)*factorial(x[t-1]-m)*factorial(x[t]-m));
    s4[t]<-s4[t]+lambda^ (x[t]-m)/(factorial(m)*factorial(x[t-1]-m))/
factorial(x[t]-m)*(m*(m-1)*alfa[t]^ (m-2)*(1-alfa[t])^ (x[t-1]-m)-2*m*alfa[t]^ (m-1)*(x[t-1]-
m)*(1-alfa[t])^ (x[t-1]-m-1)+alfa[t]^ m*(x[t-1]-m)*(x[t-1]-m-1)*(1-alfa[t])^ (x[t-1]-m-2));
    ds2_dlb[t]<-ds2_dlb[t]+(x[t]-m)*lambda^ (x[t]-m-1)/(factorial(m)*
factorial(x[t-1]-m)*factorial(x[t]-m))*(m*alfa[t]^ (m-1)*(1-alfa[t])^ (x[t-1]-m)-alfa[t]^ m*
(x[t-1]-m)*(1-alfa[t])^ (x[t-1]-m-1));
    ds3_dlb[t]<-ds3_dlb[t]+alfa[t]^ m*(1-alfa[t])^ (x[t-1]-m)*(x[t]-m)*
(x[t]-m-1)*lambda^ (x[t]-m-2)/(factorial(m)*factorial(x[t-1]-m)*factorial(x[t]-m));
  }
}
s1<-s1[2:n];
s2<-s2[2:n];
s3<-s3[2:n];

```

```

s4<-s4[2:n];
ds2_dlb<-ds2_dlb[2:n];
ds3_dlb<-ds3_dlb[2:n];
dalfa_dw<-dalfa_dw[2:n];
dalfa_dw2<-dalfa_dw2[2:n];
J[1,1]<-sum((ds3_dlb*s1-s3^ 2)/s1^ 2);
J[2,1]<-sum(dalfa_dw*(ds2_dlb/s1-s3*s2/s1^ 2));
J[1,2]<-J[2,1];
J[2,2]<-sum((dalfa_dw^ 2*s4+s2*dalfa_dw2)/s1-(s2*dalfa_dw/s1)^ 2);
f1<- -(n-1)+sum(s3/s1);
f2<-sum(dalfa_dw*s2/s1);
L[1,1]<- f1;
L[2,1]<- f2;
p<-solve(J,-L);
lambda<-lambda+p[1];
w<-w+p[2];
if(abs(p[1])<a){if(abs(p[2])<a){break}}
}
p_mv<-p;
p_mv[1]<-lambda;
p_mv[2]<-w;
return(p_mv);
}
# Estimação bayesiana
# modelo em txt
model {
  for (i in 1:(N-1)){
    zeros[i]<-0
    zeros[i] dpois(phi[i])
  }
}

```

```

soma<-0;
for (j in 0:min(x[i],x[i+1])){
  soma<-soma+log(choose(x[i],j)*pow(exp(omega*y[i+1])/
(1+exp(omega*y[i+1])),j)*pow(1-exp(omega*y[i+1])/(1+exp(omega*y[i+1])),x[i]-j)*
exp(-lambda)*pow(lambda,x[i+1]-j)/2}
  phi[i]<-soma }
  #-ln(verosimilhança da parcela i)
  omega dnorm(0,0.0001)
  lambda dgamma(0.0001,0.0001)}

mat_des_binom<-function(x){
  n<-length(x);
  Mn<-1:(n-1);
  for(i in 1:(n-1)){Mn[i]<-min(x[i],x[i+1])}
  c<-array(0,dim=c(length(x)-1,max(Mn)+1));
  for(i in 1:(n-1)){
    for(j in 1:(Mn[i]+1)){
      c[i,j]<-factorial(x[i])/(factorial(j-1)*factorial(x[i]-(j-1))*factorial(x[i+1]-(j-1)))}};
  Max.Mn<-max(Mn);
  return(Mn,c)}

binom_winbugs<-function(x,y,w_mq,l_mq){
  if(is.null(dim(x))){x<-array(x,dim=c(length(x),1))}
  c<-dim(x);
  M=c[1];
  L=5000;
  bin<-30000;
  espaco<-20;

```

```

nit<-L*espaco+bin;
sumario<-array(0,dim=c(3,7,c[2]));
am_w<-array(0,dim=c(L,c[2]));
am_l<-array(0,dim=c(L,c[2]));
desvio<-array(0,dim=c(L,c[2]));
coda_w<-array(0,dim=c(L*espaco,c[2]));
coda_l<-array(0,dim=c(L*espaco,c[2]));
coda_desv<-array(0,dim=c(L*espaco,c[2]));
PD_DIC<-array(0,dim=c(2,c[2]));
for(i in 1:c[2]){
  m<-mat_des_binom(x[i]);
  M.mn<-max(m$Mn);
  C<-m$c;
  dados<-list(x=x[i],y=y,N=M,Max.Mn=M.mn,F=C);
  par<-c("omega","lambda");
  inits<-function(){list(omega=w_mq[i],lambda=l_mq[i])};
  a<-bugs(dados,inits,par,"binom_th.txt",bugs.directory =
"c:/Programas/WinBUGS14/",n.chains=1,n.iter=nit,n.burnin=bin,n.thin=1);
  sumario[,i]<-a$summary;
  attach.bugs(a);
  coda_w[,i]<-omega;
  coda_l[,i]<-lambda;
  coda_desv[,i]<-deviance;
  PD_DIC[1,i]<-a$pD;
  PD_DIC[2,i]<-a$DIC;
  detach.bugs()}
i<-1:L;
am_w<-coda_w[1+espaco*(i-1),];
am_l<-coda_l[1+espaco*(i-1),];

```

```
desvio<-coda_desv[1+espaco*(i-1),];  
return(coda_w,coda_l,am_w,am_l,coda_desv,desvio,sumario,PD_DIC)  
}
```

A.1.3 Região de alarme e características operacionais

```
# regiões de alarme abordagem clássica  
# Esta função é função auxiliar G com parametros a,b,p,lambda  
  
funcao_g<-function(x0,x1,p,lambda){  
  aux<-0:min(x0,x1);  
  p<-dbinom(aux,x0,p)*dpois(x1-aux,lambda);  
  s<-sum(p);  
  return(s)}  
  
funcao_beta<-function(alpha,t,j){  
  # j indica o no de factores a multiplicar  
  if(j==0){  
    beta<-1}  
  else {  
    h<-t+j-1;  
    if(h>length(alpha)) stop("t+j-1 ultrapassa o tamanho do vector!")  
    aux<-alpha[(t-j+1):t];  
    beta<-prod(aux)};  
  return(beta)}
```

```

# ver pag 5 a probabilidade de C_t,1 dado D_t e x_t

#P(C_t,1|x_t,D_t) a um passo
probcats_pres<-function(u,xt,alphan_1,lambda){
  if(xt>u){prob<-0}
  else{aux2<-0;u;
  for (k in 0:u)
    {aux2[k+1]<-funcao_g(xt,k,alphan_1,lambda)};
  s<-sum(aux2);
  prob<-1-s};
  return(prob)}

# ver pag 5 a probabilidade de C_t,1 dado D_t
#P(C_t,1|D_t)
probcats_pass<-function(u,xt_1,alphan,alphan_1,lambda){
  au<-0;u;
  for(x_t in 0:u){
    au[x_t+1]<-funcao_g(xt_1,x_t,alphan,lambda)*
*probcats_pres(u,x_t,alphan_1,lambda)};
  pb<-sum(au);
  return(pb)}

#P(C_t,j|x_t,D_t)
probcats_pres_j<-function(u,j,x_t,t,alpha,lambda){
  b<-funcao_beta(alpha,t+j-1,j-1);
  aux0<-0:(j-2);
  for(z in 0:(j-2)){

```

```

    aux0[z+1]<-funcao_beta(alpha,t+j-1,z)};
s<-sum(aux0);
aux1<-0:u;
for(xtj_1 in 0:u){
    aux1[xtj_1+1]<-funcao_g(x_t,xtj_1,b,lambda*s)*
probcats_pres(u,xtj_1,alpha[t+j],lambda)};
pbj<-sum(aux1);
return(pbj)}

```

```
#P(C_t,j|D_t)
```

```

probcats_pass_j<-function(u,j,x_t_1,t,alpha,lambda){
    b<-funcao_beta(alpha,t+j-1,j);
    aux3<-0:(j-1);
    for(i in 0:(j-1)){
        aux3[i+1]<-funcao_beta(alpha,t+j-1,i)};
    s<-sum(aux3);
    aux<-0:u;
    for(xtj_1 in 0:u){
        aux[xtj_1+1]<-funcao_g(x_t_1,xtj_1,b,lambda*s)*
probcats_pres(u,xtj_1,alpha[t+j],lambda)};
    pbpj<-sum(aux);
    return(pbpj)}

```

```
regiao_1j<-function(u,p,k){
```

```

    n<-length(p);
    at1<-0:(n-1);
    for(xt in 0:(n-1)){
        if(p[xt+1]<k){
```

```

    at1[xt+1]<-NA}
else{
    at1[xt+1]<-xt}}
at1<-at1[!is.na(at1)];return(at1)}

```

#probabilidade da região de alarme

```

prob_reg<-function(at1,xt_1,alphat,lambda){
  s<-0;
  for(i in 1:length(at1)){
    s<-s+funcao_g(xt_1,at1[i],alphat,lambda)};
  return(s)}

```

#características operacionais

```

caracteristicas_op<-function(u,j,t,atj,xt_1,alfa,lambda,prob_cat){
  n<-length(atj);
  num<-1:n;
  cop<-array(0,dim=c(1,7));
  pra<-prob_reg(atj,xt_1,alfa[t],lambda);
  for (i in 1:n){
    num[i]<-funcao_g(xt_1,atj[i],alfa[t],lambda)*probcats_pres_j(u,j,atj[i],t,alfa,lambda)};
  p_int<-sum(num);
  pdet<-p_int/pra;
  pac<-p_int/prob_cat;
  pnd<-prob_cat*(1-pac)/(1-pra);
  pfa<-pra*(1-pdet)/(1-prob_cat);
  cop<-array(0,dim=c(1,7));
  cop[1,1]<-prob_cat;
  cop[1,2]<-pra;

```

```
cop[1,3]<-p_int;
cop[1,4]<-pdet;
cop[1,5]<-pac;
cop[1,6]<-pnd;
cop[1,7]<-pfa;

return(cop)}

#regiao optima classica para uma amostra
regiao_opj<-function(u,j,lambda,xt_1,alfa,t){
  h=200;
  ra<-array(NA,dim=c(h,u+22));
  co<-array(0,dim=c(h,8));
  m_zero<-array(0,dim=c(1,7));
  p<-0:(u+20);
  p_catast<-probcats_pass_j(u,j,xt_1,t,alfa,lambda);
  k<-round(p_catast,3);
  i<-seq(length=h,from=k,by=0.001);
  for(xt in 0:(u+20)){
    p[xt+1]<-probcats_pres_j(u,j,xt,t,alfa,lambda)};
  for(m in 1:h){
    aux<-regiao_1j(u,p,i[m]);
    co[m,1]<-i[m];
    ra[m,1]<-i[m];
    tam<-length(aux);
    print("tamanho regiao");
    print(tam);
    if(tam==0){
```

```

    break}
else{
  for(l in 1:tam){
    ra[m,l+1]<-aux[l];
    if(m==1){
      co[m,2:8]<-caracteristicas_op(u,j,t,aux,xt_1,alfa,lambda,p_catast);
    }
    else{
      if(tam==length(aux2)){
        if(all.equal(aux,aux2)){
          co[m,2:8]<-co[m-1,2:8]
        }
        else{
          co[m,2:8]<-caracteristicas_op(u,j,t,aux,xt_1,alfa,lambda,p_catast)
        }
      }
      else{
        co[m,2:8]<-caracteristicas_op(u,j,t,aux,xt_1,alfa,lambda,p_catast)
      }
    }
  }
  aux2<-aux;
}
}
return(co,ra)}

```

####Parte bayesiana para DSINAR com covariáveis #####

#Regiões de alarme metodologia bayesiana

```

#Probabilidade
#P(C_t,j|x_t,D_t)=1/m*soma(i in 1:m){P(C_tj|x_t,D_t,alfa(wi),lambdai)}

p_cat_pres_b_cov<-function(u,xt,t,j,y,am_w,am_lambda){
n <-length(am_lambda);
  probab<-1:n;
  for(i in 1:n){
    probab[i]<-probcats_pres_j(u,j,xt,t,exp(am_w[i]*y)/
(1+exp(am_w[i]*y)),am_lambda[i])
    media<-mean(probab);
  }
  return(media);
}

p_cat_pass_b_cov<-function(u,xt_1,t,j,y,am_w,am_lambda){
n<-length(am_lambda);
  probab<-1:n;
  for(i in 1:n){
    probab[i]<-probcats_pass_j(u,j,xt_1,t,exp(am_w[i]*y)/
(1+exp(am_w[i]*y)),am_lambda[i])
  }

  return(probab)
}

#região de alarme bayesiana

```

```
regiao_j_b<-function(u,p,k){
  atj<-0:u;
  for(xt in 0:u){
    if(p[xt+1]<k){
      atj[xt+1]<-NA}
    else{
      atj[xt+1]<-xt
    }
  }
};
atj<-atj[!is.na(atj)];return(atj)}
```

#Características operacionais metodologia bayesiana

```
cop_b_cov<-function(u,j,t,atj,xt_1,y,am_w,am_lambda,prob_cat){
  n=length(am_lambda);
  car_b<-array(0,dim=c(n,7));
  cob<-array(0,dim=c(1,7));
  for(i in 1:n){
    car_b[i,]<-caracteristicas_op(u,j,t,atj,xt_1,exp(am_w[i]*y)/
(1+exp(am_w[i]*y)),am_lambda[i],prob_cat[i])}
  cob[1,1]<-mean(car_b[,1]);
  cob[1,2]<-mean(car_b[,2]);
  cob[1,3]<-mean(car_b[,3]);
  cob[1,4]<-mean(car_b[,4]);
  cob[1,5]<-mean(car_b[,5]);
  cob[1,6]<-mean(car_b[,6]);
  cob[1,7]<-mean(car_b[,7]);
```

```

    return(cob);
}

reg_opjb_cov<-function(u,t,j,xt_1,y,am_w,am_lambda){
  h=200;
  ra<-array(NA,dim=c(h,u+22));
  co<-array(0,dim=c(h,8));
  p<-0:(u+20);
  p_catast<-p_cat_pass_b_cov(u,xt_1,t,j,y,am_w,am_lambda);
  k<-round(mean(p_catast),3);
  i<-seq(length=h,from=k-0.005,by=0.001);
  for(xt in 0:(u+20)){
    p[xt+1]<-p_cat_pres_b_cov(u,xt,t,j,y,am_w,am_lambda)};
  for(m in 1:h){
    co[m,1]<-i[m];
    ra[m,1]<-i[m];
    aux<-regiao_1j(u,p,i[m]);
    print(aux)
    tam<-length(aux);
    if(tam==0){
      break}
    else{
      for(l in 1:tam){
        ra[m,l+1]<-aux[l]
      }
      if(m==1){
        co[m,2:8]<-cop_b_cov(u,j,t,aux,xt_1,y,am_w,am_lambda,p_catast);
      }
    }
  }
}

```

```

else{
  if(tam==length(aux2)){
    if(all.equal(aux,aux2)){
      co[m,2:8]<-co[m-1,2:8]
    }
    else{
      co[m,2:8]<-cop_b_cov(u,j,t,aux,xt_1,y,am_w,am_lambda,p_catast)
    }
  }
  else{
    co[m,2:8]<-cop_b_cov(u,j,t,aux,xt_1,y,am_w,am_lambda,p_catast)
  }
}
aux2<-aux;
}
}
return(co,ra)}

```

```

r_alarm_ma_b_cov<-function(u,t,j,ams_w,am_lambda,amostrasx,y){
  c<-dim(amostrasx);
  if(is.null(c)){
    c<-0:1;
    amostrasx<-array(amostrasx,dim=c(length(amostrasx),1));
    am_lambda<-array(am_lambda,dim=c(length(am_lambda),1));
    ams_w<-array(ams_w,dim=c(length(ams_w),1))
  }
  h=200;

```

```

c_op<-array(0,dim=c(h,8,c[2]));
reg<-array(0,dim=c(h,u+22,c[2]));
for (i in 1:c[2]){
  r<-reg_opjb_cov(u,t,j,amostrasx[t-1,i],y,ams_w[i],am_lambda[i]);
  c_op[,i]<-r$co;
  reg[,i]<-r$ra}
return(reg,c_op)}

```

A.2 Funções utilizadas no Capítulo 2

A.2.1 Simulação do modelo

PINAR_1.R

Simulação de uma realização do modelo $\text{PINAR}(1)_T$

```

PINAR<-function(n,T,alpha,lambda){
  #alpha e lambda são vectores
  n1<-100*s+n;
  x<-rep(0,n1+1);
  # a primeira observação é o x0
  x[1]<-round(lambda[T],0);
  m<-100+trunc(n/T)+1;
  aux<-rep(lambda,100+trunc(n/T)+1);
  aux_l<-rep(0,T*m+1);
  aux_l[2:(T*m+1)]<-aux
  #print(aux_l)
  aux<-rep(alpha,100+trunc(n/T)+1);

```

```

aux_a<-rep(0,T*m+1);
aux_a[2:(T*m+1)]<-aux
for(i in 2:(n1+1)){
  x[i]<-rbinom(1,x[i-1],aux_a[i])+rpois(1,aux_l[i]);
}
#print(x);
#os primeiros 100 períodos são eliminados para retirar o efeito da observação inicial
return(x[(100*T+1):(n1+1)])
}

```

A.2.2 Estimação dos parâmetros do modelo

ESTIMACAO.R

Este ficheiro tem todos os métodos de estimação para uma realização do processo

#método de estimação Yule-Walker

```

est_y_w<-function(x,T){
  n<-length(x)
  m<-trunc(n/T)
  #print(m)
  l<-0
  if(!m==n/T){
    a<-x[(m*T+1):n]
    l<-length(a)
  }
  v<-rep(0,T)

```

```
mu<-rep(0,T)
gama_1<-rep(0,T)
#print(l)
for( i in 1:T){
  if(i<(l+1)){
    v[i]<-var(x[i+T*(0:m)])
    mu[i]<-mean(x[i+T*(0:m)])
    if(!i==l){
      gama_1[i]<-cov(x[i+T*(0:m)],x[i+1+T*(0:m)])
    }
  }
  else{
    gama_1[i]<-cov(x[i+T*(0:(m-1))],x[i+1+T*(0:(m-1))])
  }
}
else{
  v[i]<-var(x[i+T*(0:(m-1))])
  mu[i]<-mean(x[i+T*(0:(m-1))])
  if(!i==T){
    gama_1[i]<-cov(x[i+T*(0:(m-1))],x[i+1+T*(0:(m-1))])
  }
  else{
    if(l==0){
      gama_1[i]<-cov(x[i+T*(0:(m-2))],x[i+1+T*(0:(m-2))])
    }
    else{
      gama_1[i]<-cov(x[i+T*(0:(m-1))],x[i+1+T*(0:(m-1))])
    }
  }
}
}
```

```
}
lambda<-rep(0,T)
alfa<-rep(0,T)
alfa[1]<-gama_1[s]/v[T]
alfa[2:T]<-gama_1[1:(T-1)]/v[1:(T-1)]
mat<-array(0,dim=c(T,T))
mat[T,T]<-1
for( j in 1:(T-1)){
  mat[j,j]<-1;
  for(i in (j+1):T){
    mat[i,j]<-prod(alfa[(j+1):i])
    mat[j,i]<-prod(alfa[-(i:(j+1))])
  }
}
L<-(1-prod(alfa))*mu
lambda<-solve(mat,L)
if(!is.na(all(alfa))){
  if(!all(alfa>0)){      alfa<-rep(NA,T);
    lambda<-rep(NA,T);
  }
  else{
    if(!all(alfa<1)){
      alfa<-rep(NA,T)
      lambda<-rep(NA,T)
    }
    else{
      if(!all(lambda>0)){
        alfa<-rep(NA,T)
        lambda<-rep(NA,T)
      }
    }
  }
}
```

```

    }
  }
}
}
else{
  alfa<-rep(NA,T)
  lambda<-rep(NA,T)
}
return(alfa,lambda)
}

```

```
#####
```

```
# Método dos mínimos quadrados condicional
```

```
pinar_mq1<-function(x,T){
```

```
  n<-length(x);
```

```
  m<-trunc(n/T);
```

```
  #print(m)
```

```
  alfa<-rep(0,T)
```

```
  lambda<-rep(0,T)
```

```
  # as variaveis aux e aux1 servem para que a 1a observação não apareça como observação
```

```
x_t apenas como x_t_1
```

```
  aux<-rep(0,n+1);
```

```
  aux1<-rep(0,n);
```

```
  aux[2:(n+1)]<-x
```

```
  aux1[2:n]<-x[2:n]
```

```
  ni<-rep(m,T)
```

```
  ni[1]<-m-1
```

```
  parcelas<-array(0,dim=c(T,4))
```

```

for(j in 1:T){
  parcelas[j,3]<-sum(aux[j+T*(0:(m-1))]^ 2);
  parcelas[j,4]<-sum(aux[j+T*(0:(m-1))]);
  parcelas[j,1]<-sum(aux[j+1+T*(0:(m-1))]*aux[j+T*(0:(m-1))]);
  parcelas[j,2]<-sum(aux1[j+T*(0:(m-1))]);
}

if(!(n/T==m)){
  a<-x[(T*m+1):n];
  l<-length(a);
  for(j in 1:l){
    ni[j]<-ni[j]+1
    parcelas[j,3]<-parcelas[j,3]+x[j+T*m-1]^ 2;
    parcelas[j,4]<-parcelas[j,4]+x[j+T*m-1];
    parcelas[j,1]<-parcelas[j,1]+x[j+T*m]*x[j+T*m-1];
    parcelas[j,2]<-parcelas[j,2]+x[j+T*m];
  }
}

for (j in 1:T){
  alfa[j]<-(ni[j]*parcelas[j,1]-parcelas[j,2]*parcelas[j,4])/
  (ni[j]*parcelas[j,3]-parcelas[j,4]^ 2)
  lambda[j]<-(parcelas[j,2]-alfa[j]*parcelas[j,4])/ni[j]
}

if(!is.na(all(alfa))){
  if(!all(alfa>0)){
    alfa<-rep(NA,T);
    lambda<-rep(NA,T);
  }
  else{

```

```

    if(!all(alfa<1)){
      alfa<-rep(NA,T)
      lambda<-rep(NA,T)
    }
    else{
      if(!all(lambda>0)){
        alfa<-rep(NA,T)
        lambda<-rep(NA,T)
      }
    }
  }
}
else{
  alfa<-rep(NA,T)
  lambda<-rep(NA,T)
}
return(alfa,lambda)
}
#####
# Método dos mínimos quadrados ponderados
pinar_mq_pond<-function(x,T,a_mq,l_mq){
  n<-length(x);
  v<-rep(1,n)
  m<-trunc(n/T);
  alfa<-1:T
  lambda<-1:T
  # as variaveis aux e aux1 servem para que a 1ª observação não apareça como observa-
  ção x_t apenas como x_t_1
  aux<-rep(0,n+1);

```

```

aux1<-rep(0,n);
aux[2:(n+1)]<-x
aux1[2:n]<-x[2:n]
ni<-rep(m,T)
ni[1]<-m-1
if(!(n/T==m)){
  a<-x[(T*m+1):n];
  l<-length(a);
}
else{
  l=0
}
for(i in 1:T){
  if(i==1){
    v[i+T*(1:(m-1))]<-a_mq[i]*(1-a_mq[i])*x[i-1+T*(1:(m-1))]+l_mq[i];
  }
  else{
    v[i+T*(0:(m-1))]<-a_mq[i]*(1-a_mq[i])*x[i+T*(0:(m-1))]+l_mq[i];
  }
}
parcelas<-array(0,dim=c(T,5))
for(j in 1:T){
  parcelas[j,3]<-sum(aux[j+T*(0:(m-1))]^ 2/v[j+T*(0:(m-1))]);
  parcelas[j,4]<-sum(aux[j+T*(0:(m-1))]/v[j+T*(0:(m-1))]);
  parcelas[j,1]<-sum(aux[j+1+T*(0:(m-1))]^2*aux[j+T*(0:(m-1))]/v[j+T*(0:(m-1))]);
  parcelas[j,2]<-sum(aux1[j+T*(0:(m-1))]/v[j+T*(0:(m-1))])
  if(j==1){
    #o 1º valor n pode entrar na soma uma vez que é uma parcela ficticia
    parcelas[j,5]<-sum(1/v[j+T*(1:(m-1))])
  }
}

```

```

    }
    else{
        parcelas[j,5]<-sum(1/v[j+T*(0:(m-1))])
    }
}

if(!l==0){
    for(j in 1:l){
        v[j+T*m]<-a_mq[j]*(1-a_mq[j])*x[j-1+T*m]+l_mq[j]
        parcelas[j,3]<-parcelas[j,3]+x[j+T*m-1]^ 2/v[j+T*m];
        parcelas[j,4]<-parcelas[j,4]+x[j+T*m-1]/v[j+T*m];
        parcelas[j,1]<-parcelas[j,1]+x[j+T*m]*x[j+T*m-1]/v[j+T*m];
        parcelas[j,2]<-parcelas[j,2]+x[j+T*m]/v[j+T*m];
        parcelas[j,5]<-parcelas[j,5]+1/v[j+T*m]
    }
}

for (j in 1:T){
    alfa[j]<-(parcelas[j,5]*parcelas[j,1]-parcelas[j,2]*parcelas[j,4])/
        (parcelas[j,5]*parcelas[j,3]-parcelas[j,4]^ 2)
    lambda[j]<-(parcelas[j,2]-alfa[j]*parcelas[j,4])/parcelas[j,5]
}

if(!is.na(all(alfa))){
    if(!all(alfa>0)){
        alfa<-rep(NA,T);
        lambda<-rep(NA,T);
    }
    else{

```

```

    if(!all(alfa<1)){
      alfa<-rep(NA,T)
      lambda<-rep(NA,T)
    }
    else{
      if(!all(lambda>0)){
        alfa<-rep(NA,T)
        lambda<-rep(NA,T)
      }
    }
  }
}
else{
  alfa<-rep(NA,T)
  lambda<-rep(NA,T)
}
return(alfa,lambda)
}

#####
# As funções abaixo são necessárias para o método da máxima verosimilhança condicional

# cálculo da função de transição
funcao_p<-function(x0,x1,alpha,lambda){
  n1<-length(x1);
  n0<-length(x0);
  if(n0==n1){
    s<-rep(0,n0)
    for(i in 1:n0){

```

```

    aux<-0:min(x0[i],x1[i]);
    p<-dbinom(aux,x0[i],alpha)*dpois(x1[i]-aux,lambda);
    s[i]<-sum(p);
  } } return(s) }
#Esta função serve para calcular o logaritmo da verosimilhança consoante o período i
f_alpha<-function(alpha,i,x,T){
  if(i>s){stop}
  else{
    soma<-0;
    n<-length(x)
    N<-trunc((n-1)/T);
# As 3 expressões seguintes são para garantir que os elementos que não completam um
ciclo também entram para as estimativas
    if(!((n-1)/T==N)){
      a<-x[(s*N+2):n];
      l<-length(a);
      if(!l<i){N<-N+1}
    }
    lambda<-1/N*sum(x[i+1+T*(0:(N-1))]-alpha*x[i+T*(0:(N-1))])
    aux2<-sum(funcao_p(x[i+T*(0:(N-1))],x[i+1+T*(0:(N-1))]-1,alpha,lambda)/
funcao_p(x[i+T*(0:(N-1))],x[i+1+T*(0:(N-1))],alpha,lambda))-N
  }
  return(aux2)
}

# Esta função permite determinar o primeiro intervalo a ser considerado no método da
bissecção
intervalo_f<-function(i,x,T){

```

```

N<-trunc((length(x)-1)/T);
limsup<-min(mean(x[i+1+T*(0:(N-1))])/mean(x[i+T*(0:(N-1))]),1)
#print(limsup)
alpha<-seq(0.0000005,limsup,by=0.005)
m<-length(alpha);
aux1<-array(NA,dim=c(2,2))
f2<-rep(0,m)
for(k in 1:m){
  ##print("k=")
  ##print(k)
  f2[k]<-f_alpha(alpha[k],i,x,T);
}
for(k in 1:(m-1)){
  prod<-f2[k]*f2[k+1];
  if(!is.na(prod)){
    if(prod<0){
      aux1[1,1]<-alpha[k]
      aux1[1,2]<-f2[k]
      aux1[2,1]<-alpha[k+1]
      aux1[2,2]<-f2[k+1]
      break
    }
  }
  else{break}
}
return(aux1)
}

#metodo da bissecção para uma estação

```

```
mv_mb<-function(i,x,T){
  aux<-intervalo_f(i,x,T)
  a0<-aux[1,1];
  #print(aux)
  a<-rep(0,1000);
  b<-rep(0,1000);
  fa<-rep(0,1000);
  fb<-rep(0,1000);
  a[1]<-aux[1,1];
  b[1]<-aux[2,1];
  fa[1]<-aux[1,2];
  fb[1]<-aux[2,2];
  if(!is.na(a0)){
    for(k in 1:999){
      auxm<-mean(c(a[k],b[k]));
      fauxm<-f_alpha(auxm,i,x,T);
      #print(fauxm)
      if(fauxm==0){
        break
      }
      else{
        if(fauxm*fa[k]<0){
          fb[k+1]<-fauxm;
          b[k+1]<-auxm;
          fa[k+1]<-fa[k];
          a[k+1]<-a[k]
        }
        else{
          fa[k+1]<-fauxm;
```

```

        a[k+1]<-auxm;
        fb[k+1]<-fb[k];
        b[k+1]<-b[k]
    }
}
if(abs(fauxm)<5e-6){
    break
}
}
}
else{
    auxm<-NA
}
return(auxm)
}

```

#estimação dos parâmetros pelo método de máxima verossimilhança para o modelo inar periódico de período T

```

p_mv_biss<-function(x,T){
    N<-trunc((length(x)-1)/T);
    alpha<-rep(NA,T);
    lambda<-rep(NA,T);
    f<-array(NA,dim=c(T,2))
    for(i in 1:T){
        alpha[i]<-mv_mb(i,x,T)
        if(is.na(alpha[i])){
            break
        }
    }
}

```

```

lambda[i]<-1/N*sum(x[i+1+T*(0:(N-1))]-alpha[i]*x[i+T*(0:(N-1))]);
f[i,2]<-f_alpha(alpha[i],i,x,T);
f[i,1]<-sum(x[i+1+T*(0:(N-1))]-alpha[i]*x[i+T*(0:(N-1))])-
lambda[i]*sum(funcao_p(x[i+T*(0:(N-1))],x[i+1+T*(0:(N-1))]-1,alpha[i],lambda[i])/
funcao_p(x[i+T*(0:(N-1))],x[i+1+T*(0:(N-1))],alpha[i],lambda[i]))
}
if(any(is.na(alpha))) {
  alpha<-rep(NA,T);
  lambda<-rep(NA,T);
  f<-array(NA,dim=c(T,2))
}
return(alpha,lambda,f)
}

```

A.2.3 Previsão no modelo

Previsão.R

```

#Este tem todas as funções necessárias para a parte de previsão
funcao_beta<-function(alpha,t,j){
#j indica o nº de factores a multiplicar
  if(j==0){
    beta<-1}
  else {
    aux<-alpha[(t-j+1):t];
    beta<-prod(aux);
  }
}

```

```

    return(beta)
}

media_periodica<-function(alpha,lambda,T){
  m<-rep(0,T);
  a<-funcao_beta(alpha,T,T);
  for(i in 1:T){
    soma<-0
    for(k in 0:(i-1)){
      soma<-soma+funcao_beta(alpha,i,k)*lambda[i-k]
    }
    if(!i==T){
      for(k in 0:(T-i-1)){
        soma<-soma+funcao_beta(alpha,i,i)*funcao_beta(alpha,T,k)*lambda[T-k];
      }
    }
    m[i]<-soma/(1-a)
  }
  return(m)
}

#####

#previsao h passos à frente
previsao_h<-function(x,t,alpha,lambda,T,h,X_0=TRUE){
  #X_0 é verd ou Falso se existe a observação X_0
  # a variavel alfa_aux facilita os calculos da função beta pois é a repetição de alfa 2vezes
  #aux_alfa<-rep(alpha,2)

```

```

#aux_lambda<-rep(lambda,2)
mu<-media_periodica(alpha,lambda,T);
max_mu<-max(mu);
k<-trunc(h/T)
j<-h-k*T;
if(j==0){j=T;k<-k-1}
if(X_0){
  n<-trunc((t)/T)
  verd_x<-x[t+1+h]
}
else{
  n<-trunc((t)/T)
  verd_x<-x[t+h]
}
i<-t-n*T;
if(i==0){i=T;n<-n-1}
#print("i:")
b<-trunc((j+i)/T);
a<-j+i-b*T;
if(a==0){a=T;b<-b-1}
beta_TT<-funcao_beta(alpha,T,T)
if(!a<j){
  beta_tmais_h_h<-funcao_beta(alpha,a,j)*beta_TT^ k;
}
else{
  beta_tmais_h_h<-funcao_beta(alpha,a,a)*funcao_beta(alpha,T,j-a)*beta_TT^ k;
}
y<-0:(max_mu+100);
if(X_0){

```

```

media<-beta_tmais_h_h*(x[t+1]-mu[i])+mu[a];
py_cond_x<-funcao_p(x[t+1],y,beta_tmais_h_h,mu[a]-beta_tmais_h_h*mu[i]);
}
else{
media<-beta_tmais_h_h*(x[t]-mu[i])+mu[a];
py_cond_x<-funcao_p(x[t],y,beta_tmais_h_h,mu[a]-beta_tmais_h_h*mu[i]);
}
# mediana e moda da distribuição
fdist<-rep(0,max_mu+101);
moda<-max(py_cond_x)
#print(moda)
for(m in 1:(max_mu+101)){
fdist[m]<-sum(py_cond_x[1:m]);
if(py_cond_x[m]==moda){
#é i-1 pq a posição m tem o elemento m-1 por causa do zero
moda<-m-1
}
if(!fdist[1]<0.5){
mediana<-0
}
else{
if(m>1){
if(!(fdist[m]<0.5)){
if(fdist[m-1]<0.5){
mediana<-m-1
}
}
}
}
}
}

```

```

}
return(media,mediana,moda,verd_x)
}

##### h previsões 1 passo à frente#####
prev_h_1passo<-function(x,t,alpha0,lambda0,alpha,lambda,T,h,X_0){
#alpha0 e lambda0 são as estimativas a usar qdo h=1 porque estão numa variável à parte
p_mv_ams
estimativas<-array(0,dim=c(h,4))
aux<-previsao_h(x,t,alpha0,lambda0,T,1,X_0)
#print(aux)
estimativas[1,1]<-aux$verd_x
estimativas[1,2]<-aux$media
estimativas[1,3]<-aux$mediana
estimativas[1,4]<-aux$moda
for(i in 2: h){
aux<-previsao_h(x,t+i-1,alpha[,i-1],lambda[,i-1],T,1,X_0)
estimativas[i,1]<-aux$verd_x
estimativas[i,2]<-aux$media
estimativas[i,3]<-aux$mediana
estimativas[i,4]<-aux$moda
}
return(estimativas)
}

##### predição intervalar#####
mat_erro_prev<-function(t,alpha,lambda,T,h,pi){
#Valores que o e_{n+h} pode tomar em função de x_t e x_{t+h}
#pi=1-gamma

```

```

#vamos replicar o vector alpha e lambda para ser mais directo o calculo de beta_{j+i,i}
aux_alpha<-rep(alpha,2)
aux_lambda<-rep(lambda,2)
mu<-media_periodica(alpha,lambda,T)
#estamos a definir os valores de x e y
m<-round(max(mu),0)+500
#vamos construir a matriz do erro
erro<-array(0,dim=c(m+1,m+1))
prob_erro<-array(0,dim=c(m+1,m+1))
k<-trunc(h/T)
j<-h-k*T;
if(j==0){j=T;k<-k-1}
n<-trunc(t/T)
i<-t-n*T;
if(i==0){i=T;n<-n-1}
b<-trunc((j+i)/T); a<-j+i-b*T;
if(a==0){a=T;b<-b-1}
beta_TT<-prod(alpha);
beta_jii<-funcao_beta(aux_alpha,i+j,i);
for(x in 0:m){
  for(y in 0:m){
    erro[x+1,y+1]<-beta_jii*beta_TT^ k*x+mu[a]-beta_jii*beta_TT^ k*mu[i]-y
    prob_erro[x+1,y+1]<-dpois(x,mu[i])*
      *funcao_p(x,y,beta_jii*beta_TT^ k,mu[a]-beta_jii*beta_TT^ k*mu[i])
  }
}
#l nº de linhas para o qual a soma das probabilidades é 1 a menos de 10^-10
l=0
repeat{

```

```
l<-l+1
soma<-sum(prob_erro[1:l,1:l])
#print(soma)
if((1-soma)<10^ (-10)){
  break
}
}
l1<-l^ 2
matriz_erro<-array(0,dim=c(l1,3))
matriz_erro[,1]<-erro[1:l,1:l];
matriz_erro[,2]<-prob_erro[1:l,1:l];
#vamos ordenar o erro
erro_ord<-sort(matriz_erro[,1],method="sh",index.return=TRUE)
#nesta função sai o erro ordenado e o índice
#vamos substituir a matriz pela propria ja ordenada
matriz_erro[,]<-matriz_erro[erro_ord$ix,];
matriz_erro[1,3]<-matriz_erro[1,2];
for( u in 2: l1){
  # a 3ª coluna da matriz possui a função distribuição
  matriz_erro[u,3]<-matriz_erro[u-1,3]+matriz_erro[u,2];
}
matriz_erro<-round(matriz_erro,10)
return(soma,matriz_erro)
}

inter_conf_erro<-function(matriz_erro,pi){
  d<-dim(matriz_erro)
  l1<-d[1];
```

```

for( u in 2: l1){
  if(!matriz_erro[u-1,3]>pi/2){
    if(matriz_erro[u,3]>pi/2){
      ind_min<-u-1
    }
  }
  else{
    if(!matriz_erro[u,3]<(1-pi/2)){
      if(matriz_erro[u-1,3]<(1-pi/2)){
        ind_max<-u
      }
    }
  }
}

val_int<-array(0,dim=c(2,2))
val_int[1,1]<-matriz_erro[ind_min,1];
val_int[1,2]<-matriz_erro[ind_min,3];
val_int[2,1]<-matriz_erro[ind_max,1];
val_int[2,2]<-matriz_erro[ind_max,3];
return(val_int)
}

#apenas um valor de h
int_conf_prev_h<-function(media_prev,t,alpha,lambda,T,h,pi){
  mat_erro<-mat_erro_prev(t,alpha,lambda,T,h,pi)
  if(abs(mat_erro$soma-1)<10^ (-10)){
    int<-inter_conf_erro(mat_erro$matriz_erro,pi)
  }
  else{#print("a soma das probabilidades não dá 1")
  }
}

```

```
int[1,1]<-media_prev+int[1,1]
int[2,1]<-media_prev+int[2,1]
round(int,10)
int<-t(int)
return(int)
}
```

A.3 Funções usadas no Capítulo 3

A.3.1 Simulação do modelo SETINAR(2;1)

```
#Simulação de uma amostra do modelo threshold
sim_threshold<-function(a1,a2,lambda,R,n){
  x<-rep(0,n+501);
  print(x)
  phi<-rep(0,n+501);
  x[1]<-0;
  print(R)
  for(i in 2:(n+501)){
    print(x[i-1])
    if(x[i-1]>R){
      phi[i]<-a2;
      x[i]<-rbinom(1,x[i-1],a2)+rpois(1,lambda2)
    }
    else{
      x[i]<-rbinom(1,x[i-1],a1)+rpois(1,lambda1)
      phi[i]<-a1;
    }
  }
}
```

```

    }
  }
#deitamos a 500 observações fora
  x<-x[502:(n+501)]
  phi<-phi[502:(n+501)]
  return(x,phi)
}

```

A.3.2 Métodos de estimação

```

# Método dos mínimos quadrados para R fixo
threshold_mq_r<-function(x,r){
  n<-length(x);
  Ind_1<-rep(0,n-1);
  Ind_2<-rep(0,n-1);
  for(i in 1:(n-1)){
    if(x[i]>r){
      Ind_2[i]<-1}
    else{
      Ind_1[i]<-1}
  }
  prop<-sum(Ind_1)/(n-1)
#vamos construir a matriz
  Mat_sist<-array(0,dim=c(3,3));
  Mat_sist[1,1]<-sum(Ind_1*x[1:(n-1)]^ 2);
  Mat_sist[2,2]<-sum(Ind_2*x[1:(n-1)]^ 2);
  Mat_sist[1,3]<-sum(Ind_1*x[1:(n-1)]);

```

```

Mat_sist[2,3]<-sum(Ind_2*x[1:(n-1)]);
Mat_sist[3,1]<-sum(Ind_1*x[1:(n-1)]);
Mat_sist[3,2]<-sum(Ind_2*x[1:(n-1)]);
Mat_sist[3,3]<-n-1;
L<-array(0,dim=c(3,1));
L[1,1]<-sum(Ind_1*x[1:(n-1)]*x[2:n]);
L[2,1]<-sum(Ind_2*x[1:(n-1)]*x[2:n]);
L[3,1]<-sum(x[2:n]);
if(det(Mat_sist)==0){
  p<-rep(0,3)
}
else{
  p<-solve(Mat_sist,L);
}
if(any(p[1:3]<0)){
  p[1:3]<-NA
}
else{
  if(any(p[1:2]>1)){
    p[1:3]<-NA
  }
}
sqe<-sum((x[2:n]-p[1]*Ind_1*x[1:(n-1)]-p[2]*Ind_2*x[1:(n-1)]-p[3])^ 2)
return(p,Tqe,prop)
}

#para uma grelha de vários valores de R
threshold_mq<-function(x){

```

```
rmin<-trunc(quantile(x,0.1));
rmax<-trunc(quantile(x,0.9))+1;
r<-rmin:rmax;
n<-length(r);
parametros_sqe<-array(0,dim=c(5,n))
parametros_sqe[4,]<-r;
aux<-threshold_mq_r(x,r[1]);
parametros_sqe[1:3,1]<-aux$p;
parametros_sqe[5,1]<-aux$sqe;
if(is.na(parametros_sqe[5,1])){
  rmin_mq<-0
}
else{
  rmin_mq<-1
}
for(i in 2:n){
  aux<-threshold_mq_r(x,r[i]);
  parametros_sqe[1:3,i]<-aux$p;
  parametros_sqe[5,i]<-aux$sqe;
  #parametros_sqe[6,i]<-aux$prop;
  if(rmin_mq==0){
    if(!is.na(parametros_sqe[5,i])){
      rmin_mq<-i
    }
  }
}
else{
  if(!is.na(parametros_sqe[5,i])){
    if(parametros_sqe[5,i]<parametros_sqe[5,rmin_mq]){
      rmin_mq<-i
    }
  }
}
```

```

    }
  }
}
}
par<-parametros_sqe[,rmin_mq];
return(par,parametros_sqe,rmin_mq);
}

```

mínimos quadrados para várias amostras para a simulação

```

threshold_mq_ma<-function(am_x,r){
  d<-dim(am_x)
  if(is.null(am_x)){d<-c(length(am_x),1);am_x<-array(am_x,dim=d)}
  am_inv<-array(NA,d[2])
  par_am<-array(0,dim=c(5,d[2]))
  for(i in 1:d[2]){
    aux<-threshold_mq(am_x[,i]);
    #print(aux)
    m<-dim(aux$parametros_sqe);
    c<-0;
    if(any(is.na(aux$par))){
      par_am[,i]<-rep(NA,5);
      am_inv[i]<-i;
    }
    else{
      if(!any(aux$parametros_sqe[4,]==r)){
        par_am[,i]<-rep(NA,5);
        am_inv[i]<-i;
      }
    }
  }
}

```

```

else{
  for(j in 1:m[2]){
    if(aux$parametros_sqe[4,j]==r){
      if(!is.na(aux$parametros_sqe[5,j])){
        par_am[i]<-aux$par
      }
    }
    else{
      par_am[i]<-rep(NA,5);
      am_inv[i]<-i;
    }
  }
}
}
}
}
return(par_am,am_inv)
}

```

para a maxima verosimilhança condicional

função de transição

```
funcao_p_threshold<-function(x0,x1,alpha,beta,lambda){
```

se alpha tem tamanho 1 estamos no inar normal se alpha tem tamanho n1 então
estamos no caso de threshold

```
pb<-alpha+beta;
```

```
#print(pb)
```

```
n1<-length(x1);
```

```
n0<-length(x0);
```

```
n2<-length(pb);
```

```

if(n0==n1){
  if(n1==n2){
    s<-rep(0,n0)
    for(i in 1:n0){
      aux<-0:min(x0[i],x1[i]);
      if(x0[i]<0){
        p<-rep(0,length(aux))
      }
      else{
        p<-dbinom(aux,x0[i],pb[i])*dpois(x1[i]-aux,lambda);
      }
      s[i]<-sum(p);
    }
  }
}
return(s)
}

# gradient para aplicar depois no método de Newton
gradiente_hessiano<-function(alpha,beta,lambda,x,I1,I2){
  n<-length(x)
  grad<-rep(0,3);
  num_1<-funcao_p_threshold(x[1:(n-1)],x[2:n]-1,alpha*I1,beta*I2,lambda)
  num_2<-funcao_p_threshold(x[1:(n-1)],x[2:n]-2,alpha*I1,beta*I2,lambda)
  den<-funcao_p_threshold(x[1:(n-1)],x[2:n],alpha*I1,beta*I2,lambda)
  quociente1<-num_1/den;
  quociente2<-num_2/den;
  grad[1]<-sum(1/(alpha*(1-alpha))*I1*(x[2:n]-alpha*x[1:(n-1)]-lambda*quociente1))
  grad[2]<-sum(1/(beta*(1-beta))*I2*(x[2:n]-beta*x[1:(n-1)]-lambda*quociente1))
  grad[3]<-sum(quociente1)-(n-1)
}

```

```

hess<-array(0,dim=c(3,3))
p1<-sum(((1-2*alpha)/alpha^ 2/(1-alpha)^ 2*I1*(x[2:n]-
  -alpha*x[1:(n-1)]-lambda*quociente1));
p2<-sum(1/alpha/(1-alpha)*I1*(-x[1:(n-1)]-lambda/alpha/(1-alpha)*
  (lambda*quociente1^ 2-quociente1-lambda*quociente2)))
hess[1,1]<-p1+p2
p1<-sum(((1-2*beta)/beta^ 2/(1-beta)^ 2*I2*(x[2:n]-
  beta*x[1:(n-1)]-lambda*quociente1));
p2<-sum(1/beta/(1-beta)*I2*(-x[1:(n-1)]-lambda/beta/(1-beta)*
  (lambda*quociente1^ 2-quociente1-lambda*quociente2)))
hess[2,2]<-p1+p2
hess[1,3]<-sum(I1/alpha/(1-alpha)*(-quociente1-lambda*quociente2+
  lambda*quociente1^ 2))
hess[3,1]<-hess[1,3]
hess[2,3]<-sum(I2/beta/(1-beta)*(-quociente1-lambda*quociente2+
  lambda*quociente1^ 2))
hess[3,2]<-hess[2,3]
hess[3,3]<-sum(quociente2-quociente1^ 2)
return(grad,hess)
}

```

método de máxima verosimilhança para um valor de R fixo e usando o método de Newton

```

threshold_mv_r<-function(x,r,tol){
  n<-length(x);
  Ind_1<-rep(0,n-1);
  Ind_2<-rep(0,n-1);
  for(i in 1:(n-1)){

```

```
if(x[i]>r){
  Ind_2[i]<-1}
else{
  Ind_1[i]<-1}
}
par<-threshold_mq_r(x,r);
parametros<-par$p
#parametros<-c(0.5,0.5,mean(x[2:n])-0.5*mean(x[1:(n-1)]))
it<-0
repeat{
  it<-it+1;
  alpha<-parametros[1]
  beta<-parametros[2]
  lambda<-parametros[3]
  aux<-gradiente_hessiano(alpha,beta,lambda,x,Ind_1,Ind_2)
  if(any(is.na(aux$hess))){
    parametros<-rep(NA,3);break}
  if(det(aux$hess)==0){
    parametros<-rep(NA,3);break}
  else{
    dx<-solve(aux$hess,aux$grad)
    parametros<-parametros-dx
    if(any(is.na(dx))){parametros<-rep(NA,3);break}
    if(all(dx<tol)){break}
    if(any(parametros<0)){
      parametros<-rep(NA,3)
      break
    }
  }
  else{
```

```

        if(any(parametros[1:2]>1)){
            parametros<-rep(NA,3)
        }
    }
}
}
if(all(!is.na(parametros))){
    aux<-gradiente_hessiano(parametros[1],parametros[2],parametros[3],x,Ind_1,Ind_2)
    gradiente<-aux$grad;
    aux<-funcao_p_threshold(x[1:(n-1)],x[2:n],alpha*Ind_1,beta*Ind_2,lambda)
    lv<-sum(log(aux))
}
else{
    gradiente<-NA
    lv<-NA
}
alpha<-parametros[1]
beta<-parametros[2]
lambda<-parametros[3]
aux<-gradiente_hessiano(alpha,beta,lambda,x,Ind_1,Ind_2)
hess<-aux$hess
return(parametros,gradiente,hess,it,lv)
}

#para vários valores de r
threshold_mv<-function(x,tol){
    rmin<-trunc(quantile(x,0.1));
    rmax<-trunc(quantile(x,0.9))+1;

```

```
if(rmin==0){rmin<-1}
r<-rmin:rmax;
n<-length(r);
parametros_lv<-array(0,dim=c(5,n))
parametros_lv[4,]<-r
gradiente<-array(0,dim=c(3,n))
aux<-threshold_mv_r(x,r[1],tol);
parametros_lv[1:3,1]<-aux$parametros;
parametros_lv[5,1]<-aux$lv;
if(is.na(parametros_lv[5,1])){
  r_max_lv<-0
}
else{
  r_max_lv<-1
}
for(i in 2:n){
  aux<-threshold_mv_r(x,r[i],tol);
  parametros_lv[1:3,i]<-aux$parametros;
  parametros_lv[5,i]<-aux$lv;
  gradiente[,i]<-aux$gradiente
  if(r_max_lv==0){
    if(!is.na(parametros_lv[5,i])){
      r_max_lv<-i
    }
  }
  else{
    if(!is.na(parametros_lv[5,i])){
      if(parametros_lv[5,i]>parametros_lv[5,r_max_lv]){
        r_max_lv<-i
      }
    }
  }
}
```

```
    }  
  }  
}  
}  
par<-parametros_lv[r_max_lv]  
return(par,parametros_lv,r_max_lv,gradiente)  
}
```

Apêndice B

Tabelas e Figuras resultantes do estudo de simulação apresentado no Capítulo 2

Neste apêndice apresentam-se as tabelas e figuras de outros modelos $\text{PINAR}(1)_T$ que foram simulados.

Tabela B.1: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (2, 4, 1, 3)$.

N		α				λ			
20	YW	0.793 (0.020)	0.485 (0.039)	0.710 (0.015)	0.603 (0.034)	2.406 (1.240)	4.119 (2.826)	1.410 (0.962)	3.178 (1.925)
	CLS	0.795 (0.017)	0.485 (0.039)	0.710 (0.015)	0.603 (0.034)	2.385 (1.068)	4.119 (2.826)	1.410 (0.962)	3.185 (1.941)
	WCLS	0.796 (0.017)	0.485 (0.038)	0.709 (0.015)	0.603 (0.034)	2.383 (1.082)	4.119 (2.798)	1.418 (0.948)	3.187 (1.926)
	CML	0.856 (0.004)	0.544 (0.019)	0.759 (0.004)	0.656 (0.014)	1.903 (0.304)	3.573 (1.448)	1.001 (0.208)	2.764 (0.762)
50	YW	0.841 (0.007)	0.494 (0.016)	0.749 (0.006)	0.630 (0.013)	2.056 (0.429)	4.026 (1.124)	1.097 (0.345)	3.000 (0.715)
	CLS	0.841 (0.007)	0.494 (0.016)	0.749 (0.006)	0.630 (0.013)	2.059 (0.400)	4.026 (1.124)	1.097 (0.345)	2.999 (0.716)
	WCLS	0.840 (0.007)	0.495 (0.016)	0.749 (0.005)	0.630 (0.013)	2.059 (0.401)	4.021 (1.116)	1.100 (0.329)	2.998 (0.709)
	CML	0.855 (0.002)	0.517 (0.009)	0.763 (0.002)	0.646 (0.006)	1.940 (0.121)	3.817 (0.657)	0.979 (0.085)	2.864 (0.322)
100	YW	0.847 (0.004)	0.497 (0.008)	0.756 (0.003)	0.628 (0.007)	2.023 (0.248)	4.029 (0.564)	1.027 (0.195)	3.013 (0.387)
	CLS	0.848 (0.004)	0.497 (0.008)	0.756 (0.003)	0.628 (0.007)	2.011 (0.239)	4.029 (0.564)	1.027 (0.195)	3.012 (0.389)
	WCLS	0.848 (0.004)	0.496 (0.007)	0.756 (0.003)	0.628 (0.007)	2.012 (0.233)	4.037 (0.543)	1.028 (0.177)	3.014 (0.381)
	CML	0.852 (0.001)	0.509 (0.005)	0.761 (0.001)	0.636 (0.003)	1.973 (0.061)	3.917 (0.336)	0.983 (0.048)	2.948 (0.164)
500	YW	0.85 (0.001)	0.500 (0.002)	0.759 (0.001)	0.630 (0.001)	2.003 (0.048)	4.002 (0.108)	1.010 (0.046)	3.000 (0.074)
	CLS	0.85 (0.001)	0.500 (0.002)	0.759 (0.001)	0.630 (0.001)	2.002 (0.048)	4.002 (0.108)	1.010 (0.046)	3.000 (0.074)
	WCLS	0.85 (0.001)	0.500 (0.001)	0.759 (0.001)	0.630 (0.001)	2.003 (0.047)	4.000 (0.104)	1.008 (0.042)	3.000 (0.074)
	CML	0.85 (0.0003)	0.502 (0.001)	0.760 (0.0002)	0.632 (0.001)	1.996 (0.013)	3.976 (0.059)	1.001 (0.009)	2.986 (0.034)

Tabela B.2: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (6, 1, 5, 3)$.

N		α				λ			
20	YW	0.743 (0.038)	0.411 (0.020)	0.696 (0.037)	0.601 (0.026)	7.016 (4.156)	2.326 (4.039)	5.511 (2.857)	3.365 (3.337)
	CLS	0.744 (0.035)	0.411 (0.020)	0.696 (0.037)	0.601 (0.026)	7.009 (3.907)	2.326 (4.039)	5.511 (2.857)	3.358 (3.351)
	WCLS	0.745 (0.036)	0.411 (0.019)	0.696 (0.038)	0.601 (0.026)	7.004 (3.941)	2.330 (3.944)	5.510 (2.886)	3.367 (3.310)
	CML	0.821 (0.012)	0.464 (0.008)	0.748 (0.016)	0.634 (0.011)	6.234 (1.526)	1.556 (1.546)	5.075 (1.382)	2.985 (1.425)
50	YW	0.816 (0.012)	0.462 (0.006)	0.744 (0.016)	0.625 (0.011)	6.300 (1.390)	1.564 (1.242)	5.123 (1.110)	3.037 (1.431)
	CLS	0.817 (0.012)	0.462 (0.006)	0.744 (0.016)	0.625 (0.011)	6.292 (1.332)	1.564 (1.242)	5.123 (1.110)	3.035 (1.429)
	WCLS	0.817 (0.012)	0.462 (0.006)	0.744 (0.015)	0.625 (0.011)	6.291 (1.333)	1.566 (1.181)	5.125 (1.100)	3.034 (1.412)
	CML	0.845 (0.005)	0.489 (0.002)	0.761 (0.006)	0.630 (0.004)	6.010 (0.546)	1.172 (0.455)	4.982 (0.482)	2.977 (0.522)
100	YW	0.839 (0.007)	0.484 (0.003)	0.754 (0.008)	0.626 (0.006)	6.095 (0.712)	1.237 (0.516)	5.051 (0.613)	3.033 (0.740)
	CLS	0.840 (0.006)	0.484 (0.003)	0.754 (0.008)	0.626 (0.006)	6.093 (0.687)	1.237 (0.516)	5.051 (0.613)	3.032 (0.739)
	WCLS	0.840 (0.006)	0.485 (0.002)	0.754 (0.008)	0.626 (0.006)	6.093 (0.688)	1.226 (0.496)	5.049 (0.607)	3.035 (0.723)
	CML	0.849 (0.002)	0.497 (0.001)	0.759 (0.003)	0.631 (0.002)	6.002 (0.290)	1.048 (0.220)	5.005 (0.263)	2.978 (0.284)
500	YW	0.849 (0.002)	0.497 (0.001)	0.759 (0.002)	0.631 (0.001)	6.011 (0.169)	1.037 (0.135)	5.005 (0.113)	2.985 (0.130)
	CLS	0.849 (0.002)	0.497 (0.001)	0.759 (0.002)	0.631 (0.001)	6.009 (0.167)	1.037 (0.135)	5.005 (0.113)	2.986 (0.130)
	WCLS	0.849 (0.002)	0.497 (0.001)	0.759 (0.002)	0.631 (0.001)	6.008 (0.167)	1.041 (0.122)	5.006 (0.111)	2.985 (0.130)
	CML	0.850 (0.001)	0.499 (0.0002)	0.760 (0.001)	0.630 (0.0004)	6.004 (0.056)	1.009 (0.047)	4.999 (0.050)	2.995 (0.054)

Tabela B.3: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (3, 6, 1, 5)$.

N		α				λ			
20	YW	0.791 (0.021)	0.482 (0.042)	0.695 (0.015)	0.604 (0.038)	3.664 (3.008)	6.258 (7.096)	1.832 (2.206)	5.243 (4.213)
	CLS	0.793 (0.019)	0.482 (0.042)	0.695 (0.015)	0.604 (0.038)	3.647 (2.636)	6.258 (7.096)	1.832 (2.206)	5.238 (4.208)
	WCLS	0.793 (0.019)	0.481 (0.042)	0.694 (0.015)	0.605 (0.038)	3.642 (2.638)	6.270 (7.015)	1.843 (2.147)	5.232 (4.169)
	CML	0.849 (0.005)	0.502 (0.026)	0.752 (0.003)	0.631 (0.019)	2.996 (0.667)	5.996 (4.448)	1.114 (0.370)	4.960 (2.160)
50	YW	0.834 (0.007)	0.494 (0.015)	0.739 (0.005)	0.617 (0.015)	3.174 (1.039)	6.051 (2.504)	1.268 (0.665)	5.132 (1.742)
	CLS	0.834 (0.007)	0.494 (0.015)	0.739 (0.005)	0.617 (0.015)	3.176 (0.966)	6.051 (2.504)	1.268 (0.665)	5.129 (1.741)
	WCLS	0.834 (0.007)	0.495 (0.015)	0.738 (0.004)	0.617 (0.015)	3.174 (0.955)	6.048 (2.494)	1.272 (0.627)	5.125 (1.719)
	CML	0.852 (0.002)	0.503 (0.009)	0.760 (0.001)	0.629 (0.007)	2.972 (0.257)	5.934 (1.504)	1.005 (0.148)	5.008 (0.840)
100	YW	0.849 (0.004)	0.495 (0.008)	0.752 (0.002)	0.629 (0.007)	3.006 (0.534)	6.055 (1.275)	1.089 (0.341)	5.016 (0.788)
	CLS	0.848 (0.004)	0.495 (0.008)	0.752 (0.002)	0.629 (0.007)	3.015 (0.517)	6.055 (1.275)	1.089 (0.341)	5.016 (0.789)
	WCLS	0.849 (0.004)	0.495 (0.008)	0.752 (0.002)	0.629 (0.007)	3.013 (0.508)	6.055 (1.256)	1.087 (0.315)	5.017 (0.782)
	CML	0.852 (0.001)	0.501 (0.005)	0.760 (0.001)	0.632 (0.003)	2.974 (0.129)	5.973 (0.791)	0.997 (0.074)	4.979 (0.393)
500	YW	0.849 (0.001)	0.497 (0.002)	0.761 (0.001)	0.632 (0.002)	3.012 (0.108)	6.040 (0.247)	0.988 (0.080)	4.977 (0.172)
	CLS	0.849 (0.001)	0.497 (0.002)	0.761 (0.001)	0.632 (0.002)	3.012 (0.107)	6.040 (0.247)	0.988 (0.080)	4.977 (0.172)
	WCLS	0.849 (0.001)	0.497 (0.001)	0.761 (0.001)	0.632 (0.002)	3.011 (0.104)	6.043 (0.243)	0.993 (0.072)	4.976 (0.171)
	CML	0.850 (0.0002)	0.500 (0.001)	0.761 (0.0001)	0.630 (0.001)	3.005 (0.026)	6.003 (0.155)	0.991 (0.014)	4.994 (0.081)

Tabela B.4: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.6, 0.4, 0.8, 0.25)$ e $\lambda = (3, 2, 4, 1)$.

N		α				λ			
20	YW	0.547 (0.056)	0.416 (0.031)	0.685 (0.050)	0.211 (0.011)	3.133 (0.640)	1.940 (0.736)	4.398 (0.966)	1.314 (0.557)
	CLS	0.548 (0.054)	0.416 (0.031)	0.685 (0.050)	0.211 (0.011)	3.129 (0.626)	1.940 (0.736)	4.398 (0.966)	1.312 (0.560)
	WCLS	0.551 (0.054)	0.411 (0.030)	0.685 (0.050)	0.206 (0.011)	3.122 (0.627)	1.962 (0.698)	4.396 (0.970)	1.340 (0.548)
	CML	0.608 (0.034)	0.445 (0.020)	0.773 (0.019)	0.253 (0.007)	2.922 (0.404)	1.784 (0.461)	4.008 (0.471)	1.000 (0.334)
50	YW	0.586 (0.029)	0.397 (0.014)	0.764 (0.020)	0.231 (0.006)	3.041 (0.277)	2.010 (0.350)	4.104 (0.342)	1.135 (0.298)
	CLS	0.586 (0.029)	0.397 (0.014)	0.764 (0.020)	0.231 (0.006)	3.040 (0.276)	2.010 (0.350)	4.104 (0.342)	1.134 (0.297)
	WCLS	0.587 (0.029)	0.396 (0.014)	0.765 (0.019)	0.230 (0.006)	3.038 (0.274)	2.015 (0.335)	4.101 (0.341)	1.147 (0.284)
	CML	0.613 (0.018)	0.411 (0.010)	0.800 (0.009)	0.249 (0.004)	2.952 (0.195)	1.935 (0.237)	3.942 (0.192)	1.007 (0.197)
100	YW	0.594 (0.016)	0.398 (0.007)	0.793 (0.010)	0.246 (0.003)	3.027 (0.149)	2.008 (0.163)	4.026 (0.196)	1.022 (0.159)
	CLS	0.594 (0.016)	0.398 (0.007)	0.793 (0.010)	0.246 (0.003)	3.027 (0.150)	2.008 (0.163)	4.026 (0.196)	1.023 (0.159)
	WCLS	0.593 (0.016)	0.397 (0.007)	0.794 (0.010)	0.245 (0.003)	3.028 (0.148)	2.013 (0.157)	4.024 (0.195)	1.030 (0.149)
	CML	0.604 (0.010)	0.406 (0.005)	0.806 (0.005)	0.254 (0.002)	2.991 (0.099)	1.966 (0.107)	3.968 (0.121)	0.963 (0.111)
500	YW	0.605 (0.044)	0.400 (0.002)	0.799 (0.003)	0.250 (0.001)	2.982 (0.324)	2.004 (0.037)	4.005 (0.044)	0.998 (0.039)
	CLS	0.605 (0.044)	0.400 (0.002)	0.799 (0.003)	0.250 (0.001)	2.982 (0.324)	2.004 (0.037)	4.005 (0.044)	0.998 (0.039)
	WCLS	0.605 (0.043)	0.400 (0.002)	0.800 (0.003)	0.251 (0.001)	2.982 (0.316)	2.004 (0.035)	4.004 (0.044)	0.997 (0.038)
	CML	0.602 (0.024)	0.401 (0.001)	0.803 (0.001)	0.251 (0.001)	2.987 (0.170)	2.000 (0.024)	3.989 (0.023)	0.991 (0.025)

Tabela B.5: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (2, 4, 1, 3)$.

N		α				λ			
20	YW	0.588 (0.033)	0.423 (0.048)	0.774 (0.014)	0.293 (0.031)	2.041 (0.687)	3.910 (1.215)	1.156 (0.448)	2.753 (1.069)
	CLS	0.590 (0.032)	0.423 (0.048)	0.774 (0.014)	0.293 (0.031)	2.032 (0.664)	3.910 (1.215)	1.156 (0.448)	2.755 (1.061)
	WCLS	0.591 (0.031)	0.421 (0.046)	0.775 (0.013)	0.291 (0.030)	2.031 (0.651)	3.921 (1.184)	1.156 (0.445)	2.771 (1.023)
	CML	0.613 (0.019)	0.432 (0.038)	0.804 (0.004)	0.299 (0.026)	1.935 (0.411)	3.867 (0.999)	0.982 (0.157)	2.723 (0.919)
50	YW	0.598 (0.015)	0.401 (0.026)	0.793 (0.007)	0.248 (0.014)	2.009 (0.316)	3.999 (0.609)	1.047 (0.219)	3.003 (0.498)
	CLS	0.598 (0.015)	0.401 (0.026)	0.793 (0.007)	0.248 (0.014)	2.010 (0.314)	3.999 (0.609)	1.047 (0.219)	3.002 (0.499)
	WCLS	0.597 (0.014)	0.400 (0.025)	0.794 (0.006)	0.248 (0.014)	2.014 (0.302)	4.001 (0.597)	1.042 (0.209)	3.006 (0.487)
	CML	0.603 (0.008)	0.402 (0.019)	0.802 (0.002)	0.254 (0.012)	1.989 (0.172)	3.992 (0.466)	0.991 (0.062)	2.970 (0.432)
100	YW	0.595 (0.009)	0.396 (0.012)	0.798 (0.004)	0.248 (0.008)	2.016 (0.189)	4.013 (0.293)	1.010 (0.124)	3.015 (0.269)
	CLS	0.595 (0.009)	0.396 (0.012)	0.798 (0.004)	0.248 (0.008)	2.015 (0.187)	4.013 (0.293)	1.010 (0.124)	3.015 (0.269)
	WCLS	0.595 (0.008)	0.396 (0.012)	0.798 (0.003)	0.248 (0.007)	2.016 (0.179)	4.016 (0.293)	1.011 (0.114)	3.018 (0.263)
	CML	0.599 (0.004)	0.398 (0.009)	0.799 (0.001)	0.251 (0.007)	1.995 (0.092)	4.006 (0.238)	1.001 (0.029)	3.000 (0.239)
500	YW	0.601 (0.001)	0.398 (0.002)	0.799 (0.001)	0.249 (0.001)	1.994 (0.030)	4.005 (0.052)	1.004 (0.025)	2.996 (0.052)
	CLS	0.601 (0.001)	0.398 (0.002)	0.799 (0.001)	0.249 (0.001)	1.995 (0.030)	4.005 (0.052)	1.004 (0.025)	2.996 (0.052)
	WCLS	0.601 (0.001)	0.398 (0.002)	0.799 (0.001)	0.249 (0.001)	1.994 (0.030)	4.006 (0.051)	1.004 (0.023)	2.997 (0.051)
	CML	0.600 (0.001)	0.397 (0.002)	0.799 (0.0002)	0.249 (0.001)	1.997 (0.015)	4.011 (0.041)	1.001 (0.006)	2.994 (0.045)

Tabela B.6: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (1, 6, 3, 5)$.

N		α				λ			
20	YW	0.555 (0.016)	0.427 (0.051)	0.757 (0.022)	0.289 (0.029)	1.395 (0.871)	5.864 (1.907)	3.348 (1.628)	4.663 (2.806)
	CLS	0.556 (0.015)	0.427 (0.051)	0.757 (0.022)	0.289 (0.029)	1.390 (0.811)	5.864 (1.907)	3.348 (1.628)	4.653 (2.826)
	WCLS	0.557 (0.015)	0.428 (0.051)	0.758 (0.022)	0.288 (0.029)	1.381 (0.770)	5.860 (1.906)	3.341 (1.625)	4.662 (2.777)
	CML	0.595 (0.007)	0.444 (0.043)	0.798 (0.008)	0.295 (0.025)	1.095 (0.345)	5.768 (1.612)	3.008 (0.663)	4.599 (2.387)
50	YW	0.582 (0.007)	0.415 (0.029)	0.793 (0.010)	0.253 (0.013)	1.130 (0.376)	5.905 (0.963)	3.048 (0.688)	4.974 (1.273)
	CLS	0.582 (0.007)	0.415 (0.029)	0.793 (0.010)	0.253 (0.013)	1.133 (0.364)	5.905 (0.963)	3.048 (0.688)	4.975 (1.269)
	WCLS	0.581 (0.007)	0.414 (0.029)	0.793 (0.010)	0.252 (0.013)	1.139 (0.333)	5.910 (0.958)	3.048 (0.690)	4.982 (1.256)
	CML	0.597 (0.003)	0.413 (0.022)	0.801 (0.004)	0.255 (0.012)	1.023 (0.137)	5.917 (0.750)	2.984 (0.278)	4.960 (1.149)
100	YW	0.598 (0.004)	0.403 (0.015)	0.798 (0.005)	0.252 (0.007)	1.013 (0.213)	5.991 (0.482)	3.007 (0.387)	4.987 (0.691)
	CLS	0.598 (0.004)	0.403 (0.015)	0.798 (0.005)	0.252 (0.007)	1.013 (0.211)	5.991 (0.482)	3.007 (0.387)	4.987 (0.689)
	WCLS	0.597 (0.004)	0.402 (0.015)	0.798 (0.005)	0.251 (0.007)	1.018 (0.183)	5.993 (0.482)	3.007 (0.382)	4.992 (0.680)
	CML	0.601 (0.002)	0.403 (0.012)	0.801 (0.002)	0.253 (0.006)	0.989 (0.070)	5.986 (0.395)	2.986 (0.135)	4.969 (0.591)
500	YW	0.601 (0.001)	0.402 (0.003)	0.800 (0.001)	0.252 (0.001)	0.991 (0.045)	5.991 (0.087)	3.001 (0.082)	4.982 (0.140)
	CLS	0.601 (0.001)	0.402 (0.003)	0.800 (0.001)	0.252 (0.001)	0.992 (0.044)	5.991 (0.087)	3.001 (0.082)	4.982 (0.140)
	WCLS	0.600 (0.001)	0.401 (0.003)	0.800 (0.001)	0.252 (0.001)	0.997 (0.040)	5.992 (0.086)	3.001 (0.081)	4.983 (0.139)
	CML	0.600 (0.0003)	0.402 (0.002)	0.800 (0.0004)	0.252 (0.001)	0.995 (0.014)	5.989 (0.071)	2.999 (0.027)	4.981 (0.120)

Tabela B.7: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (4, 1, 3, 2)$.

N		α				λ			
20	YW	0.440 (0.056)	0.200 (0.011)	0.492 (0.057)	0.187 (0.024)	3.901 (0.544)	1.031 (0.285)	3.012 (0.390)	1.687 (0.415)
	CLS	0.440 (0.055)	0.200 (0.011)	0.492 (0.057)	0.187 (0.024)	3.898 (0.532)	1.060 (0.269)	3.012 (0.388)	1.699 (0.408)
	WCLS	0.441 (0.055)	0.194 (0.010)	0.491 (0.057)	0.186 (0.024)	3.901 (0.533)	1.031 (0.285)	3.012 (0.390)	1.694 (0.409)
	CML	0.451 (0.051)	0.206 (0.010)	0.513 (0.049)	0.192 (0.024)	3.872 (0.500)	0.998 (0.257)	2.968 (0.346)	1.673 (0.408)
50	YW	0.402 (0.035)	0.198 (0.007)	0.497 (0.035)	0.134 (0.009)	3.995 (0.282)	1.012 (0.188)	3.008 (0.193)	1.866 (0.174)
	CLS	0.402 (0.035)	0.198 (0.007)	0.497 (0.035)	0.134 (0.009)	3.996 (0.280)	1.020 (0.172)	3.008 (0.191)	1.866 (0.177)
	WCLS	0.401 (0.035)	0.196 (0.007)	0.497 (0.035)	0.134 (0.009)	3.995 (0.281)	1.012 (0.188)	3.008 (0.193)	1.866 (0.175)
	CML	0.411 (0.032)	0.203 (0.006)	0.502 (0.029)	0.135 (0.008)	3.974 (0.260)	0.988 (0.160)	2.998 (0.164)	1.859 (0.165)
100	YW	0.408 (0.020)	0.201 (0.004)	0.499 (0.020)	0.118 (0.005)	3.968 (0.145)	0.991 (0.103)	2.991 (0.102)	1.919 (0.095)
	CLS	0.408 (0.020)	0.201 (0.004)	0.499 (0.020)	0.118 (0.005)	3.968 (0.142)	0.999 (0.095)	2.992 (0.100)	1.920 (0.095)
	WCLS	0.408 (0.019)	0.199 (0.004)	0.499 (0.020)	0.118 (0.005)	3.968 (0.144)	0.991 (0.103)	2.991 (0.102)	1.919 (0.095)
	CML	0.409 (0.017)	0.201 (0.003)	0.501 (0.014)	0.118 (0.005)	3.967 (0.128)	0.989 (0.085)	2.986 (0.078)	1.919 (0.092)
500	YW	0.399 (0.004)	0.199 (0.001)	0.501 (0.004)	0.101 (0.001)	4.005 (0.030)	1.003 (0.020)	2.999 (0.020)	1.997 (0.024)
	CLS	0.399 (0.004)	0.199 (0.001)	0.501 (0.004)	0.101 (0.001)	4.006 (0.029)	1.001 (0.019)	2.999 (0.019)	1.998 (0.024)
	WCLS	0.399 (0.004)	0.200 (0.001)	0.502 (0.004)	0.101 (0.001)	4.005 (0.029)	1.003 (0.020)	2.999 (0.020)	1.997 (0.024)
	CML	0.400 (0.003)	0.201 (0.001)	0.502 (0.003)	0.101 (0.001)	4.004 (0.025)	0.997 (0.016)	2.998 (0.015)	1.999 (0.023)

Tabela B.8: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.40, 0.20, 0.50, 0.1)$ e $\lambda = (6, 1, 5, 3)$.

N		α				λ			
20	YW	0.437 (0.061)	0.174 (0.009)	0.485 (0.065)	0.185 (0.023)	5.856 (1.078)	1.228 (0.464)	5.002 (0.714)	2.503 (0.939)
	CLS	0.435 (0.060)	0.174 (0.009)	0.485 (0.065)	0.185 (0.023)	5.859 (1.062)	1.228 (0.464)	5.002 (0.714)	2.498 (0.957)
	WCLS	0.435 (0.060)	0.172 (0.008)	0.484 (0.065)	0.184 (0.022)	5.858 (1.062)	1.241 (0.453)	5.003 (0.716)	2.506 (0.937)
	CML	0.446 (0.056)	0.186 (0.008)	0.507 (0.059)	0.188 (0.022)	5.821 (0.989)	1.138 (0.400)	4.943 (0.656)	2.475 (0.933)
50	YW	0.403 (0.035)	0.191 (0.0054)	0.505 (0.042)	0.138 (0.009)	5.986 (0.584)	1.071 (0.293)	4.994 (0.362)	2.775 (0.375)
	CLS	0.402 (0.035)	0.191 (0.0054)	0.505 (0.042)	0.138 (0.009)	5.988 (0.576)	1.071 (0.293)	4.994 (0.362)	2.776 (0.376)
	WCLS	0.402 (0.034)	0.188 (0.0050)	0.505 (0.042)	0.137 (0.009)	5.989 (0.570)	1.094 (0.272)	4.994 (0.360)	2.778 (0.376)
	CML	0.409 (0.031)	0.195 (0.004)	0.510 (0.035)	0.143 (0.010)	5.964 (0.518)	1.044 (0.2380)	4.983 (0.308)	2.746 (0.405)
100	YW	0.405 (0.019)	0.200 (0.003)	0.509 (0.027)	0.115 (0.004)	5.989 (0.310)	0.997 (0.157)	4.982 (0.210)	2.909 (0.198)
	CLS	0.405 (0.019)	0.200 (0.003)	0.509 (0.027)	0.115 (0.004)	5.990 (0.310)	0.997 (0.157)	4.982 (0.210)	2.910 (0.198)
	WCLS	0.404 (0.019)	0.198 (0.003)	0.509 (0.027)	0.115 (0.005)	5.992 (0.310)	1.009 (0.148)	4.982 (0.210)	2.910 (0.199)
	CML	0.409 (0.016)	0.200 (0.002)	0.507 (0.021)	0.116 (0.004)	5.975 (0.277)	0.994 (0.125)	4.987 (0.167)	2.909 (0.188)
500	YW	0.399 (0.004)	0.200 (0.001)	0.504 (0.005)	0.101 (0.001)	6.006 (0.061)	1.002 (0.036)	4.989 (0.039)	2.991 (0.050)
	CLS	0.399 (0.004)	0.200 (0.001)	0.504 (0.005)	0.101 (0.001)	6.006 (0.061)	1.002 (0.036)	4.989 (0.039)	2.991 (0.050)
	WCLS	0.399 (0.004)	0.199 (0.001)	0.504 (0.005)	0.101 (0.001)	6.007 (0.060)	1.007 (0.032)	4.989 (0.038)	2.991 (0.050)
	CML	0.399 (0.003)	0.200 (0.001)	0.503 (0.004)	0.101 (0.001)	6.006 (0.052)	1.001 (0.026)	4.992 (0.031)	2.988 (0.050)

Tabela B.9: Média e EQM (em parêntesis) das estimativas de $\alpha = (0.40, 0.20, 0.50, 0.1)$ e $\lambda = (3, 5, 1, 6)$.

N		α				λ			
20	YW	0.408 (0.034)	0.288 (0.043)	0.455 (0.018)	0.285 (0.075)	2.961 (1.413)	4.533 (1.528)	1.312 (0.630)	5.214 (1.528)
	CLS	0.410 (0.034)	0.288 (0.043)	0.455 (0.018)	0.285 (0.075)	2.948 (1.418)	4.533 (1.528)	1.312 (0.630)	5.211 (1.541)
	WCLS	0.408 (0.033)	0.286 (0.042)	0.456 (0.017)	0.283 (0.074)	2.958 (1.375)	4.544 (1.494)	1.305 (0.602)	5.216 (1.543)
	CML	0.419 (0.025)	0.291 (0.037)	0.493 (0.010)	0.297 (0.078)	2.894 (1.067)	4.514 (1.321)	1.079 (0.340)	5.161 (1.583)
50	YW	0.404 (0.0155)	0.224 (0.0175)	0.487 (0.0088)	0.198 (0.0282)	2.980 (0.6610)	4.866 (0.6392)	1.093 (0.3111)	5.613 (0.5506)
	CLS	0.404 (0.0153)	0.224 (0.0175)	0.487 (0.0088)	0.198 (0.0282)	2.979 (0.6513)	4.866 (0.6392)	1.093 (0.3111)	5.615 (0.5503)
	WCLS	0.403 (0.0149)	0.223 (0.0172)	0.485 (0.0079)	0.197 (0.0277)	2.987 (0.6360)	4.871 (0.6271)	1.103 (0.2778)	5.619 (0.5462)
	CML	0.406 (0.0108)	0.229 (0.0163)	0.498 (0.0045)	0.201 (0.0273)	2.965 (0.4561)	4.837 (0.5940)	1.025 (0.1431)	5.605 (0.5346)
100	YW	0.397 (0.008)	0.211 (0.010)	0.499 (0.004)	0.147 (0.012)	3.006 (0.340)	4.932 (0.374)	1.008 (0.154)	5.813 (0.246)
	CLS	0.397 (0.008)	0.211 (0.010)	0.499 (0.004)	0.147 (0.012)	3.007 (0.337)	4.932 (0.374)	1.008 (0.154)	5.813 (0.246)
	WCLS	0.396 (0.008)	0.211 (0.010)	0.497 (0.004)	0.147 (0.012)	3.013 (0.323)	4.934 (0.372)	1.018 (0.140)	5.812 (0.246)
	CML	0.400 (0.005)	0.211 (0.009)	0.502 (0.002)	0.148 (0.011)	2.987 (0.225)	4.932 (0.338)	0.987 (0.074)	5.807 (0.240)
500	YW	0.399 (0.001)	0.201 (0.002)	0.5 (0.001)	0.109 (0.003)	3.002 (0.064)	4.994 (0.075)	0.998 (0.033)	5.967 (0.059)
	CLS	0.399 (0.001)	0.201 (0.002)	0.5 (0.001)	0.109 (0.003)	3.002 (0.064)	4.994 (0.075)	0.998 (0.033)	5.967 (0.059)
	WCLS	0.399 (0.001)	0.201 (0.002)	0.5 (0.001)	0.109 (0.003)	3.001 (0.062)	4.994 (0.074)	0.999 (0.029)	5.967 (0.059)
	CML	0.399 (0.001)	0.202 (0.002)	0.5 (0.0004)	0.109 (0.003)	3.002 (0.047)	4.986 (0.070)	0.995 (0.014)	5.966 (0.058)

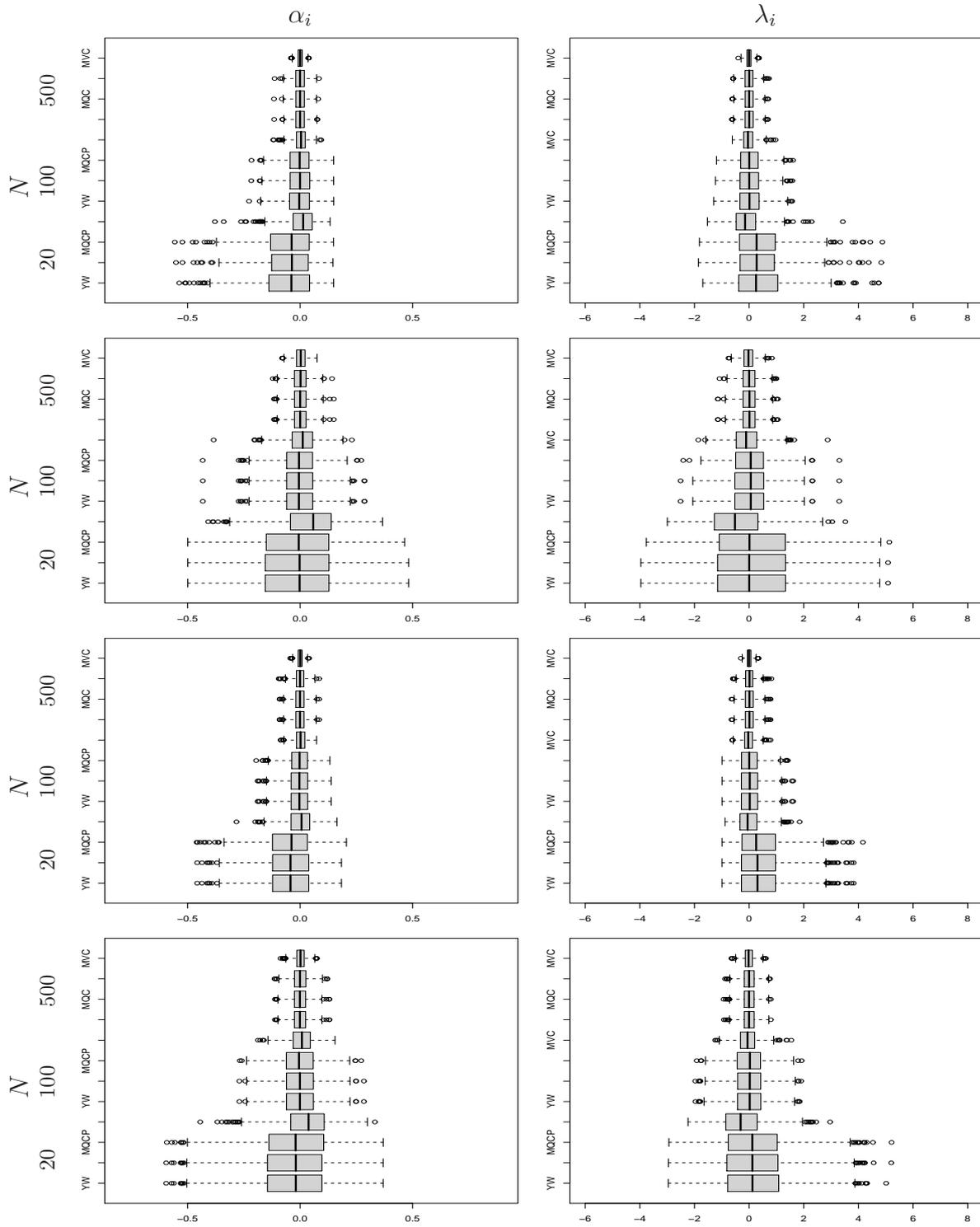


Figura B.1: Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (2, 4, 1, 3)$.

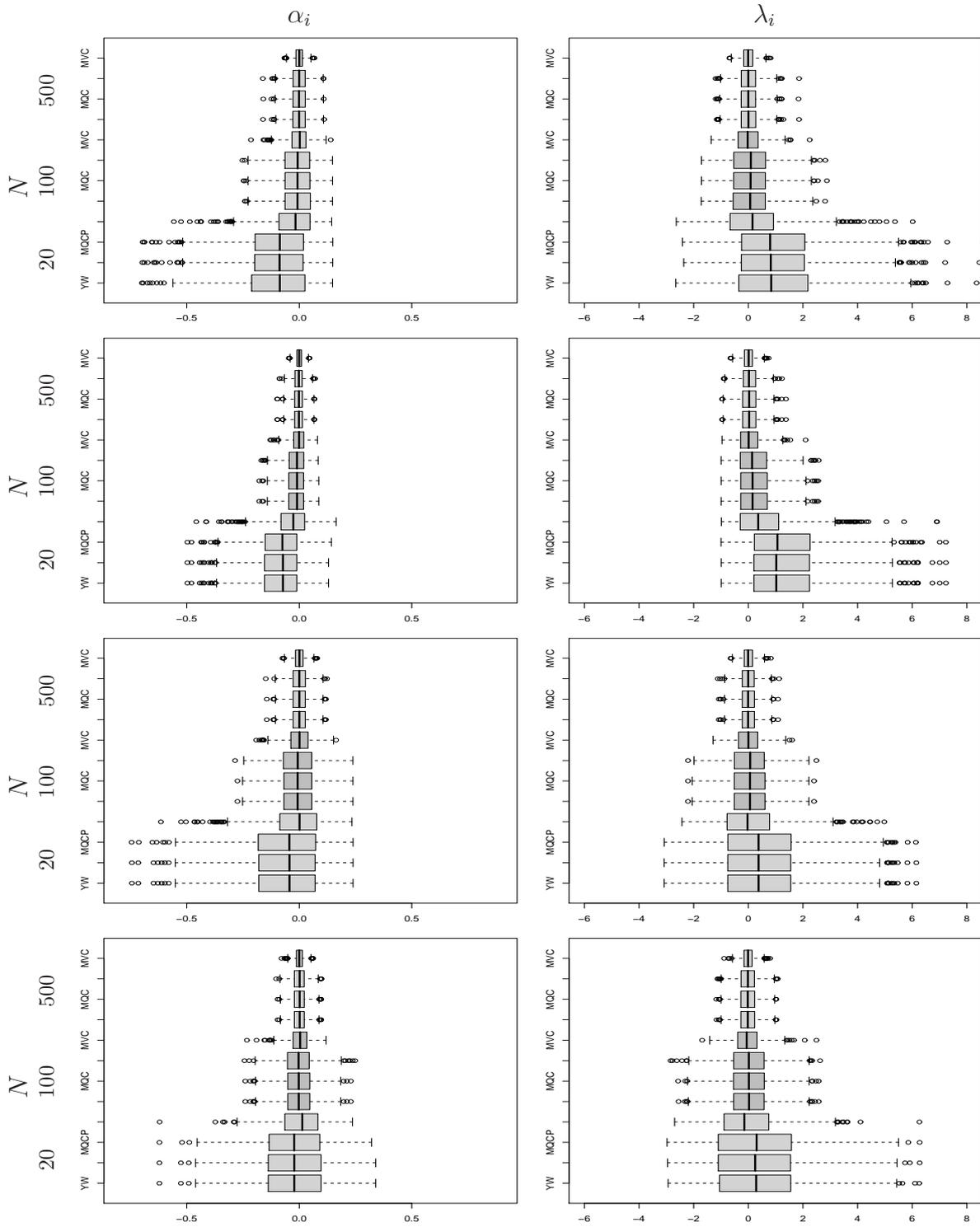


Figura B.2: Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (6, 1, 5, 3)$.

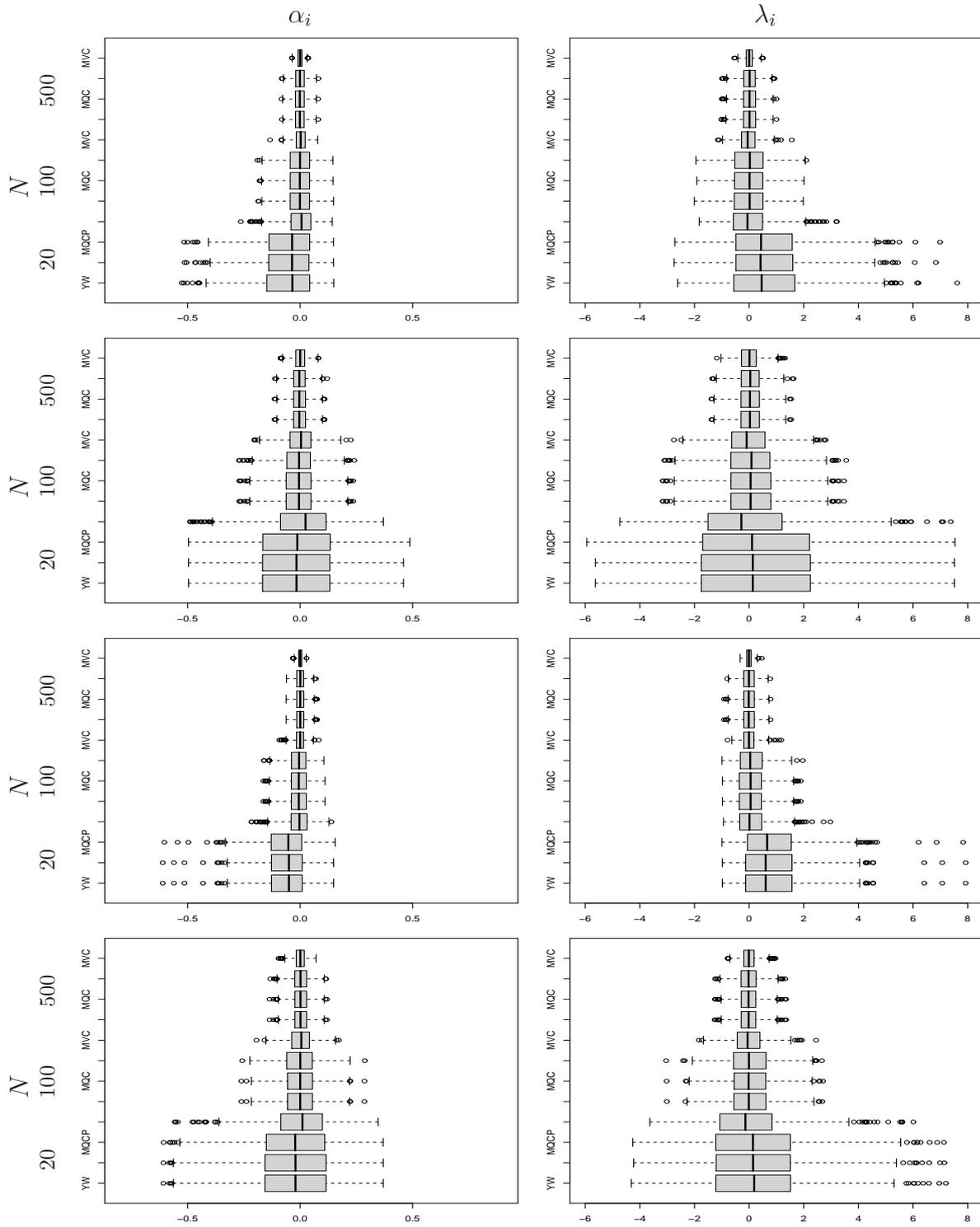


Figura B.3: Viés das estimativas de $\alpha = (0.85, 0.50, 0.76, 0.63)$ e $\lambda = (3, 6, 1, 5)$.

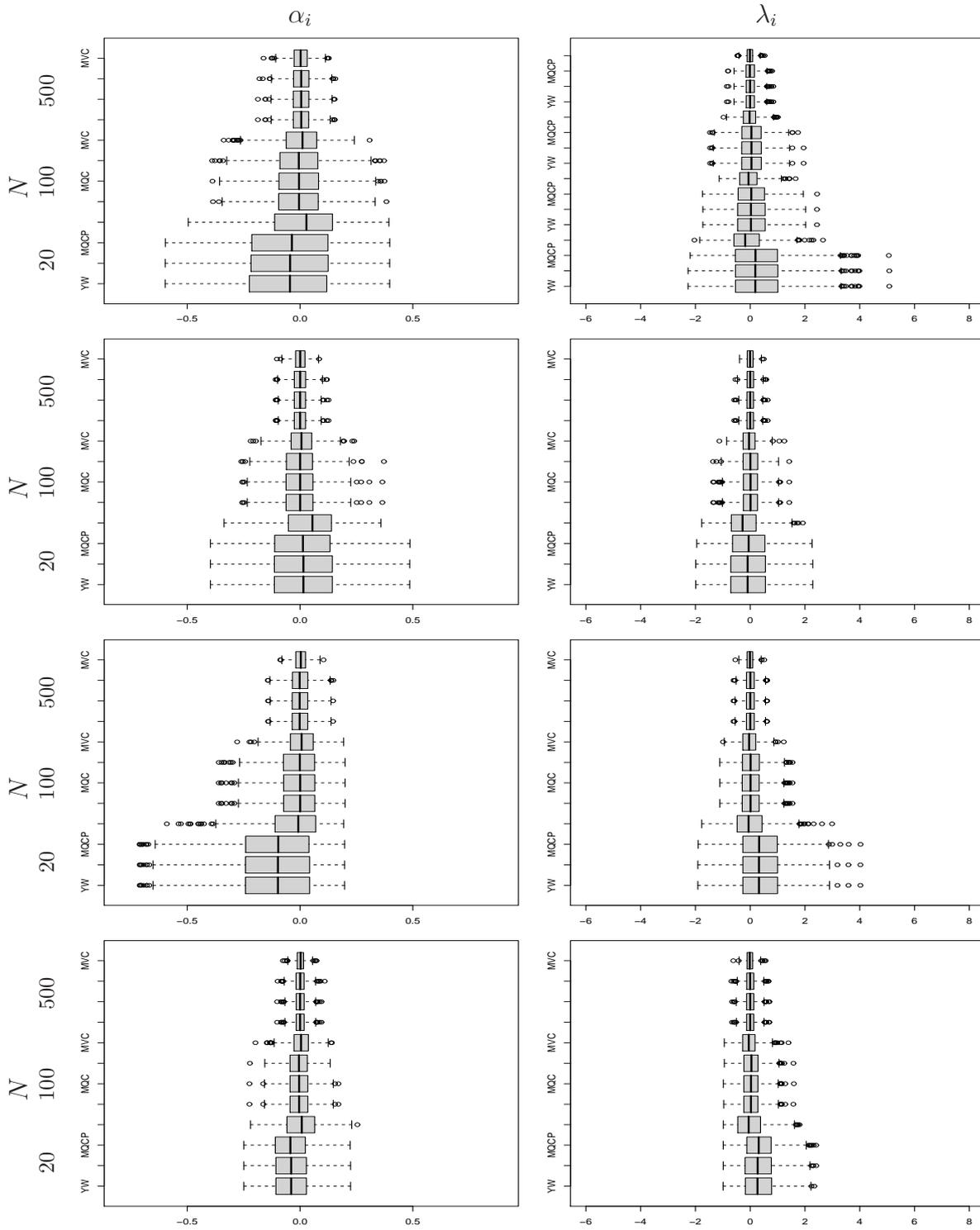


Figura B.4: Viés das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (3, 2, 4, 1)$.

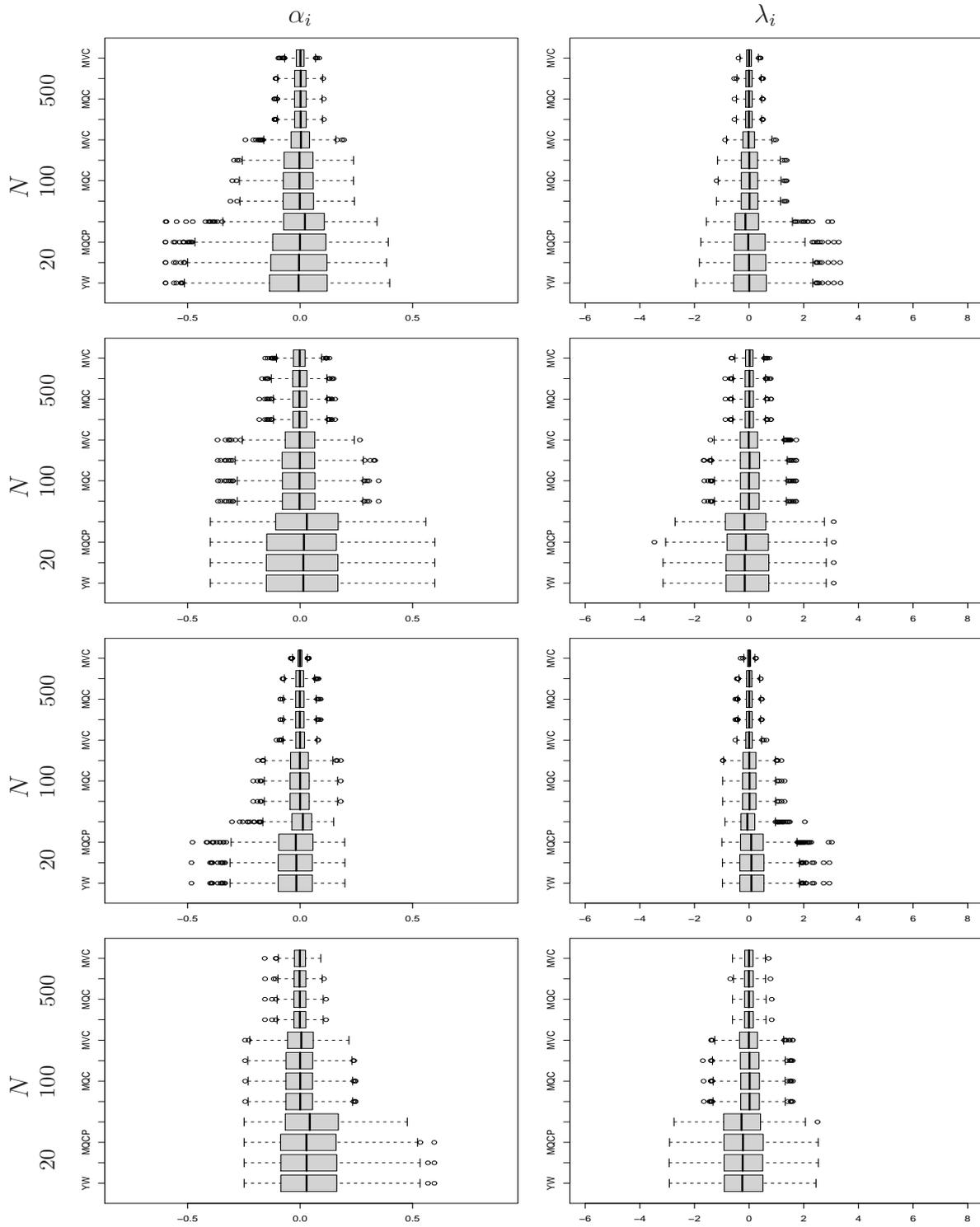


Figura B.5: Viés das estimativas de $\alpha = (0.6, 0.4, 0.8, 0.25)$ e $\lambda = (2, 4, 1, 3)$.

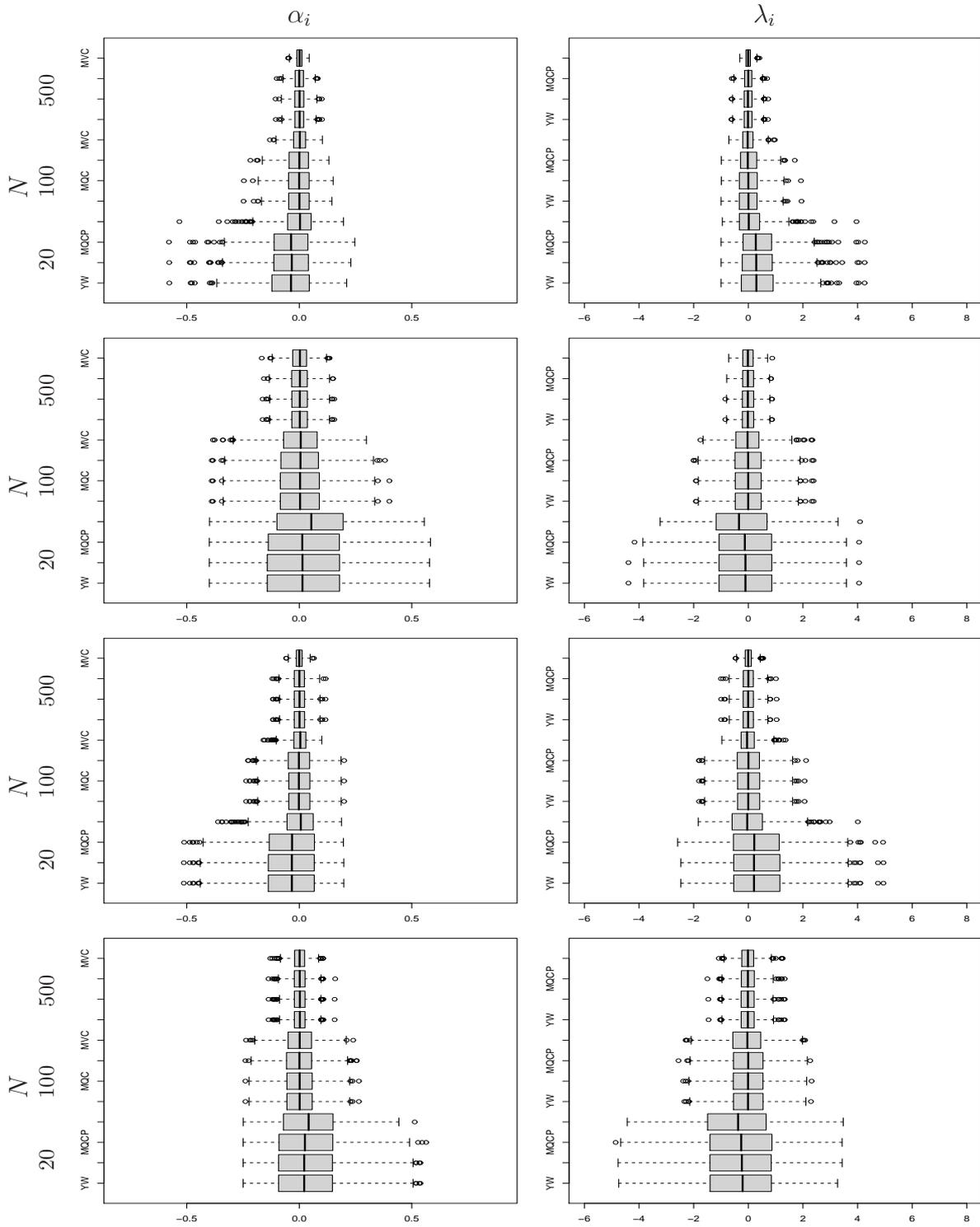


Figura B.6: Viés das estimativas de $\alpha = (0.60, 0.40, 0.80, 0.25)$ e $\lambda = (1, 6, 3, 5)$.

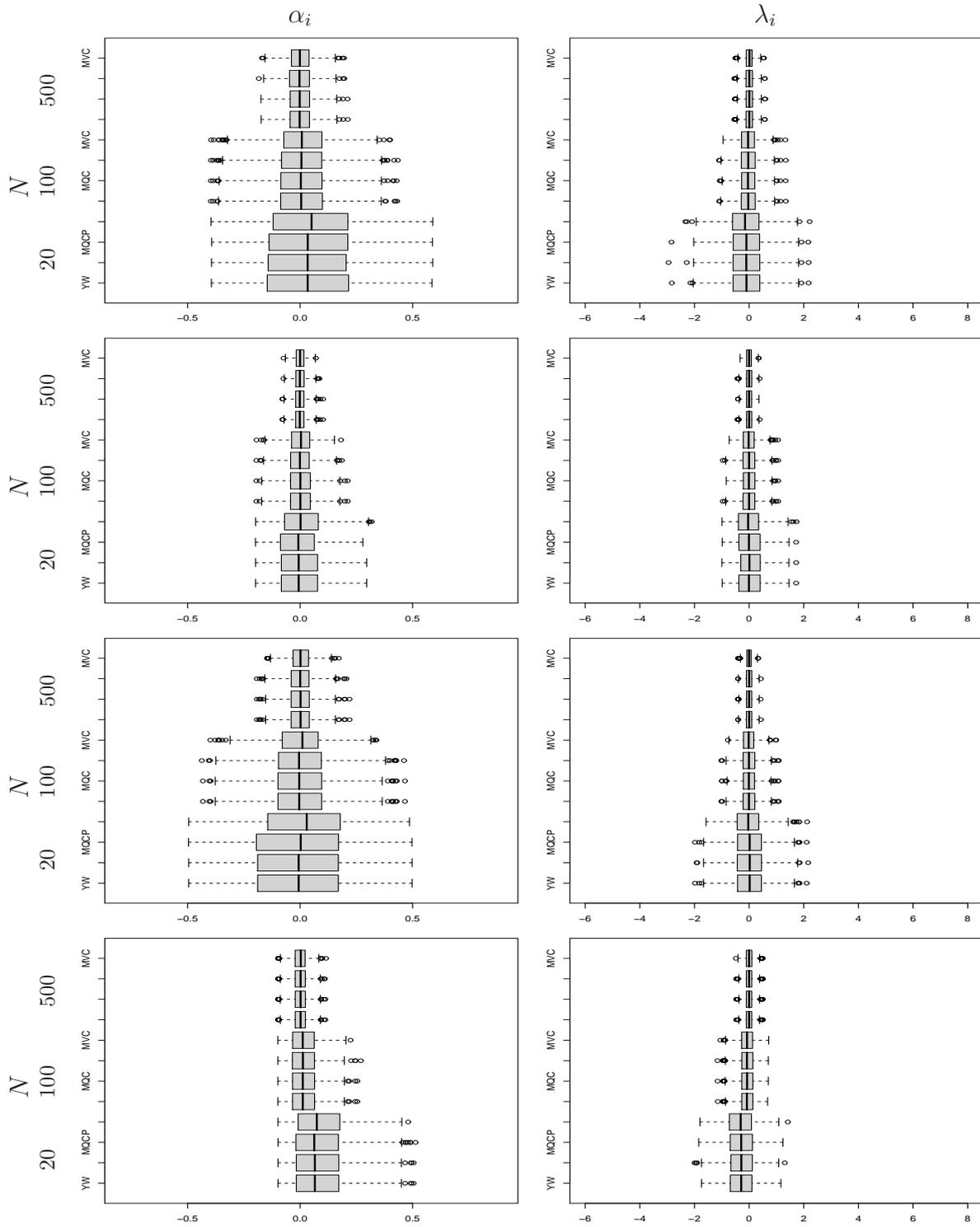


Figura B.7: Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (4, 1, 3, 2)$.

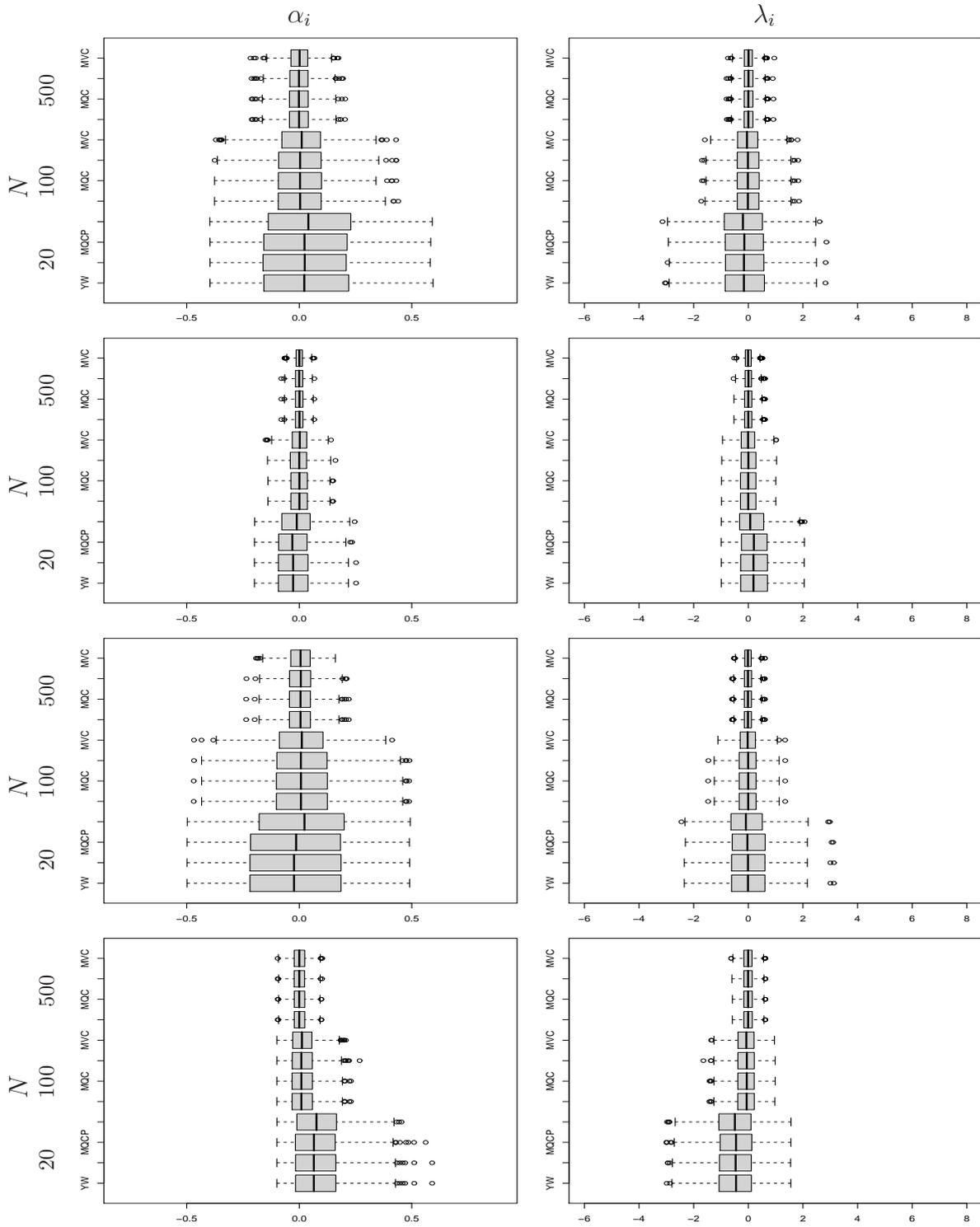


Figura B.8: Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (6, 1, 5, 3)$.

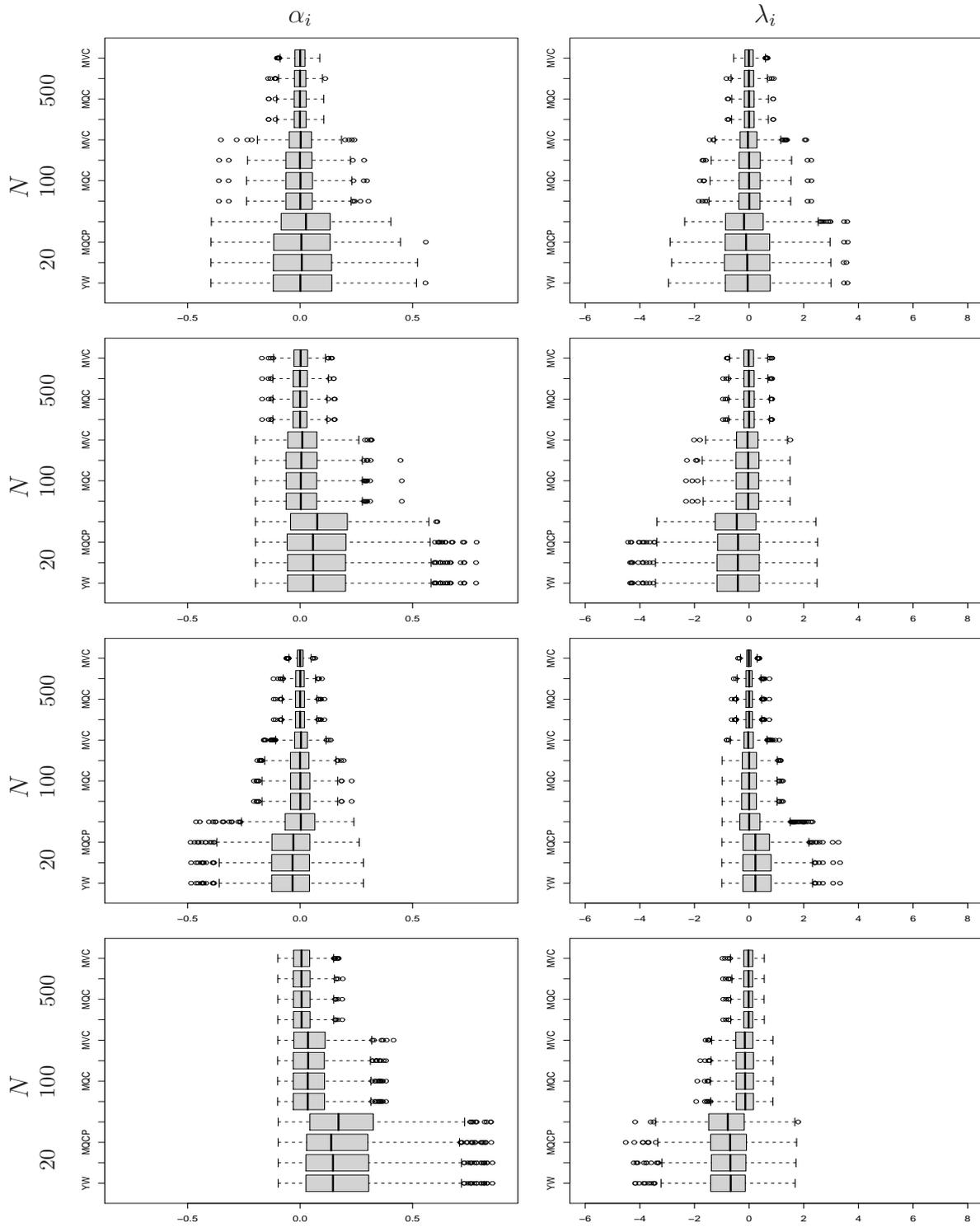


Figura B.9: Viés das estimativas de $\alpha = (0.4, 0.2, 0.5, 0.1)$ e $\lambda = (3, 5, 1, 6)$.

Bibliografia

- [1] Ahn, S., Gyemin, L. e Jongwoo, J. (2000). Analysis of the M/D/1-type queue based on an integer-valued autoregressive process. *Oper. Res. Lett.* **27**, 235–241.
- [2] Alish, M. (2009). The impact of missing data in a generalized integer-valued autoregression model for count data. *J. Biopharm. Statist.* **19**, 1039–1054.
- [3] Al-Osh, M.A. e Alzaid, A.A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Analysis* **8**, 261–275.
- [4] Al-Osh, M.A. e Alzaid, A.A. (1988). Integer-valued moving average (INMA) process. *Stat. Pap.* **29**, 281–300.
- [5] Alzaid, A.A. e Al-Osh, M.A. (1988). First-order integer-valued autoregressive (INAR(1)) process: distributional and regression properties. *Statist. Neerlandica* **42**, 53–61.
- [6] Alzaid, A.A. e Al-Osh, M.A. (1990). An integer-valued p th-order autoregressive structure (INAR(p)) process. *J. Appl. Probab.* **27**, 314–324.
- [7] Alzaid, A.A. e Al-Osh, M.A. (1993). Some autoregressive moving average process with generalized poisson marginal distribution. *Ann. Inst. Statist. Math.* **45**, 223–232.
- [8] Aly, E.-E. e Bouzar, N. (1994). Explicit stationary distributions for some Galton-Watson processes with immigration. *Commun. Statist.-Stochastic Models* **10**, 499–517.

- [9] Aly, E.-E. e Bouzar, N. (2005). Stationary solutions for integer-valued autoregressive processes. *Int. J. of Math. Math. Sci.* **1**, 1–18.
- [10] Amaral, L.F., Souza, R.C. e Stevenson, M. (2008). A smooth transition periodic autoregressive (STPAR) model for short-term load forecasting. *Int. J. Forecast.* **24**, 603–615.
- [11] Amaral-Turkman, M.A. e Turkman, K.F. (1990). Optimal alarm systems for autoregressive processes; a Bayesian approach. *Comp. Stat. and Data Analysis* **10**, 307–314.
- [12] Antunes, M., Amaral-Turkman, M.A. e Turkman, K.F. (2003). A Bayesian approach to event prediction intervals. *J. Time Ser. Anal.* **24**, 631–646.
- [13] Basawa, I.V. e Lund, R.B. (2001). Large sample properties of parameter estimates for periodic ARMA models. *J. Time Ser. Anal.* **22**, 651–666.
- [14] Basawa, I.V., Lund, R.B. e Shao, Q. (2004). First-order seasonal autoregressive processes with periodically varying parameters. *Statist. Probab. Lett.* **67**, 299–306.
- [15] Bennett, W.R. (1958). Statistics of regenerative digital transmission. *Bell Sys Tech. J.* **37**, 1501–1542.
- [16] Bentarzi, M. e Hallin, M. (1994). On the invertibility of periodic moving-average models. *J. Time Ser. Anal.* **15**, 263–268.
- [17] Billingsley, P. (1961). *Statistical Inference for Markov Processes*. University of Chicago Press, Chicago.
- [18] Bloomfield, P., Hurd, H.L. e Lund, R.B. (1994). Periodic correlation in stratospheric ozone data. *J. Time Ser. Anal.* **15**, 127–150.
- [19] Blundell, R., Griffith, R. e Windmeijer, F. (2002). Individual effects and dynamics in count data models. *J. Econometrics* **108**, 113–131.
- [20] Boero, G. e Marrocu, E. (2004). The performance of SETAR models: a regime conditional evaluation of point, interval and density forecasts. *Int. J. Forecast.* **20**, 305–320.

- [21] Braga, A.L., Zanobetti, A., Schwartz, J. (2001). The time course of weather-related deaths. *Epidemiology* **12**, 662–667.
- [22] Brännäs, K. (1995). Explanatory variables in the AR(1) count data model. *Umeå Economic Studies* 381.
- [23] Brännäs, K. e Hall, A. (2001). Estimation in integer-valued moving average models. *Appl. Stochastic Models Bus. Ind.* **17**, 277–291.
- [24] Brännäs, K. e Hellström, J. (2001). Generalized integer-valued autoregression. *Econometric Rev.* **20**, 425–443.
- [25] Brännäs, K., Hellström, J. e Nordström, J. (2002). A new approach to modelling and forecasting monthly guest nights in hotels. *Int. J. Forecast.* **18**, 9–30.
- [26] Brännäs, K. e Nordström, J. (2006). Tourist accomodation effects of festivals. *Tourism Econ.* **12**, 291–302.
- [27] Campbell, M.J., Julious, S.A., Peterson, C.K. and Tobias, A. (2001). Atmospheric pressure and sudden infant death syndrome in Cook County, Chicago. *Paediatr Perinat Epidemiol.* **15**, 287–289.
- [28] de Maré, J. (1980). Optimal prediction of catastrophes with application to Gaussian process. *Ann. Probab.* **8**, 841–850.
- [29] Doukhan, P., Latour, A. e Oraichi, D. (2006). Simple integer-valued bilinear time series model. *Adv. Appl. Probab.* **38**, 559–578.
- [30] Drost, F.C., van den Akker, R. e Werker, B. (2008). Note on integer-valued bilinear time series models. *Statist. Probab. Lett.* **78**, 992–996.
- [31] Du, J.-G. e Li, Y. (1991). The integer valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.* **12**, 129–142.

- [32] Fu, Q., Fu, H. e Sun, Y. (2004). Self-exciting threshold auto-regressive model (SETAR) to forecast the well irrigation rice water requirement. *Nature and Sciences* **2**, 36–43.
- [33] Franke, J. e Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In, Subba Rao, T. (Ed.), *Developments in Time Series Analysis*. Chapman & Hall, London, 310–330.
- [34] Franke, J. e Subba Rao, T. (1995). Multivariate first order integer valued autoregressions. *Technical Report*. Math Dep., UMIST.
- [35] Franses, P.H. (1994). A multivariate approach to modeling univariate seasonal time series. *J. Econometrics* **63**, 133–151.
- [36] Franses, P.H. e Paap, R. (2004). *Periodic Time Series*. Oxford University Press. Oxford.
- [37] Freeland, R.K.(1988). *Statistical analysis of discrete time series with application to the analysis of Worker's Compensation Claims Data*. Ph.D. thesis. The University of British Columbia, Canada.
- [38] Freeland, R.K. e McCabe, B. (2004). Forecasting discrete valued low count time series. *Int. J. Forecast.* **20**, 427–434.
- [39] Freeland, R.K. e McCabe, B. (2005). Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Statist. Probab. Lett.* **73**, 147–153.
- [40] Garcia-Ferrer, A. e Queralt, R.A. (1997). A note on forecasting international tourism deman in Spain. *Int. J. Forecast.* **13**, 539–549.
- [41] Gardner, W.A., Napolitano, A. e Paura, L., 2006. Cyclostationary, half a century of research. *Signal Process.* **86**, 639–697.
- [42] Gladyshev, E.G. (1961). Periodically correlated random sequences. *Soviet Math.* **2**, 385–388.

- [43] Gladyshev, E.G. (1963). Periodically and almost periodically correlated random processes with continuous time parameter. *Theory Probab. Appl.* **8**, 173–177.
- [44] Gomes, D. (2005). *Processos Autoregressivos de Coeficientes Aleatórios na Modelação de Dados de Contagem*. Tese de Doutoramento, Universidade de Évora, Portugal.
- [45] Gomes, D., Canto e Castro, L. (2009). Generalized integer-valued random coefficient for a first order structure autoregressive (RCINAR) process. *J. Statist. Plann. Inference* **139**, 4088–4097.
- [46] Hall, A. (2001). Extremes of integer-valued moving averages models with regularly varying tails. *Extremes* **4**, 219–239.
- [47] Hall, A., Scotto, M.G. e Cruz, J.P. (2009). Extremes of integer-valued moving average sequences. *Test* (aceite para publicação)
- [48] Haldrup, N., Hyllerberg, S., Pons, G. e Sansó, A. (2007). Common periodic correlation features and the interaction of stocks and flows in daily airport data. *J. Bus. Econom. Statist.* **25**, 21–32.
- [49] Hall, P. e Heyde, C.C. (1980). *Martingale Limit Theory and its Applications*. Academic Press, New York.
- [50] Heyde, C.C. e Lin, Y.X. (1992). On quasi-likelihood methods and estimation for branching processes and heterocedastic regression models. *Aust. J. Stat.* **34**, 199–206.
- [51] Hipel, K.W. e McLeod, A.I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam.
- [52] Hurd, H.L. e Miamee, A. (2007). *Periodically Correlated Random Sequences, Spectral Theory and Practice*. John Wiley & Sons, New Jersey.
- [53] Jacobs, P.A. e Lewis, P.A.W. (1978). Discrete time series generated by mixtures I: Correlational and runs properties. *J. R. Statist. Soc. B* **40**, 222–228.

- [54] Jung, R.C. e Tremayne, A.R. (2006). Binomial thinning models for integer time series. *Stat. Model.* **6**, 81–96.
- [55] Karlin, S. e Taylor, H. (1981). *A Second Course in Stochastic Processes*. Academic Press.
- [56] Katsouyanni, K., Touloumi, G., Samoli, E., Gryparis, A., Le Tertre, A., Monopolis Y. and Rossi, G. (2002). Confounding and effect modification in the short-term effects of ambient particles on total mortality: results from 29 European cities within the APHEA 2 project. *Epidemiology* **12**, 521–531.
- [57] Klimko, L.A. e Nelson, P.I. (1978). On conditional least squares estimation for stochastic process. *Ann. Statist.* **6**, 629–642.
- [58] Koop, G. e Tale, L. (2004). Measuring the health effects of air pollution, to what extent can we really say that people are dying from bad air? *J. Environ. Econ. Manage.* **47**, 30–54.
- [59] Lambert, D. e Liu, C. (2006). Adaptive thresholds: Monitoring streams of network counts. *J. Am. Stat. Assoc.* **101**, 78–88.
- [60] Latour, A. (1988). Existence and stochastic structure of a non-negative integer-valued autoregressive processes. *J. Time Ser. Anal.* **4**, 439–455.
- [61] Latour, A. (1997). The multivariate GINAR(p) process. *Adv. Appl. Probab.* **29**, 228–248.
- [62] Lindgren, G. (1975a). Prediction for a random point time. *Ann. Probab.* **3**, 412–423.
- [63] Lindgren, G. (1975b). Prediction of catastrophes and high level crossings. *Bull. Internat. Statist. Inst.* **46**, 225–240.
- [64] Lindgren, G. (1980). Model processes in non-linear prediction, with application to detection and alarm. *Ann. Probab.* **8**, 775–792.

- [65] Lindgren, G. (1985). Optimal prediction of level crossings in Gaussian processes and sequences. *Ann. Probab.* **13**, 804–824.
- [66] Lund, R.B. e Basawa, I.V. (2000). Recursive prediction and likelihood evaluation for periodic ARMA models. *J. Time Ser. Anal.* **21**, 75–93.
- [67] Lund, R.B., Hurd, H.L., Bloomfield, P. e Smith, R.L. (1995). Climatological time series with periodic correlation. *J. Climate* **11**, 2787–2809.
- [68] Lund, R.B., Shao, Q. e Basawa, I.V. (2006). Parsimonious periodic time series modeling. *Aust. N. Z. J. Stat.* **48**, 33–47.
- [69] Lund, R.B., Wang, X.L., Lu, Q.Q. Reeves, J., Gallagher, C. e Feng, Y. (2007). Changepoint detection in periodic and autocorrelated time series. *J. Climate* **11**, 5178–5190.
- [70] McCabe, B.P.M. e Martin, G.M. (2005). Bayesian prediction of low count time series. *Int. J. Forecast.* **21**, 315–330.
- [71] McDonald, I.L. e Zucchini, W. (1997). Hidden Markov and other models for discrete-valued time series. *Monographs on Statistics and Applied Probability* **70**. Chapman & Hall, London.
- [72] McKenzie, E. (1986). Autoregressive-moving average processes with negative-binomial and geometric marginal distributions. *Adv. Appl. Probab.* **18**, 679–705.
- [73] McKenzie, E. (1988). Some ARMA for dependent sequences of Poisson counts. *Handbook of Statistics* **20**, 822–835.
- [74] McKenzie, E. (2003). Discrete variate time series. In, Shanbhag, D.N. Rao, C.R. ed., *Handbook of Statistics*. Elsevier Science, 573–606.
- [75] Meyn, S.P. e Tweedie, R.L. (1996). *Markov Chains and Stochastic Stability*. Springer-Verlag, London.

- [76] Monteiro, M., Pereira, I. e Scotto, M.G. (2008). Optimal alarm systems for count processes. *Comm. Statist. Theory Methods* **37**, 3054–3076.
- [77] Monteiro, M., Scotto, M.G. e Pereira, I. (2010). Integer-valued autoregressive processes with periodic structure. *J. Statist. Plann. Inference* **140**, 1529–1541.
- [78] Neal, P. e Subba Rao, T. (2007). MCMC for integer-valued ARMA processes. *J. Time Series Anal.* **28**, 92–110.
- [79] Nordström, J. (1996). Tourism satellite account for Sweden 1992-93. *Tourism Econ.* **2**, 13–42.
- [80] Quoreshi, A.M.M.S. (2006). Bivariate time series modelling of financial count data. *Comm. Statist. Theory Methods* **35**, 1343–1358.
- [81] Pai, J. e Pedersen, H.W. (1999). Threshold models of the term structure of interest rate. *Proceedings of the 9th International AFIR Colloquium.* 387–400.
- [82] Paulino, C.D., Turkman M.A.A. e Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.
- [83] Rohatgi, V.K. (1976). *An Introduction to Probability Theory and Mathematical Statistics*. John Wiley & Sons Inc.
- [84] Ross, S.M. (1983). *Stochastic Processes*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., Canada.
- [85] Salas, J.D. (1993). Analysis and modeling of hydrologic time series. In, Maidment, D.R. (Ed.), *Handbook of Hydrology*, Chapter 19. McGraw-Hill, New York.
- [86] Schott, J. R. (1997). *Matrix Analysis for Statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., London.
- [87] Schwartz, J., Levin, R. e Hodge, K. (1997). Drinking water turbidity and paediatric hospital use for gastrointestinal illness in Philadelphia. *Epidemiology* **8**, 615–620.

- [88] Scotto, M.G. e Guedes Soares, C. (2000). Modelling the long-term time series of significant wave height with non-linear threshold models. *Coastal Engineering* **40**, 313–327.
- [89] Silva, M.E. e Oliveira, V.L. (2004). Difference equations for the higher-order moments and cumulants of the INAR(1) model. *J. Time Ser. Anal.* **25**, 317–333.
- [90] Silva, I. (2005). *Contributions to the analysis of discrete-valued time series*. Tese de Doutoramento, Universidade do Porto, Portugal.
- [91] Silva, I., Silva, M.E., Pereira, I. e Silva, N. (2005). Replicated INAR(1) processes. *Methodol. Comput. Appl. Probab.* **7**, 517–542.
- [92] Shao, Q. e Ni, P.P. (2004). Least-squares estimation and ANOVA for periodic autoregressive time series. *Statist. Probab. Lett.* **69**, 287–297.
- [93] Shao, Q. (2006). Mixture periodic autoregressive time series models. *Statist. Probab. Lett.* **76**, 609–618.
- [94] Steutel, F. e Harn, K.V. (1979). Discrete analogues of self-decomposability and stability *Ann. Prob.* **7**, 893–899.
- [95] Svensson, A., Lindquist, R. e Lindgren, G. (1996). Optimal prediction of catastrophes in autoregressive moving average processes. *J. Time Ser. Anal.* **17**, 511–531.
- [96] Svensson, A. e Holst, J. (1997). Optimal prediction of level-crossing in piecewise linear Gaussian processes with changing catastrophe level. *Tech. rep. 1997:14*, Department of Mathematical Statistics, Lund Institute of Technology.
- [97] Svensson, A. e Holst, J. (1998). Prediction of the high water levels in the Baltic. *Istatistik, Journal of the Turkish Statistical Association* **1**, 39–46.
- [98] Tesfaye, Y.G., Meerschaert, M.M. e Anderson, P.L. (2006). Identification of PARMA models and their application to the modeling of riverflows. *Water Resour. Res.* **42**, W01419 (11pp.).

- [99] Thyregod, P., Carstensen, J., Madsen, H. e Arnbjerg-Nielsen, K. (1999). Integer valued autoregressive models for tipping bucket rainfall measurements. *Environmetrics* **10**, 395–411.
- [100] Tong, H. (1977). Some comments on the Canadian lynx data. *J. Roy. Statist. Soc. Ser. A* **140**, 432–435.
- [101] Tong, H. (1990). *Non-linear Time Series: a Dynamical System Approach*. Oxford: Clarendon Press.
- [102] Touloumi, G., Atkinson, R., Le Tertre, A., Samoli, E., Schwartz, J., Schlinder, C., Vonk, J.M., Rossi, G., Saez. M., Rabszenko, D. and Katsouyanni K. (2004). Analysis of health outcome time series data in epidemiological studies. *Environmetrics* **15**, 101–117.
- [103] Vecchia, A.V. (1985). Periodic autoregressive-moving average (PARMA) modelling with applications to water resources. *Water Res. Bull.* **21**, 721–730.
- [104] Weiß, C.H. (2008a). The combined INAR(p) models for time series of counts. *Statist. Probab. Lett.* **78**, 1817–1822.
- [105] Weiß, C.H. (2008b). Thinning operations for modelling time series of counts—a survey. *Adv. Statist. Anal.* **92**, 319–341.
- [106] Zheng, H.T., Basawa, I.V. e Datta, S. (2006). Inference for p th-order random coefficient integer-valued autoregressive processes. *J. Time Ser. Anal.* **27**, 411–440.
- [107] Zheng, H.T., Basawa, I.V. e Datta, S. (2007). First-order random coefficient integer-valued autoregressive processes. *J. Statist. Plann. Inference* **173**, 212–229.
- [108] Zhou, J. e Basawa, I.V. (2005). Least-squared estimation for bifurcation autoregressive processes. *Statist. Probab. Lett.* **74**, 77–88.

- [109] Zhu, R. e Joe, H. (2003). A new type of discrete self-decomposability and its applications to continuous-time Markov processes for modeling count data time series. *Stoch. Models* **19**, 235–254.

