MODELOS INAR(1) E ESTRUTURAIS PARA SÉRIES TEMPORAIS DE CONTAGENS: UM ESTUDO COMPARATIVO, UTILIZANDO AS DISTRIBUIÇÕES POISSON E GEOMÉTRICA.

JOEL MAURÍCIO CORRÊA DA ROSA

ORIENTADOR : Prof. Dr. Luiz Koodi Hotta

Dissertação apresentada ao Instituto de Matemática Estatística e Computação Científica da Universidade Estadual de Campinas para a obtenção de título de Mestre em Estatística

Campinas
1997

UNICAMP
BIBLIOTECA CENTRAL

MODELOS INAR(1) E ESTRUTURAIS PARA SÉRIES TEMPORAIS DE CONTAGENS: UM ESTUDO COMPARATIVO, UTILIZANDO AS DISTRIBUIÇÕES POISSON E GEOMÉTRICA.

Este exemplar corresponde a redação final da dissertação devidamente corrigida e defendida pelo Sr. Joel Maurício Corrêa da Rosa e aprovada pela Comissão Julgadora.

Campinas, 13 de novembro de 1997

Prof. Dr. Luiz Koodi Hotta
Orientador

Dissertação apresentada ao Instituto de Matemática Estatística e Computação Científica da Universidade Estadual de Campinas para a obtenção de título de Mestre em Estatística

Dissertação de Mestrado defendida e aprovada em 03 de novembro de 1997

pela Banca Examinadora composta pelos Profs. Drs.

Prof (a). Dr (a). LUIZ KOODI HOTTA

Prof (a). Dr (a). PEDRO LUIZ VALLS PEREIRA

Prof (a). Dr (a). CRISTIANO AUGUSTO COELHO FERNANDES

Dedico este trabalho a Joel Corrêa da Rosa e Maria Luíza da Costa Rosa, meus pais. Nada seria possível sem eles

AGRADECIMENTOS

- Ao Prof. Dr. Luiz Koodi Hotta pelo total apoio, compreensão, paciência e orientação.
- A minha esposa, Cida, por toda ajuda, carinho e incentivo que foram fundamentais na conclusão deste trabalho.
- Aos meus pais e irmãos pelo incentivo, orações e auxílio nas horas mais difíceis.
- Ao Departamento de Estatística da Universidade Federal do Paraná pela compreensão e o tempo concedido para a realização deste trabalho.
- Ao Prof. Dr. Cristiano Fernandes da PUC-RJ pelas sugestões e idéias.
- Ao CNPq pelo auxílio financeiro concedido.
- Ao fiel Deus pelo amor e proteção.

RESUMO

Analisar séries temporais constituídas por dados de contagens requer a utilização de técnicas estatísticas mais específicas, principalmente em situações nas quais a magnitude das observações é pequena. Dentre as várias propostas emergentes para a análise destas séries, duas são apresentadas nesta dissertação; os modelos INAR(1) (Al-Osh & Alzaid, 1987) e os modelos estruturais para dados de contagens (Harvey e Fernandes, 1989). As duas metodologias representam adaptações de modelos originalmente Gaussianos , à situações envolvendo dados de contagens. No presente trabalho, inicialmente são mostradas as principais propriedades destes modelos e aspectos de suas distribuições preditivas 1- passo-à-frente, utilizando as distribuições Poisson, geométrica e binomial negativa. Em uma segunda etapa, é realizada uma comparação de caráter empírico entre 3 dos modelos apresentados, utilizando séries reais e simuladas. Como resultados, apresentamos o comportamento de cada modelo em termos de ajuste e características da distribuição preditiva 1-passo-à-frente.

ABSTRACT

Count data in Time Series Analysis must be treated with special techniques, specially when the counts are too small. Among the recent proposals to analyse this kind of data, we present two of them in this work : INAR(1) models (Al-Osh and Alzaid, 1987) e structural models (Harvey e Fernandes, 1989). Both models modify theories originally build to deal with Gaussian data. In the present work, we firstly show the main properties of these models and some aspects of their one-step-ahead predictive distribution, making use of Poisson, binomial and negative binomial distributions. After this, we make an empiric comparison, in real series and simulated ones, among 3 of the models presented. As results of this study, we show the goodness of fit of each model and some aspects of their one- step- ahead predictive distribution, considering real and artificial data.

SUMÁRIO

Capítulo 1 – Introdução 1

Capítulo 2 - Modelos INAR(1) 6

2.1 – Introdução 6
2.2 – Proriedades 8
2.3 - Estimação dos parâmetros 10

2.3.1 - Estimadores de Yule-Walker 11
2.3.2 - Estimadores de Quadrados Mínimos 11
2.3.3 - Estimadores de Máxima verossimilhança 12

2.4 - Distribuição preditiva 1-passo-à-frente 14
2.5 - Auto-decomponibilidade de uma distribuição discreta 15
2.6 - Modelo INAR(1)-Poisson 17

2.6.1 - Distribuição Preditiva 1-passo-à-frente do modelo INAR(1)-Poisson 21
2.6.2 - Estimadores de máxima verossimilhança com o uso da distribuição Poisson 22

2.7 - Modelo INAR(1)-Geométrica 23

2.7.1 - Distribuição preditiva 1-passo-à-frente do modelo INAR(1)-Geométrica 28
2.7.2 - Estimadores de Máxima Verossimilhança do modelo INAR(1)-Geométrica 28

2.8 - Modelo GINAR(1) 30
2.9 - Modelo INAR(p) 34

Capítulo 3 - Modelos Estruturais para dados de contagem 37

3.1 – Introdução 37
3.2 - O modelo Poisson-Gama 40

3.2.1 - Propriedades do modelo Poisson-Gama 44
3.2.2 - Distribuição preditiva 1-passo-à-frente 45
3.2.3 - Estimação por máxima verossimilhança 47
3.2.4 - Correção de Shephard para o modelo Poisson-Gama 48

3.3 - O modelo Binomial Negativa-Beta 50

3.3.1 - Propriedades do modelo Binomial Negativa-Beta 53
3.3.2 - Distribuição preditiva 1-passo-à-frente do modelo Binomial Negativa-Beta 53
3.3.3 - Estimação dos parâmetros do modelo Binomial Negativa-Beta 54

Capítulo 4 - Comparação entre os modelos 55

4.1 – Introdução 55

4.1.1- Escolha dos processos geradores para a simulação das séries 56
4.1.2- Afinidade entre as distribuições preditivas 1-passo-à-frente 57

4.2 - Comparação utilizando séries reais 58

4.2.1 - Série de gols marcados pela Inglaterra em partidas internacionais contra a Escócia 58
4.2.2 - Série de furtos no Hyde Park 63

4.3 - Resultados obtidos para a simulação do processo INAR(1)-Poisson 67

4.3.1.-Ajuste dos modelos 69
4.3.1.1 - Ajuste do modelo INAR(1)-Geométrica ao processo INAR(1)-Poisson simulado 69
4.3.1.2 - Ajuste do modelo Poisson-Gama ao processo INAR(1)-Poisson simulado 71
4.3.2 - Estatísticas principais da distribuição preditiva 1-passo-à-frente 73
4.3.2.1 - Média da distribuição preditiva 1-passo-à-frente 73
4.3.2.2 - Variância da distribuição preditiva 1-passo-à-frente 77
4.3.2.3 - Erro Quadrático Médio da previsão 1-passo-à-frente 80
4.3.3 - Afinidade entre as distribuições preditivas 1-passo-à-frente 83

4.4 - Resultados obtidos a partir da simulação do processo INAR(1)-Geométrica 87

4.4.1 - Ajuste dos modelos 87
4.4.1.1 - Ajuste do modelo INAR(1)-Poisson ao processo INAR(1)-Geométrica simulado 88
4.4.1.2 - Ajuste do modelo Poisson-Gama ao processo INAR(1)-Geométrica simulado 90
4.4.2 - Estatísticas principais da distribuição preditiva 1-passo-à-frente 91
4.4.2.1 - Média da distribuição preditiva 1-passo-à-frente 91
4.4.2.2 - Variância da distribuição preditiva 1-passo-à-frente 97
4.4.2.3 - Erro Quadrático Médio da previsão 1-passo-à-frente 102
4.4.3 - Afinidade entre as distribuições preditivas 1-passo-à-frente 106

4.5 - Resultados obtidos a partir da simulação do processo Poisson-Gama 111

4.5.1 - Ajuste dos modelos 111
4.5.1.1 - Ajuste do modelo INAR(1)-Poisson ao processo Poisson-Gama simulado 111
4.5.1.2 - Ajuste do modelo INAR(1)-Geométrica a processos Poisson-Gama simulados 112
4.5.2 - Principais estatísticas da distribuição preditiva 1-passo-à-frente 113
4.5.2.1 – Média da distribuição preditiva 1-passo-à-frente 113
4.5.2.2 – Variância da distribuição preditiva 1-passo-à-frente 115
4.5.2.3 - Erro Quadrático Méido da previsão 1-passo-à-frente 116
4.5.3 - Afinidade entre as distribuições preditivas 1-passo-à-frente 118

4.6 - Seleção de modelos 120

4.6.1 - Seleção de modelos - simulação do processo INAR(1)-Poisson 120
4.6.2 - Seleção de modelos- simulação do processo INAR(1)-Geométrica 124
4.6.3 - Seleção de modelos - simulação do processo Poisson-Gama 128
4.6.4 Considerações finais sobre a seleção de modelos 129

Capítulo 5 - Conclusões 131

APENDICE A - Resultados para a simulação do processo INAR(1)-Poisson 133

APÊNDICE B - Resultados para a simulação do processo INAR(1)-Geométrica 141

APÊNDICE C - Resultados para a simulação do processo Poisson-Gama 156

APÊNDICE D - Programas em MATLAB 159

REFERÊNCIAS BIBLIOGRÁFICAS 164

LISTA DE TABELAS

Tabela 4.2.1 - Principais estatísticas descritivas da série de gols. 59

Tabela 4.2.2 - Frequência absoluta e relativa (percentual em relação ao valor anterior) de Gols em partidas consecutivas. 60

Tabela 4.2.3 - Ajuste dos modelos à séries de gols marcados pela Inglaterra contra a Escócia 60

Tabela 4.2.4 - Ajuste dos modelos a séries de furtos em Hyde Park 64

Tabela 4.3.1 - Erro relativo médio cometido ao utilizar a média da distribuição preditiva 1-passo-à-frente dos modelos ajustados ao invés da média da distribuição verdadeira. 75

Tabela 4.4.1 - Erro relativo médio da média da distribuição preditiva 1-passo-à-frente dos modelos ajustados em relação à média da distribuição preditiva correta. 95

Tabela 4.6.1 - Freqüência de escolhas , através dos critérios AIC e BIC, de cada modelo em 20 replicações do processo INAR(1)-Poisson. 121

Tabela 4.6.2 - Freqüência de escolhas dos modelos, através dos critérios AIC e BIC, segundo o parâmetro α , em 60 replicações do processo INAR(1)-Poisson simulado 122

Tabela 4.6.3 - Freqüência de escolhas dos modelos, através dos critério AIC e BIC, segundo o parâmetro de locação λ , em 100 replicações do processo INAR(1)-Poisson 123

Tabela 4.6.4 - Freqüência de escolhas, através dos critérios AIC e BIC, de cada modelo nas 20 replicações do processo INAR(1)-Geométrica - séries com pequena correlação serial (ϕ=0,1 e ϕ=0,3). 124

Tabela 4.6.5 - Freqüência de escolhas, através dos critérios AIC e BIC, de cada modelo nas 20 replicações do processo INAR(1)-Geométrica - séries com moderada correlação serial (ϕ=0,5 e ϕ=0,7). 125

Tabela 4.6.6 - Freqüência de escolhas, através dos critérios AIC e BIC, de cada modelo nas 20 replicações do processo INAR(1)-Geométrica - séries com alta correlação serial (ϕ=0,9). 126

Tabela 4.6.7 - Freqüência de escolhas dos modelos, através dos critérios AIC e BIC, segundo a correlação serial ϕ, em 100 replicações do processo INAR(1)-Geométrica 127

Tabela 4.6.8 - Freqüência de escolhas através dos critérios AIC e BIC, segundo o parâmetro de locação , em 100 replicações do processo INAR(1)-Geométrica 127

Tabela 4.6.9 - Frequência de escolhas através dos critérios AIC e BIC, de cada modelo ajustado em 20 replicações do processo Poisson-Gama 128

Tabela 4.6.10 a - Frequência relativa percentual de escolhas corretas -Simulação do processo INAR(1)-Poisson 129

Tabela 4.6.10b - Frequência relativa percentual de escolhas corretas -Simulação do processo INAR(1)-Geométrica 129

Tabela 4.6.10 c - Frequência relativa percentual de escolhas corretas - Simulação do processo Poisson-Gama 130

Tabela A1 - Estimativas médias e desvios-padrões obtidos para os parâmetros no ajuste dos modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama em 20 replicações do processo INAR(1)-Poisson 133

Tabela A2 - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Poisson. 135

Tabela A3 - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Poisson. 136

Tabela A4 - Erro quadrático da previsão 1-passo-à-frente utilizando a média da distribuição preditiva 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Poisson. 137

Tabela A5 - Divergência de Kullback-Leibler entre as distribuições preditivas 1-passo-à-frente do modelo ajustado e do processo simulado - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Poisson 138

Tabela A6 - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Poisson 139

Tabela A7 - Resultados obtidos pelos 3 modelos ajustados no critério BIC - média e desvio padrão em 20 replicações do processo INAR(1)-Poisson. 140

Tabela B1.a - Estimativas médias e desvios-padrões obtidos para os parâmetros no ajuste dos modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama em processos INAR(1)-Geométrica simulados com valores fixos de ϕ = 0,1 e 0,3. 141

Tabela B1.b - Estimativas médias e desvios-padrões obtidos para os parâmetros no ajuste dos modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama em processos INAR(1)-Geométrica simulados com valores fixos de ϕ = 0,5 e 0,7 142

Tabela B1.c - Estimativas médias e desvios-padrões obtidos para os parâmetros no ajuste dos modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama em processos INAR(1)-Geométrica simulados com valores fixos de ϕ = 0,9 142

Tabela B.2a - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidas em 20 replicações do processo INAR(1)-Geométrica (*ϕ=0,1 e ϕ=0,3)* 144

Tabela B.2b - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidas em 20 replicações do processo INAR(1)-Geométrica (*ϕ=0,5 e ϕ=0,7)* *145*

Tabela B.2c - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidas em 20 replicações do processo INAR(1)-Geométrica (*ϕ=0,9)* *145*

Tabela B.3a - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica (*ϕ = 0,1 e ϕ=0,3)*. 146

Tabela B.3b - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica (*ϕ = 0,5 e ϕ=0,7)*. *147*

Tabela B.3c - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica (*ϕ = 0,9)* *147*

Tabela B4.a - Erro quadrático da previsão 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,1$ e $\phi=0,3$) *148*

Tabela B4.b - Erro quadrático da previsão 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,5$ e $\phi=0,7$) *149*

Tabela B4.b - Erro quadrático da previsão 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,9$) *149*

Tabela B5.a - Divergência de kullback leibler entre a distribuição preditiva real e a distribuição preditiva do modelo ajustado. - média e desvio padrão obtidos em 20 eplicações do processo INAR(1)-Geométrica ($\phi=0,1$ e $\phi=0,3$) *150*

Tabela B5.b - Divergência de Kullback Leibler entre a distribuição preditiva real e a distribuição preditiva do modelo ajustado. - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,5$ e $\phi=0,7$) *151*

Tabela B5.c - Divergência de Kullback Leibler entre a distribuição preditiva real e a distribuição preditiva do modelo ajustado. - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,9$) ..... .151

Tabela B6.a - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,1$ e $\phi=0,3$) *152*

Tabela B6.b - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,5$ e $\phi=0,7$) 153

Tabela B6.c - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,9$) 153

Tabela B7.a - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,1$ e $\phi=0,3$) *154*

Tabela B7.b - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,5$ e $\phi=0,7$) *155*

Tabela B7.c - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica ($\phi=0,9$) *155*

Tabela C1 - Estimativas médias e desvios-padrões obtidos para os parâmetros no ajuste dos modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama em processos INAR(1)-Poisson simulados..156

Tabela C2 - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo Poisson-Gama 156

Tabela C3 - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo Poisson-Gama 157

Tabela C4 - Erro quadrático da previsão 1-passo-à-frente utilizando a média da distribuição preditiva 1-passo-à-frente - média e desvio padrão em 20 replicações do processo Poisson-Gama 157

Tabela C5 - Divergência de Kullback-Leibler entre as distribuições preditivas 1-passo-à-frente do modelo ajustado e do processo simulado - média e desvio padrão obtidos em 20 replicações do processo Poisson-Gama. 157

Tabela C6 - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo Poisson-Gama 158

Tabela C7 - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo Poisson-Gama. 158

LISTA DE FIGURAS

Figura 1.1a,b - Ruídos gerados através das distribuições normal e Poisson, ambas com média e variância iguais a 2 2

Figura 2.1.1 - Função de autocorrelação do modelo INAR(1) com valor p = 0,8 9

Figura 2.6.1 - Simulações de tamanho n=150 do processo INAR(1)-Poisson com valor fixo do parâmetro de locação *(λ =1)* 18

Figura 2.6.2 - Simulações de tamanho n=150 do processo INAR(1)-Poisson com valor fixo do parâmetro de locação *(λ =2)* 19

Figura 2.6.3 - Simulações de tamanho n=150 do processo INAR(1)-Poisson com valor fixo do parâmetro de locação *(λ =3)* 20

Figura 2.7.1 - Simulações de tamanho n=150 do processo INAR(1)-Geométrica com valor fixo do parâmetro de locação *(θ=0,1)* 24

Figura 2.7.2 - Simulações de tamanho n=150 do processo INAR(1)-Geométrica com valor fixo do parâmetro de locação *(θ=0,5)* 25

Figura 2.7.3 - Simulações de tamanho n=150 do processo INAR(1)-Geométrica com valor fixo do parâmetro de locação *(θ=0,9)* 26

Figura 2.8.1 - Simulações de tamanho n = 150 do processo GINAR(1), utilizando valor fixo do parâmetro de locação *(θ=0,1)* 31

Figura 2.8.2 - Simulações de tamanho de tamanho n = 150 do processo GINAR(1), utilizando valor fixo do parâmetro de locação*(θ=0,5)* 32

Figura 2.8.3 - Simulações de tamanho de tamanho n = 150 do processo GINAR(1), utilizando valor fixo do parâmetro de locação *(θ=0,9)* 33

Figura 3.2.1 - Simulações de 150 valores do processo Poisson-Gama com valores iniciais a_0 = 10 e b_0 = 1 43

Figura 3.3.1 - Simulações do modelo Binomial Negativa-Beta para o valor fixo ν = 3 e distribuição do parâmetro de estado no instante t = 0, BETA(9,10) 52

Figura 4.2.1 - Número de gols marcados pela Inglaterra contra a Escócia em partidas intrnacionais no período compreendido entre 1872 e 1987 59

Figura 4.2.2 - Comportamento da previsão pontual 1-passo-à-frente dos modelos ajustados à série de gols marcados pela Inglaterra 61

Figura 4.2.3 - Distribuições preditivas 1-passo-à-frente dos modelos ajustados para a série de gols da Inglaterra 62

Figura 4.2.4 - Série de furtos em Hyde Park em Chicago no período compreendido entre janeiro de 1968 e novembro de 1973 63

Figura 4.2.5 - Previsão pontual 1-passo-à-frente obtida pelos modelos na série de furtos em Hyde Park 65

Figura 4.2.6 - Distribuição preditiva 1-passo-à-frente dos modelos ajustados à série de furtos em Hyde Park 66

Figura 4.3.1 - Razão entre a média e a variância do processo INAR(1)-Geométrica 68

Figura 4.3.2 - Estimativas médias em 20 replicações obtidas para o parâmetro ϕ do modelo INAR(1)-Geométrica a partir da simulação do processo INAR(1)-Poisson com parâmetros α e λ 70

Figura 4.3.3 - Estimativas médias em 20 replicações obtidas para o parâmetro θ do modelo INAR(1)-Geométrica a partir da simulação do processo INAR(1)-Poisson com parâmetros α e λ 70

Figura 4.3.4 - Estimativas médias em 20 replicações obtidas para o hiperparâmetro ω do modelo Poisson-Gama a partir da simulação do processo INAR(1)-Poisson com parâmetros α e λ. 71

Figura 4.3.5 - Diagrama de dispersão entre as médias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a média verdadeira na simulação do processo INAR(1)-Poisson 74

Figura 4.3.6 - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Poisson 78

Figura 4.3.7a - Comportamento do erro quadrático - média em 20 replicações - obtido ao utilizar como previsão pontual 1-passo-à-frente a média das distribuições preditivas 1-passo-à-frente dos modelos ajustados ao processo INAR(1)-Poisson simulado 81

Figura 4.3.7b- Comportamento do erro quadrático - média em 20 replicações - obtido ao utilizar como previsão pontual 1-passo-à-frente a média das distribuições preditivas 1-passo-à-frente dos modelos ajustados ao processo INAR(1)-Poisson simulado 82

FIGURA 4.3.8a - Divergência de Kullback-Leibler - média em 20 replicações - entre as distribuições preditivas do modelo ajustado e modelo simulado - simulação do processo INAR(1)-Poisson 84

FIGURA 4.3.8b - Divergência de Kullback-Leibler - média em 20 replicações - entre as distribuições preditivas do modelo ajustado e modelo simulado - simulação do processo INAR(1)-Poisson 85

Figura 4.4.1 - Estimativas médias em 20 replicações obtidas para o parâmetro α do modelo INAR(1)-Poisson a partir da simulação do processo INAR(1)- Geométrica com parâmetros ϕ e θ 88

Figura 4.4.2 - Estimativas médias em 20 replicações obtidas para o parâmetro α do modelo INAR(1)-Poisson a partir da simulação do processo INAR(1)- Geométrica com parâmetros ϕ e θ 89

Figura 4.4.3 - Estimativas médias em 20 replicações obtidas para o hiperparâmetro ω do modelo Poisson-Gama a partir da simulação do processo INAR(1)-Geométrica com parâmetros ϕ e θ *90*

Figura 4.4.4a - Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo INAR(1)-Geométrica com θ =0,1 92

Figura 4.4.4b- Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo INAR(1)-Geométrica com θ =0,5. 93

Figura 4.4.4c - Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo INAR(1)-Geométrica com θ =0,9. 94

Figura 4.4.5a - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Geométrica com $\theta = 0.1$ 98

Figura 4.4.5b - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Geométrica com $\theta = 0.5$ 99

Figura 4.4.5c - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Geométrica com $\theta = 0.9$. 100

Figura 4.4.6a - Erro Quadrático de Previsão 1-passo-à-frente - média para os modelos ajustados em 20 replicações do processo INAR(1)-Geométrica simulado. 102

Figura 4.4.6b - Erro Quadrático de Previsão 1-passo-à-frente - média para os modelos ajustados em 20 replicações do processo INAR(1)-Geométrica. 104

Figura 4.4.6c - Erro Quadrático de Previsão 1-passo-à-frente - média para os modelos ajustados em 20 replicações do processo INAR(1)-Geométrica. 105

FIGURA 4.4.7a - Divergência de Kullback-Leibler entre a distribuição preditiva real e as distribuições preditivas dos modelos ajustados 107

FIGURA 4.4.7b - Divergência de Kullback-Leibler entre a distribuição preditiva real e as distribuições preditivas dos modelos ajustados - simulação do processo INAR(1)-Geométrica. 108

FIGURA 4.4.7c - Divergência de Kullback-Leibler entre a distribuição preditiva real e as distribuições preditivas dos modelos ajustados - simulação do processo INAR(1)-Geométrica. 109

Figura 4.5.1 - Parâmetros estimados no ajuste do modelo INAR(1)-Poisson a processos Poisson-Gama simulados - média com base em 20 replicações. 111

Figura 4.5.2 - Parâmetros estimados no ajuste do modelo INAR(1)-Geométrica a processos Poisson-Gama simulados - média com base em 20 replicações 112

Figura 4.5.3 - Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo Poisson-Gama com hiperparâmetro ω =0,9 114

Figura 4.5.4 - Desvios das variâncias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo Poisson-Gama. 115

Figura 4.5.5 - Comportamento do erro quadrático - média em 20 replicações - obtido ao utilizar como previsão pontual 1-passo-à-frente a média das distribuições preditivas 1-passo-à-frente dos modelos ajustados ao processo Poisson-Gama simulado 116

Figura 4.5.6 - Divergência de Kullback-Leibler - média em 20 replicações entre as distribuições preditivas 1-passo-à-frente dos modelos ajustados e a distribuição preditiva real - simulação do processo Poisson-Gama 118

CAPÍTULO 1

INTRODUÇÃO

Os dados de contagem estão presentes no dia-a-dia sendo estes exemplificados com relativa facilidade. Número de pacientes internados em um ambulatório, número de acidentes rodoviários, número de inadimplentes , constituem informações cujo interesse ou curiosidade nos motivam a buscar um tratamento estatístico adequado para sua análise.

A Análise de Séries Temporais, enquanto ferramenta de análise estatística , tem nos modelos ARIMA desenvolvidos por Box & Jenkins (1970) sua técnica mais difundida. Entretanto, a teoria que propiciou o desenvolvimento deste e outros modelos está fortemente amparada no pressuposto de que a distribuição normal fornece a estrutura geradora das inovações. Quando consideramos dados obtidos através de contagens, especialmente aquelas cuja magnitude é pequena, como por exemplo : 3 pacientes internados ao mês, 2 acidentes diários, encontramos dificuldades para a implementação de modelos Gaussianos tais como os modelos ARIMA. Através de um exemplo, podemos visualizar melhor a diferença de comportamento entre uma série temporal com estrutura geradora Gaussiana e outra com estrutura geradora não-Gaussiana, mais especificamente aquelas com distribuição adequadas a dados de contagem, como a distribuição de Poisson ou a binomial negativa. Através das Figuras 1.1a e 1.1b, podemos evidenciar o quanto um processo estocástico de estado contínuo difere de um processo com mesmas média e variância , porém de estado discreto.

Figuras 1.1a,b - Ruídos gerados através das distribuições normal e Poisson, ambas com média e variância iguais a 2.

FIGURA 1.1.a - 500 ruídos gerados através da distribuição normal com média e variância iguais a 2

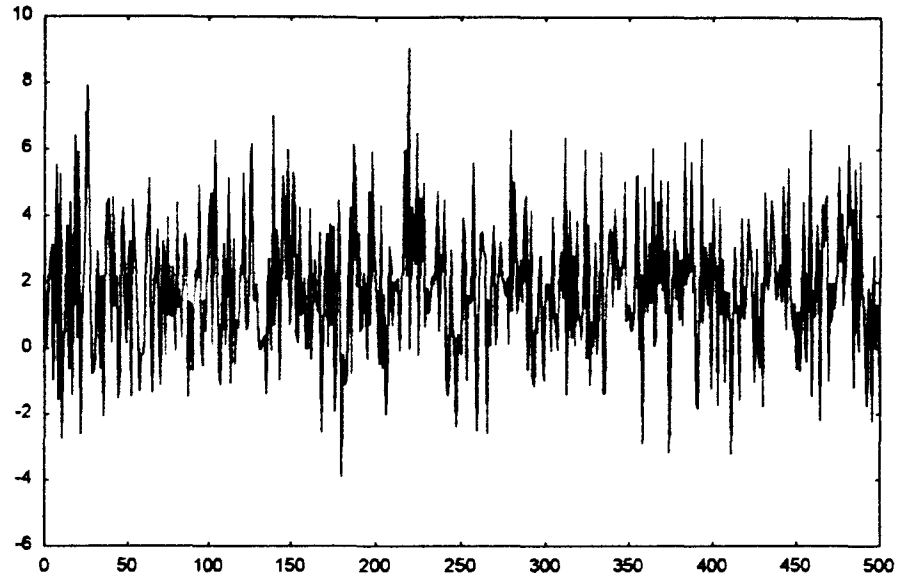
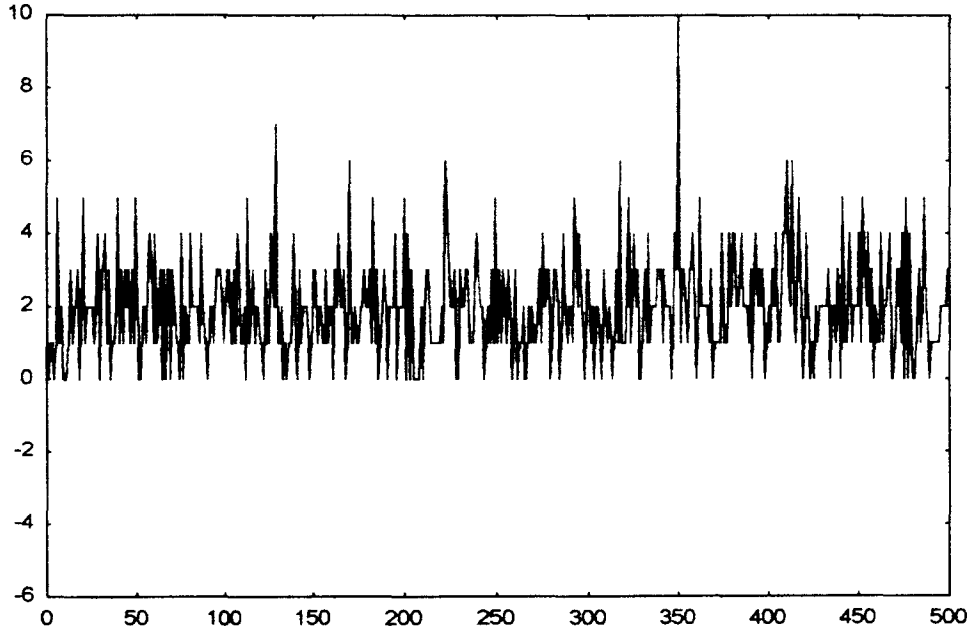


FIGURA 1.1.b - 500 ruídos gerados através da distribuição Poisson com média igual a 2



A Figura 1.1a apresenta o comportamento de um ruído gerado através da distribuição normal com média e variância iguais a 2, enquanto a Figura 1.1b apresenta um ruído gerado através da distribuição de Poisson com média e conseqüentemente variância iguais a 2 . Através deste exemplo, fica ilustrado que embora os dois processos apresentem a mesma média e variância, o processo da Figura 1.1.a apresenta maior oscilação em relação ao processo da Figura 1.1.b . Soma-se a tudo isso o fato do primeiro processo assumir valores negativos, o que não ocorre ao se trabalhar com contagens. Uma clara diferença entre os dois processos pode ser encontrada ao contabilizarmos os pontos de mudança , ou seja, quantidade de picos e vales nas séries. Encontramos no processo exibido na Figura 1.1.a um total de 337 pontos de mudança, enquanto no ruído gerado através da distribuição Poisson são encontrados apenas 233, não contabilizando as observações empatadas.

Para que os modelos probabilísticos, inicialmente desenvolvidos para analisar dados contínuos, possam ser utilizados em situações envolvendo dados originários de distribuições discretas, recorre-se, na maioria das situações, a transformações do tipo potência, tais como a logaritmização e a radiciação. Entretanto, nenhum princípio matemático garante que uma transformação feita com determinado objetivo, traga uma melhoria ao conjunto de dados sob outros aspectos. Uma transformação que melhore o conjunto de dados em um sentido pode piorá-lo em outro (Hoaglin, 1983). Além disto, a re-expressão dos dados, nos leva a uma escala menos familiar que pode dificultar a interpretação dos resultados, principalmente para o usuário. Estes problemas se agravam ainda mais quando a variável aleatória expressa uma contagem pois além de existir a possibilidade de observar um grande número de zeros, esta não assume valores negativos. Estas restrições inviabilizam, em muitas situações, a utilização de modelos Gaussianos.

Com o intuito de preservar a interpretação dos resultados e criar modelos específicos para situações onde a distribuição de probabilidade é supostamente discreta, surgem recentemente, várias propostas para a modelagem de séries temporais constituídas de dados de contagem. Uma tentativa de reproduzir a modelagem de Box & Jenkins (1970) é encontrada em Jacobs & Lewis (1978a,b) , através dos modelos DARMA (Auto-regressivos Médias-móveis Discretos) que representam uma combinação linear aleatória de

de variáveis discretas independentes e identicamente distribuídas. Em Zeger (1988) é encontrada uma proposta para a modelagem de séries temporais de contagens que utiliza um modelo de regressão log-linear para a média do processo, condicionada a um processo latente que introduz a estrutura de autocorrelação entre as observações. Neste modelo é possível a utilização de variáveis explanatórias e a estimação do vetor de parâmetros pode ser feita pelo método da máxima verossimilhança ou por mínimos quadrados ponderados.

Dentre as várias propostas emergentes, duas merecem especial atenção neste trabalho; os modelos INAR(1), auto-regressivos de primeira ordem para valores inteiros, desenvolvidos por Al Osh e Alzaid (1987) e os modelos estruturais, uma adaptação de Harvey e Fernandes (1989) dos modelos originalmente Gaussianos propostos por Harvey (1984), a situações que envolvem dados de contagem. O primeiro dos modelos citados apresenta propriedades semelhantes àquelas dos modelos AR(1) desenvolvidos por Box and Jenkins, baseando a sua construção na estrutura de autocorrelação da série temporal. A outra abordagem segue a linha de modelagem dinâmica, tal como nos modelos de Espaço de Estados, nos quais é permitido aos parâmetros variarem ao longo do tempo , podendo a série ser descrita em função das suas componentes de interesse; sazonalidade, ciclo e tendência, além de utilizar variáveis explanatórias. Entretanto, neste trabalho utilizaremos somente modelos univariados para séries temporais. Utilizando a classificação atribuída por Cox (1981) aos procedimentos de modelagem de uma série temporal, os modelos INAR(1) representam aqueles construídos com base nas observações enquanto a modelagem estrutural é conduzida pelos parâmetros. Em face a esta oposição é natural o questionamento tal aquele encontrado em Kendall (1990) sobre qual modelagem utilizar ARIMA ou Estrutural ?

Adaptando a indagação de Kendall à situações que envolvem dados de contagem, este trabalho não pretende responder a questão de forma absoluta, mas tem como objetivos gerais : a) apresentar os dois enfoques de modelagem como alternativas para analisar séries temporais de contagens, sem que haja a necessidade de re-expressão; b) comparar , de forma empírica, alguns aspectos destes processos de contagens.

O primeiro objetivo é justificado por tratarmos de procedimentos que podem ser implementados computacionalmente com relativa facilidade. Este fato é percebido nos capítulos 2 e 3 que apresentam respectivamente os modelos INAR(1) e Estruturais para dados de contagem, mostrando suas propriedades, métodos para a estimação dos parâmetros e aspectos da distribuição preditiva. O capítulo 4 apresenta a metodologia utilizada para comparar os modelos propostos e traz os resultados obtidos com base em séries reais e simulações. As conclusões do estudo são apresentadas no capítulo 5. Nos apêndices, são apresentadas tabelas mais detalhadas com os resultados e também os programas utilizados em MATLAB para a simulação das séries e obtenção das estimativas para os parâmetros dos modelos utilizados.

CAPÍTULO 2

Modelos INAR(1)

2.1 - INTRODUÇÃO

Os modelos INAR(1), auto regressivos de primeira ordem para valores inteiros, apresentados por Al-Osh e Alzaid (1987), representam uma modificação da metodologia dos modelos de Box & Jenkins (1970) para analisar dados de contagem. Estes modelos são construídos com base na estrutura de autocorrelação da série temporal, utilizando uma convolução de distribuições discretas para se chegar a distribuição preditiva. Nesta seção apresentaremos a forma geral do modelo INAR(1) e suas propriedades e em seguida apresentamos os casos específicos das inovações provenientes das distribuições Poisson e geométrica.

Seja Z uma variável aleatória que assume valores inteiros não-negativos. Para a construção dos modelos INAR(1), antes se faz necessária a definição do operador 'o'. Steutel e Van Harn (1979) definem a ação deste operador em Z como :

$$p \circ Z = \sum_{i=1}^{Z} Y_i$$

na equação acima, $p \in [0,1]$ e Y_i é uma sequência de variáveis aleatórias independentes com a distribuição de Bernoulli(p). Em decorrência da definição do operador, temos :

(a) $0 \circ Z = 0$;

(b) $1 \circ Z = Z$;

(c) $E(p \circ Z) = pE(Z)$;

para $\beta \in [0,1]$, as variáveis $\beta \circ (p \circ Z)$ e $\beta p \circ Z$ apresentam a mesma distribuição de probabilidade.

Utilizando o operador definido anteriormente e sendo Z_t uma variável aleatória discreta que assume valores no conjunto dos inteiros não-negativos, temos o modelo INAR(1) dado por :

$$Z_t = p \circ Z_{t-1} + \varepsilon_t \qquad (2.1.1)$$

onde ε_t ; $t = 1,2,...$, é uma seqüência de variáveis aleatórias discretas I.I.D com média μ , variância σ^2 e $Cov(Z_s, \varepsilon_t) = 0$ para s < t.

No modelo (2.1.1), Z_t é obtido através da soma de duas variáveis aleatórias com distribuições discretas. A primeira delas, $W_t = p \circ Z_{t-1}$, tem distribuição binomial, condicionada ao conhecimento de Z_{t-1} . A segunda variável aleatória, ε_t, é chamada de inovação do processo e esta é geralmente não observável, seguindo qualquer distribuição de probabilidades adequada a contagens tais como a Poisson, binomial negativa, geométrica, entre outras. W_t e ε_t são independentes e assim temos :

$W_t \mid Z_{t-1} \sim Binomial(p, Z_{t-1})$

$P(\varepsilon_t = i) = q_t(i) \quad i=0,1,2,3,.....$

O modelo INAR(1) tem forte apelo interpretativo, visto que seus termos são variáveis aleatórias de uso frequente nas aplicações da estatística no dia-a-dia. Dentre as possibilidades de aplicação do modelo INAR(1) , podemos citar duas:

■ Se considerarmos Z_t como o número de clientes em uma fila de banco no instante t, o termo $p \circ Z_{t-1}$ representa o número de clientes que estiveram na fila no instante t e continuam no instante t-1, enquanto ε_t descreve a entrada aleatória de clientes no intervalo (t-1,t], sendo nesta situação um processo observável.

■ Para Z_t, número de pacientes internados no mês t em um determinado hospital, podemos representar este processo em função dos pacientes que durante o mês imediatamente anterior estavam internados e permanecem no instante atual, somando-se a entrada de pacientes, ε_t, causada por fontes aleatórias.

Pela sua natureza, o modelo INAR(1) é apropriado para situações onde a escala de mensuração dos dados não possui um limite superior. Uma desvantagem que surge neste modelo é que estes são impróprios para a modelagem de uma série binária.

2.2 - PROPRIEDADES

Em decorrência do que foi exposto em (2.1.1), a média e a variância incondicionais do modelo INAR(1) são dadas por :

$$E[Z_t] = p \circ E[Z_{t-1}] + \mu = p^t E[Z_0] + \mu \sum_{j=0}^{t-1} p^j \tag{2.2.1}$$

$$Var[Z_t] = p^2 Var[Z_{t-1}] + pqE[Z_{t-1}] + \sigma^2 =$$

$$= p^{2t} Var[Z_0] + q \sum_{j=1}^{t} p^{2j-1} E[Z_{t-j}] + \sigma^2 \sum_{j=1}^{t} p^{2(j-1)} \tag{2.2.2}$$

A estacionariedade de segunda ordem é obtida a partir do valor inicial com :

a) $E[Z_0] = \dfrac{\mu}{1-p}$ (2.2.3)

b) $Var[Z_0] = \dfrac{(p\mu + \sigma^2)}{1-p^2}$ (2.2.4)

Neste caso, para k inteiro não negativo, temos as funções de autocovariância e autocorrelação fornecidas respectivamente por :

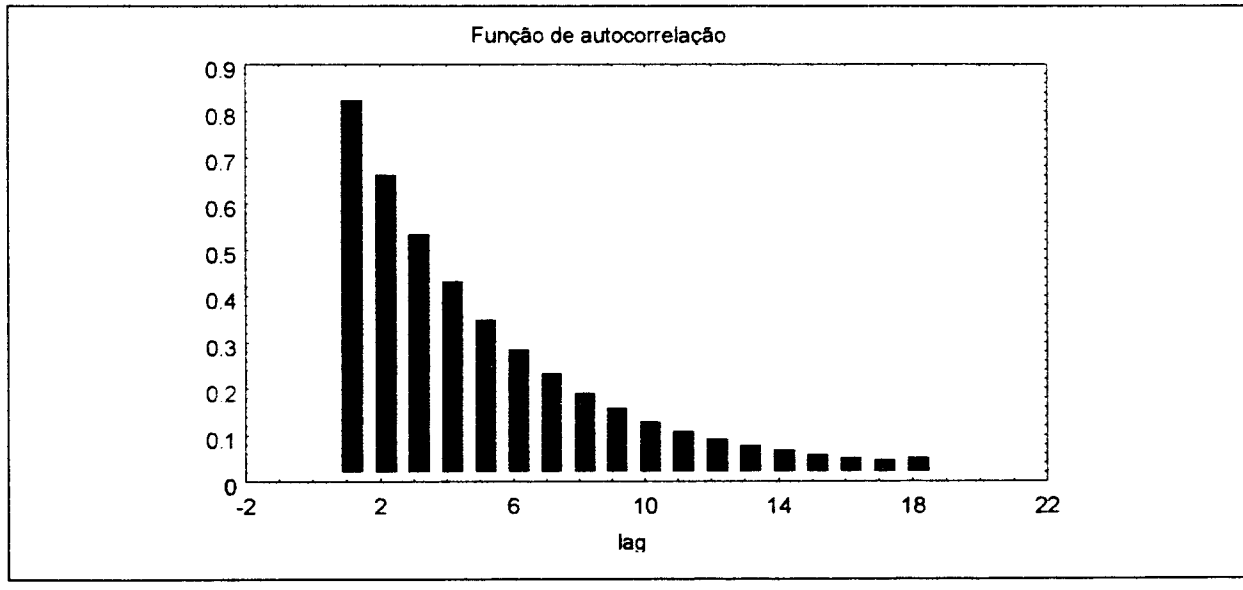
$$\gamma(k) = \operatorname{cov}(Z_t, Z_{t-k}) = Cov(Z_{t-k}, p^k Z_{t-k}) + \operatorname{cov}(Z_{t-k}, \sum_{j=0}^{k-1} p^j \varepsilon_{t-j}) =$$

$$p^k Var(Z_{t-k}) + \sum_{j=0}^{k-1} p^j Cov(Z_{t-k}, \varepsilon_{t-j}) = p^k \gamma(0) \tag{2.2.5}$$

$$\rho(k) = \frac{\gamma(k)}{\gamma(0)} = p^k; \qquad \rho(k) \in [0,1] \tag{2.2.6}$$

Tal como os modelos AR(1), a função de autocorrelação decai exponencialmente com o lag k. Entretanto, ao contrário dos modelos AR, a função assume somente valores positivos. A Figura 2.1 mostra o gráfico da função de autocorrelação do modelo INAR(1) com parâmetro p=0,8.

Figura 2.1.- Função de autocorrelação do modelo INAR(1) com valor de $p = 0,8$



Nos modelos de Box & Jenkins, quando satisfeitas as condições de estacionariedade e invertibilidade, podemos representar o modelo AR(1) através de um modelo MA(∞). De modo análogo, a distribuição marginal de Z_t pode ser escrita em função da distribuição das inovações.

$$Z_t \overset{d}{=} \sum_{j=0}^{\infty} p^j \circ \varepsilon_{t-j}, \qquad \text{para} \quad p \in [0,1] \tag{2.2.7}$$

Como observado acima, a dependência de Z_t na sequência de inovações decai exponencialmente com a defasagem de tempo.

2.3 - ESTIMAÇÃO DOS PARÂMETROS

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

Na presente situação, o procedimento preferencial para a estimativa dos parâmetros do modelo INAR(1) será o uso da estimação por máxima verossimilhança, condicionada ao conhecimento do valor inicial da série, Z_0. Sprott (1983) oferece meios para alcançar a estimativa de máxima verossimilhança dos parâmetros de uma convolução de duas distribuições discretas, em especial da convolução das distribuições de Poisson e binomial. A solução para o vetor que maximiza a função de log-verossimilhanca é encontrada iterando uma equação apenas, já que é possível a eliminação de um elemento do vetor de parâmetros. O procedimento proposto se extende para quaisquer distribuições da família série de potências. Supondo que o modelo INAR(1) esteja completamente especificado através do vetor de parâmetros $\Theta = (p, \mu)$ e que tenha sido observada a série $\mathbf{Z} = (z_1, z_2, \ldots, z_n)$ apresentamos em seguida os estimadores de Yule Walker, Quadrados

mínimos e máxima verossimilhança, estes dois últimos condicionados ao conhecimento do valor inicial Z_0.

2.3.1 - Estimadores de Yule-Walker

Através do método dos momentos, chegamos aos estimadores de Yule-Walker que utilizam as informações da autocovariância e média amostrais.

$$\hat{p} = \frac{\sum_{t=0}^{n-1}(z_t - \bar{z})(z_{t+1} - \bar{z})}{\sum_{t=0}^{n}(z_t - \bar{z})^2} \quad ; \tag{2.3.1}$$

$\bar{z} = \frac{1}{n}\sum_{t=1}^{n} z_t$ é a média aritmética da série e desta forma, obtemos uma estimativa para a média das inovações conforme o resultado abaixo :

$$\hat{\mu} = \frac{\sum_{j=1}^{t}\hat{\varepsilon}_j}{n} \quad ; \quad \hat{\varepsilon}_j = z_t - \hat{p}z_{t-1} \tag{2.3.2}$$

2.3.2 - Estimadores de Quadrados Mínimos

A estimativa de quadrados mínimos para o vetor de parâmetros $\Theta = (p, \mu)$ é encontrada minimizando a soma dos quadrados dos erros obtidos através do ajuste do modelo :

$$E(Z_t | Z_{t-1}) = pZ_{t-1} + \mu = g(\Theta, Z_{t-1}) \tag{2.3.3}$$

Estaremos desta forma, encontrando o mímimo para a função :

$$Q_n(\Theta) = \sum_{t=1}^{n} [Z_t - g(\Theta, Z_{t-1})]^2 \tag{2.3.4}$$

Derivando $Q_n(\Theta)$ em relação a p e μ e posteriormente igualando as duas equações a zero, obtemos os seguintes resultados :

$$\hat{p} = \frac{\sum_{t=1}^{n} z_t z_{t-1} - (\sum_{t=1}^{n} z_t \sum_{t=1}^{n} z_{t-1}) / n}{\sum_{t=1}^{n} z_{t-1}^2 - (\sum_{t=1}^{n} z_{t-1})^2 / n} \tag{2.3.5}$$

$$\hat{\mu} = \frac{1}{n}(\sum_{t=1}^{n} z_t - \hat{p}\sum_{t=1}^{n} z_{t-1}) \tag{2.3.7}$$

Sobre as propriedades destes estimadores, podemos encontrar em Klimko e Nelson (1978) que quando as inovações seguem a distribuição Poisson, estes são consistentes e o vetor $(\hat{p}, \hat{\mu})$ têm distribuição assintótica normal bivariada.

2.3.3 - Estimadores de Máxima Verossimilhança

Al-Osh e Alzaid (1987) propõe duas alternativas para a obtenção das estimativas dos parâmetros do modelo INAR(1) através a maximização da função de verossimilhança : estimadores condicionados e não condicionados ao conhecimento do valor incial Z_0. Segundo a forma não condicional e supondo a distribuição de Poisson no instante t =0, a função de verossimilhança, calculada com base em n+1 observações é escrita como :

$$L(\mathbf{z}, \Theta) = [\prod_{t=0}^{n} P_t(z_t)] \frac{[\mu / (1-p)]^{z_0}}{z_0!} e^{-\frac{\mu}{(1-p)}} \quad ; \mathbf{Z} = (z_0, z_1, \ldots, z_n) \tag{2.3.8}$$

$$P_t(z_t) = P(Z_t = z_t | Z_{t-1} = z_{t-1})$$

Alternativamente, a função pode ser maximizada de forma condicional ao conhecimento prévio de z_0.

$$L(\mathbf{z},\Theta|z_0) = \prod_{t=1}^{n} P_t(z_t) \tag{2.3.9}$$

As estimativas de máxima verossimilhança são encontradas, derivando $\log L$ em relação aos parâmetros p e μ e posteriormente igualando as duas equações obtidas a zero.

$$\frac{\partial \log L(\mathbf{z},\Theta)}{\partial p} = 0$$

$$\frac{\partial \log L(\mathbf{z},\Theta)}{\partial \mu} = 0$$

Sprott(1983) descreve meios para encontrar estas estimativas, iterando apenas uma das equações acima, após a eliminação de um dos parâmetros, escrevendo, por exemplo, p em função de μ. O autor demonstra este procedimento para a convolução das distribuições Poisson e binomial e comenta que é possível extendê-lo para qualquer convolução de duas distribuições da família série de potências , como por exemplo a binomial e a geométrica. Um estudo empírico é realizado em Al-Osh e Alzaid (1987) para determinar qual o melhor dentre os métodos apresentados quando as inovações são provenientes da distribuição Poisson. O resultado obtido demonstra que o método de máxima verossimilhança é aquele que fornece as melhores propriedades para os estimadores.

2.4 - DISTRIBUIÇÃO PREDITIVA 1-PASSO-À-FRENTE

A distribuição preditiva um-passo-à-frente do modelo INAR(1) resulta da convolução da distribuição binomial e uma distribuição para uma variável aleatória Q. discreta , tal que $P(Q_t = i) = q_t(i)$; $i = 0,1,2,3,\ldots$. Como resultado, temos :

$$P(Z_t = z | Z_{t-1} = m) = \sum_{i=0}^{min(z,m)} q_t(z-i)\binom{m}{i} p^i (1-p)^{m-i} \qquad (2.4.1)$$

A média da distribuição preditiva pode ser utilizada como previsão pontual 1-passo-à-frente e esta segue o mesmo formato das previsões obtidas nos modelos de Box & Jenkins :

$$E(Z_t | Z_{t-1} = z_{t-1}) = pz_{t-1} + \mu$$

A previsão m-passos a frente é obtida através de :

$$E(Z_{t+m} | Z_t = z_{t-1}) = E(p \circ z_{t+m-1} | Z_t = z_t) + \mu$$

Uma propriedade desejável para o processo INAR(1) é a obtenção, tal como acontece no clássico AR(1) das mesmas distribuições de probabilidade para a marginal e as inovações. Nos modelos INAR(1), apenas a escolha da distribuição Poisson, dentre as mais conhecidas, para as inovações faz com que seja obtida esta mesma distribuição marginal para o processo Z_t , devido as propriedades de auto-decomponibilidade e estabilidade desta distribuição (Steutel e Van Harn, 1979).

2.5 - AUTO-DECOMPONIBILIDADE DE UMA DISTRIBUIÇÃO DISCRETA

Em Steutel e Van Harn (1979), encontramos a adaptação da propriedade de auto-decomponibilidade de uma distribuição para o caso envolvendo variáveis aleatórias discretas.

Uma distribuição de probabilidade para uma variável aleatória discreta Z assumindo valores no conjunto dos inteiros não negativos é auto-decomponível se a sua função geratriz de probabilidade pode ser escrita da seguinte forma :

$$G_Z(s) = G_Z(1-\beta+\beta s)G_\beta(s) \quad ; \quad |s| \leq 1 \; ; \; \beta \in (0,1) \qquad (2.5.1)$$

$G_\beta(s)$ é uma função geratriz de probabilidade.

Utilizando o operador 'o' definido na seção 2.1, a propriedade de auto-decomponibilidade pode ser representada em termos da variável aleatória Z através da relação :

$$Z = \beta o Z' + Z_\beta$$

$\beta o Z'$ e Z_β são variáveis aleatórias independentes , enquanto Z' e Z têm a mesma distribuição de probabilidade.

Alternativamente, podemos representar Z_n, o processo INAR(1) , em função de seu valor inicial e das inovações :

$$Z_n = p^n \, o Z_0 + \sum_{k=0}^{n-1} p^k \, o \varepsilon_{n-k}$$

A função geratriz deste processo é encontrada através de :

$$G_{z_n}(s) = G_{z_o}(1-p^n+p^n s)\prod_{k=0}^{n-1} G_\varepsilon(1-p^k+p^k s)$$

Utilizando inovações da distribuição Poisson com parâmetro λ e supondo adicionalmente que $Z_0 \sim Poisson(\frac{\lambda}{1-p})$, ou seja, assumindo que o processo é fracamente estacionário, a função geratriz de probabilidades é escrita da seguinte forma:

$$G_{z_n}(s) = e^{\frac{\lambda}{1-p}(p^n s - p^n)} \prod_{k=0}^{n-1} e^{\lambda(p^k s - p^k)}$$

A função geratriz no limite é dada por :

$$G_{z_n}(s) = e^{\frac{\lambda}{1-p}(s-1)}$$

$$n \to \infty$$

A distribuição marginal de Z_n é Poisson com média igual a $\frac{\lambda}{1-p}$ que é a mesma distribuição de Z_{n-1} . Em consequencia da definição do processo INAR(1) na qual $\alpha \circ Z_{n-1}$ e ε_t são independentes, fica desta forma caracterizada a auto-decomponibilidade do processo Z_n, quando utilizada a distribuição Poisson. Podemos verificar que tal propriedade não ocorre para as inovações da distribuição geométrica.

. A distribuição geométrica não sendo auto-decomponível, um processo INAR(1) com distribuição marginal geométrica é obtido através da adoção de um procedimento similar ao discutido em Lawrence (1980) para uma distribuição contínua. Nesta situação, busca-se encontrar a distribuição das inovações que consequentemente acarrete a distribuição marginal geométrica.

Nas próximas seções, apresentamos três possíveis situações para processo INAR(1). Nas duas primeiras são atribuídas as distribuições Poisson e Geométrica para as inovações; denominando estes modelos, respectivamente, por INAR(1)-Poisson e INAR(1)-Geométrica. A terceira formulação apresenta o processo GINAR(1), (McKenzie,1986) , no qual a distribuição marginal do processo é geométrica.

2.6 - O MODELO INAR(1)-POISSON

Supondo que as inovações ε_t são oriundas da distribuição Poisson com parâmetro λ , não se alteram as propriedades do modelo INAR(1) descritas na seção 2.1 e, pela auto-decomponibilidade desta distribuição, é obtida uma seqüência de variáveis dependentes com distribuição marginal Poisson.

A distribuição Poisson é apresentada como um modelo probabilístico para o número de ocorrências de um determinado evento em um intervalo de tempo. Esta distribuição possui apenas um parâmetro que representa a média e a variância, sendo este chamado parâmetro de intensidade. Uma variável aleatória Z_t com distribuição Poisson apresenta a seguinte função massa de probabilidade :

$$P(Z_t = z) = \frac{e^{-\lambda}\lambda^z}{z!} \qquad z = 0,1,2....$$

$$\lambda > 0$$

Definimos o modelo INAR(1)-Poisson como :

$$Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim poisson(\lambda)$$

A média e variância incondicionais deste processo são obtidas através de (2.2.1) e (2.2.2) efetuando as seguintes substituições : $\mu = \lambda$; $\sigma^2 = \lambda$.

$$E(Z_t) = \alpha E(Z_{t-1}) + \lambda$$

$$Var(Z_t) = \alpha^2 Var(Z_{t-1}) + \alpha(1-\alpha)E(Z_{t-1}) + \lambda$$

As figuras 2.6.1, 2.6.2 e 2.6.3 mostram algumas simulações do modelo INAR(1) com inovações seguindo a distribuição Poisson (λ) com valor inicial $Z_0 = \dfrac{\lambda}{1-\alpha}$.

Figura 2.6.1 - Simulações de tamanho n=150 do processo INAR(1)-Poisson com valor fixo do parâmetro de locação $\lambda = 1$.

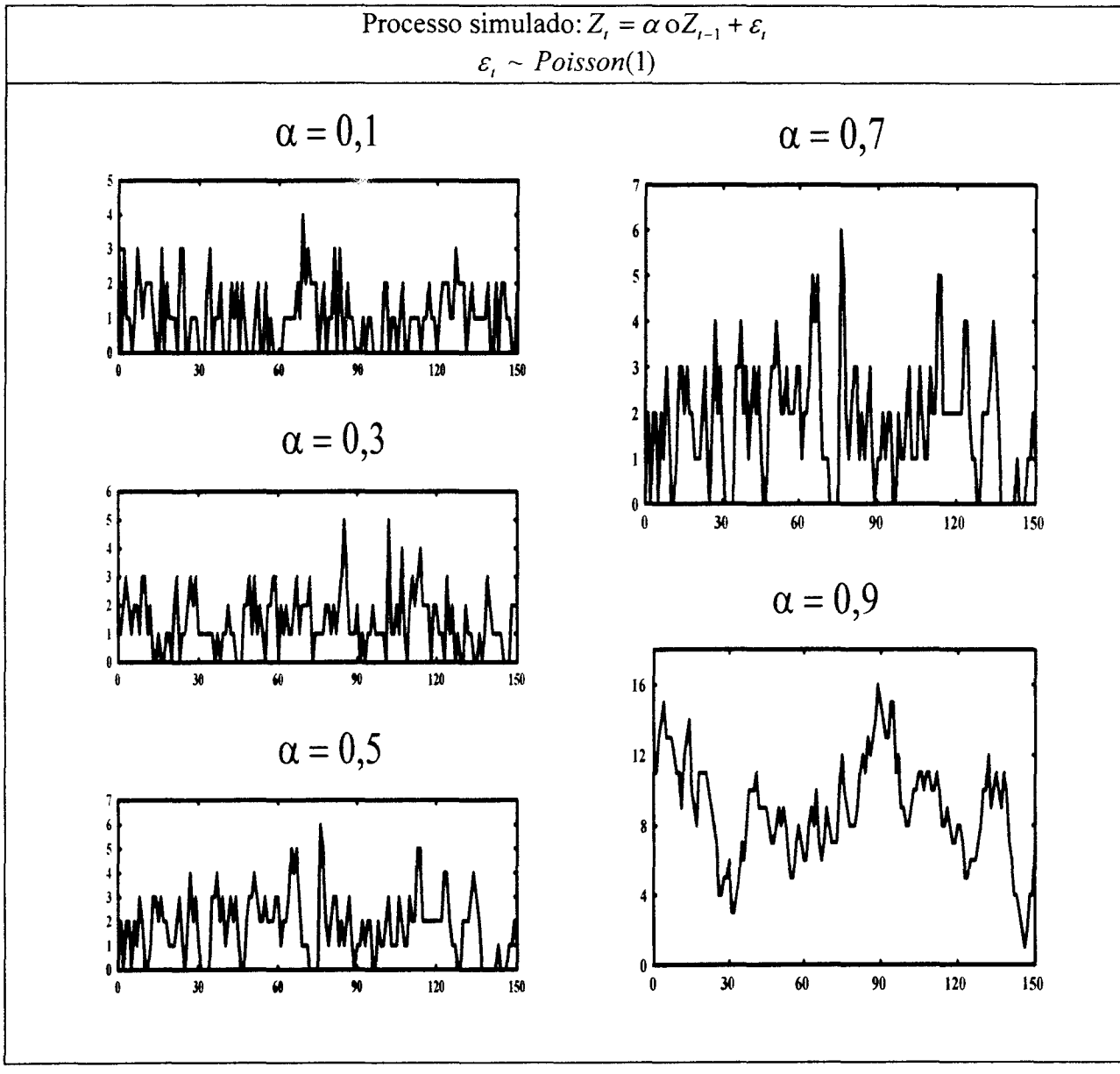


Figura 2.6.2 - Simulações de tamanho n=150 do processo INAR(1)-Poisson com valor fixo do parâmetro de locação $\lambda = 2$.

Processo simulado : $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim poisson(2)$

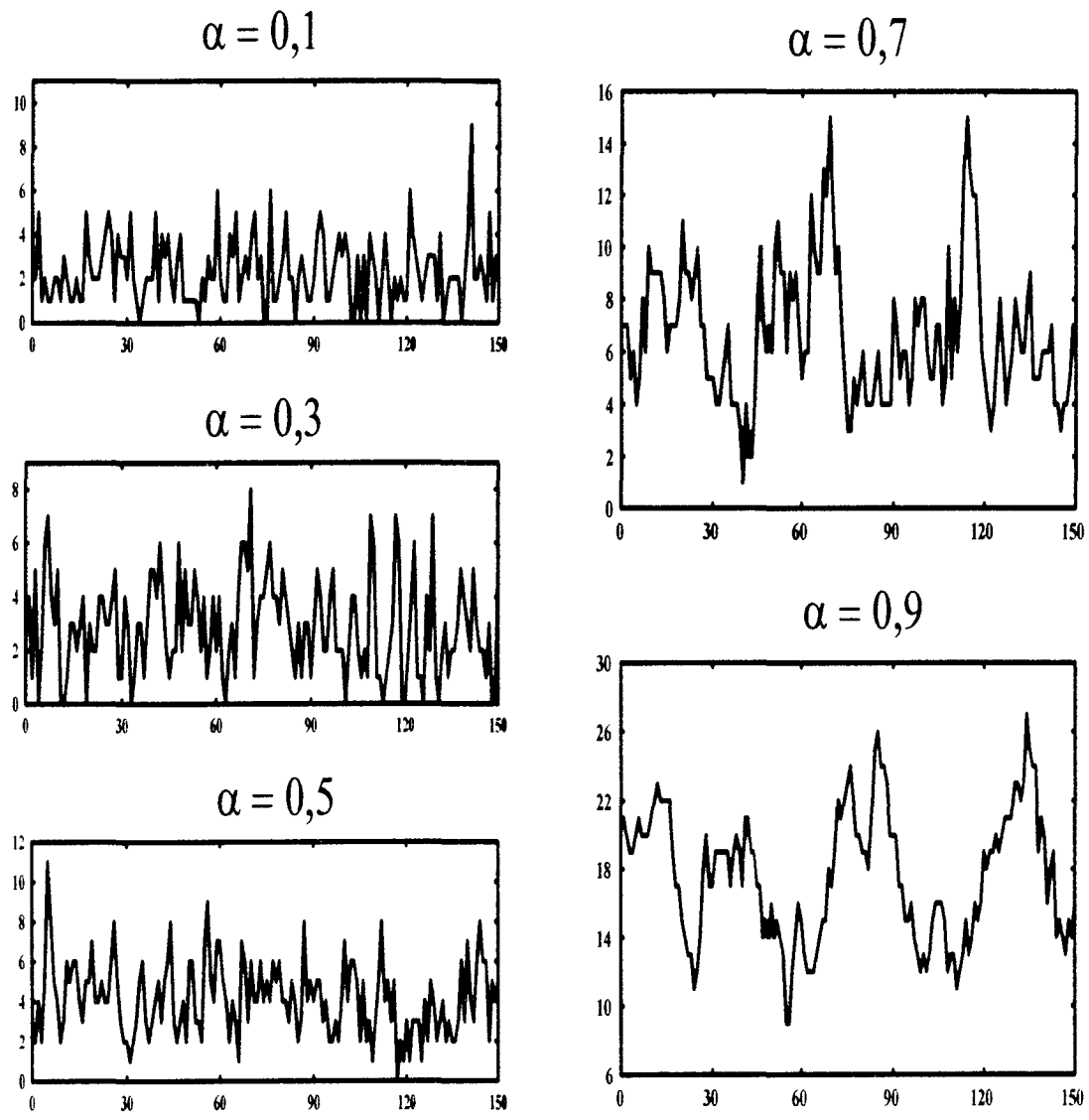
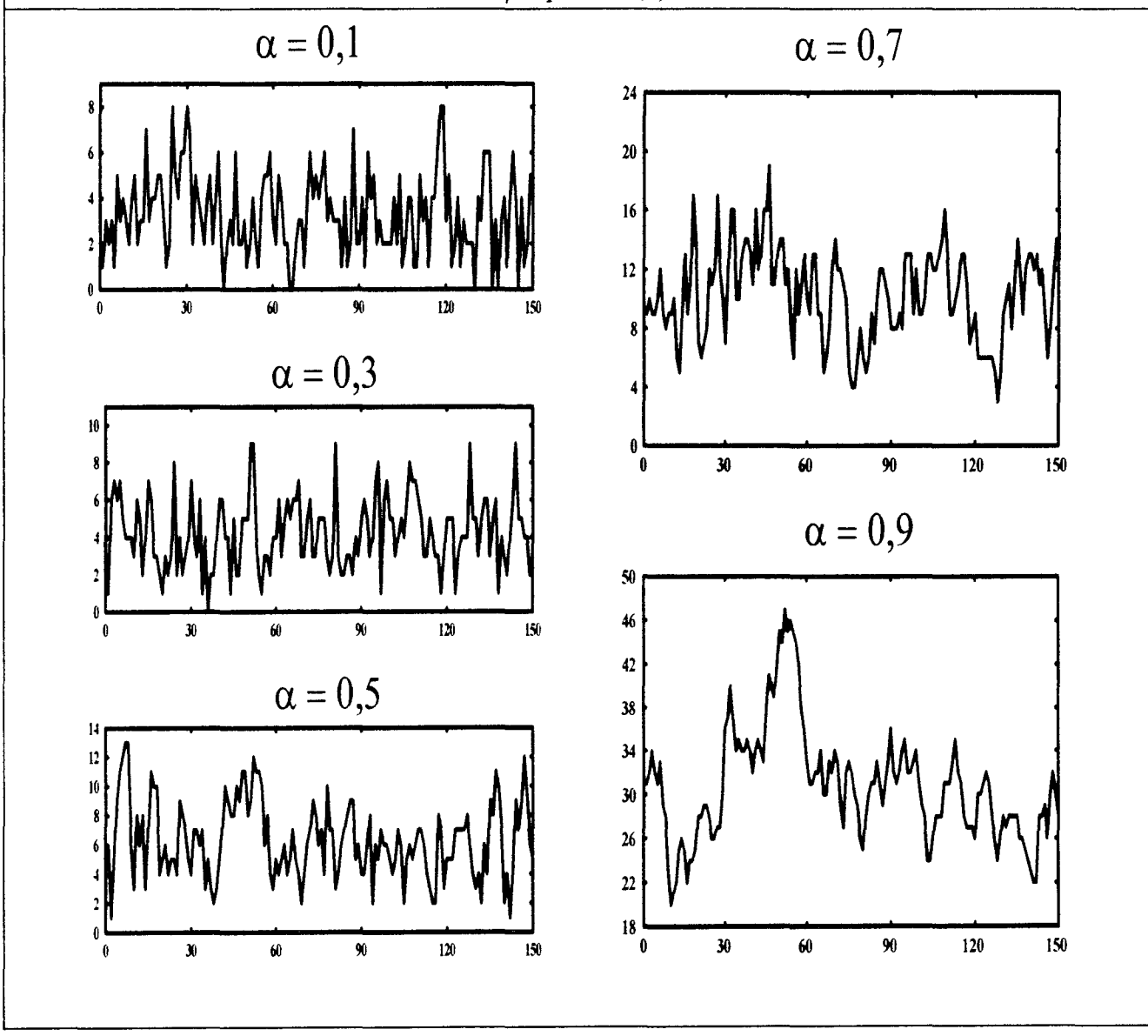


Figura 2.6.3 - Simulações de tamanho n=150 do processo INAR(1)-Poisson com valor fixo do parâmetro de locação $\lambda = 3$.

$$Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$$
$$\varepsilon_t \sim poisson(3)$$



Através das simulações para o processo INAR(1)-Poisson, podemos verificar que este se aproxima mais de um processo estocástico de estado contínuo a medida em que aumentamos o valor do parâmetro α que traz a informação da correlação de primeira ordem entre as contagens . Em processos simulados com valor fixo de $\alpha = 0{,}9$, a série oscila de forma bem mais acentuada, em especial na situação em que o parâmetro de locação assume o valor $\lambda = 3$.

De forma contrária a esta situação, verificamos que o processo que combina os parâmetros $\alpha = 0.1$ e $\lambda = 1$ é extremamente repetitivo; e nesta simulação, o processo assumiu somente quatro estados, apresentando uma grande quantidade de zeros.

2.6.1- Distribuição Preditiva 1-passo-à-frente do modelo INAR(1)-Poisson

A convolução das distribuições binomial e Poisson tem como resultado para a probabilidade em (2.4.1) :

$$P_t(z_t) = \sum_{i=0}^{min(z_t, z_{t-1})} \frac{\exp(-\lambda)\lambda^{z_t - i}}{(z_t - i)!} \binom{z_{t-1}}{i} \alpha^i (1-\alpha)^{z_{t-1}-i}$$

onde $P_t(z_t) = P(Z_t = z_t | Z_{t-1} = z_{t-1})$

Ao utilizarmos a média da distribuição preditiva como previsão pontual 1-passo-à-frente iremos obter :

$$E[Z_t | Z_{t-1}] = \alpha Z_{t-1} + \lambda$$

A variância da previsão 1-passo-à-frente será dada por :

$$Var[Z_t | Z_{t-1}] = \alpha(1-\alpha)Z_{t-1} + \lambda$$

2.6.2 - Estimadores de máxima verossimilhança com o uso da distribuição Poisson

Utilizando os procedimentos da seção 2.3, ou seja, derivando a função de log-verossimilhança em relação aos parâmetros α e λ , obtendo :

$$S_\alpha = \partial \frac{\log L}{\partial \alpha} = \sum_{t=1}^{n} H(t) - n \tag{2.6.1}$$

$$S_\lambda = \partial \frac{\log L}{\partial \lambda} = \sum_{t=1}^{n} \frac{(z_t - \alpha z_{t-1}) - \lambda H(t)}{\alpha(1-\alpha)} \tag{2.6.2}$$

$$H(t) = \frac{P_t(z_t - 1)}{P_t(z_t)}$$

igualando S_α e S_λ a zero, encontramos a seguinte relação entre os parâmetros α e λ :

$$\lambda = (\sum_{t=1}^{n} z_t - \alpha \sum_{t=1}^{n-1} z_{t-1}) / n$$

Substituindo λ na equação (2.6.2) , o procedimento iterativo para encontrar a estimativa de máxima verossimilhança será feito apenas nesta equação.

2.7 - MODELO INAR(1)-GEOMÉTRICA

A distribuição geométrica é utilizada para uma variável aleatória Z que pode ser definida como número de fracassos até a obtenção do primeiro sucesso em uma sequência de experimentos independentes de Bernoulli com probabilidade θ de sucesso. Desta forma, temos que sua função massa de probabilidade é :

$$P(Z = z) = \theta(1-\theta)^z$$

$$0 < \theta < 1$$

A esperança e variância desta distribuição são dadas respectivamente por :

$$E(Z) = \frac{(1-\theta)}{\theta}$$

$$Var(Z) = \frac{(1-\theta)}{\theta^2}$$

Quando as inovações são provenientes da distribuição geométrica com parâmetro θ, ao contrário do caso da distribuição Poisson, não obtemos aqui uma sequência de variáveis dependentes com distribuição marginal geométrica. O modelo é formulado da seguinte forma:

$$Z_t = \phi \circ Z_{t-1} + \varepsilon_t \qquad \phi \in (0,1)$$

$\varepsilon_t \sim$ geométrica (θ)

A características distribucionais deste processo são obtidas efetuando na seção 2.2, as seguintes substituições :

$$\mu = \frac{(1-\theta)}{\theta}$$

$$\sigma^2 = \frac{(1-\theta)}{\theta^2}$$

$$p = \phi$$

Através das Figuras 2.7.1, 2.7.2 e 2.7.3 , visualizamos algumas simulações do processo INAR(1)-Geométrica com 150 observações utilizando como valor inicial $Z_0 = \frac{(1-\theta)}{\theta(1-\phi)}$, que é o valor esperado de Z no instante t=0 para a obtenção de estacionariedade de segunda ordem.

Figura 2.7.1 - Simulações de tamanho n=150 do processo INAR(1)-Geométrica com valor fixo do parâmetro de locação $\theta=0,1$.

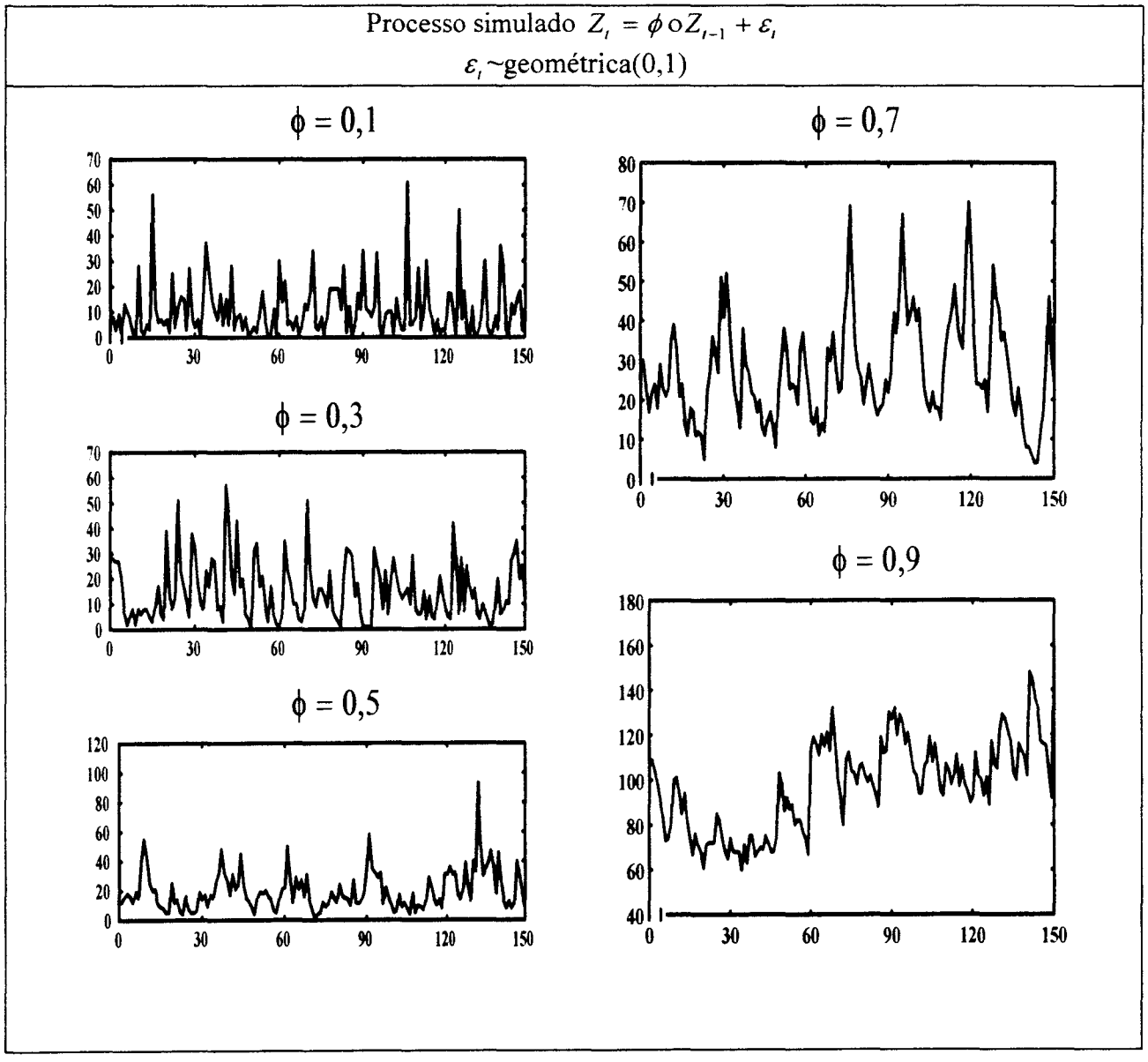


Figura 2.7.2 - Simulações de tamanho n=150 do processo INAR(1)-Geométrica com valor fixo do parâmetro de locação $\theta = 0,5$.

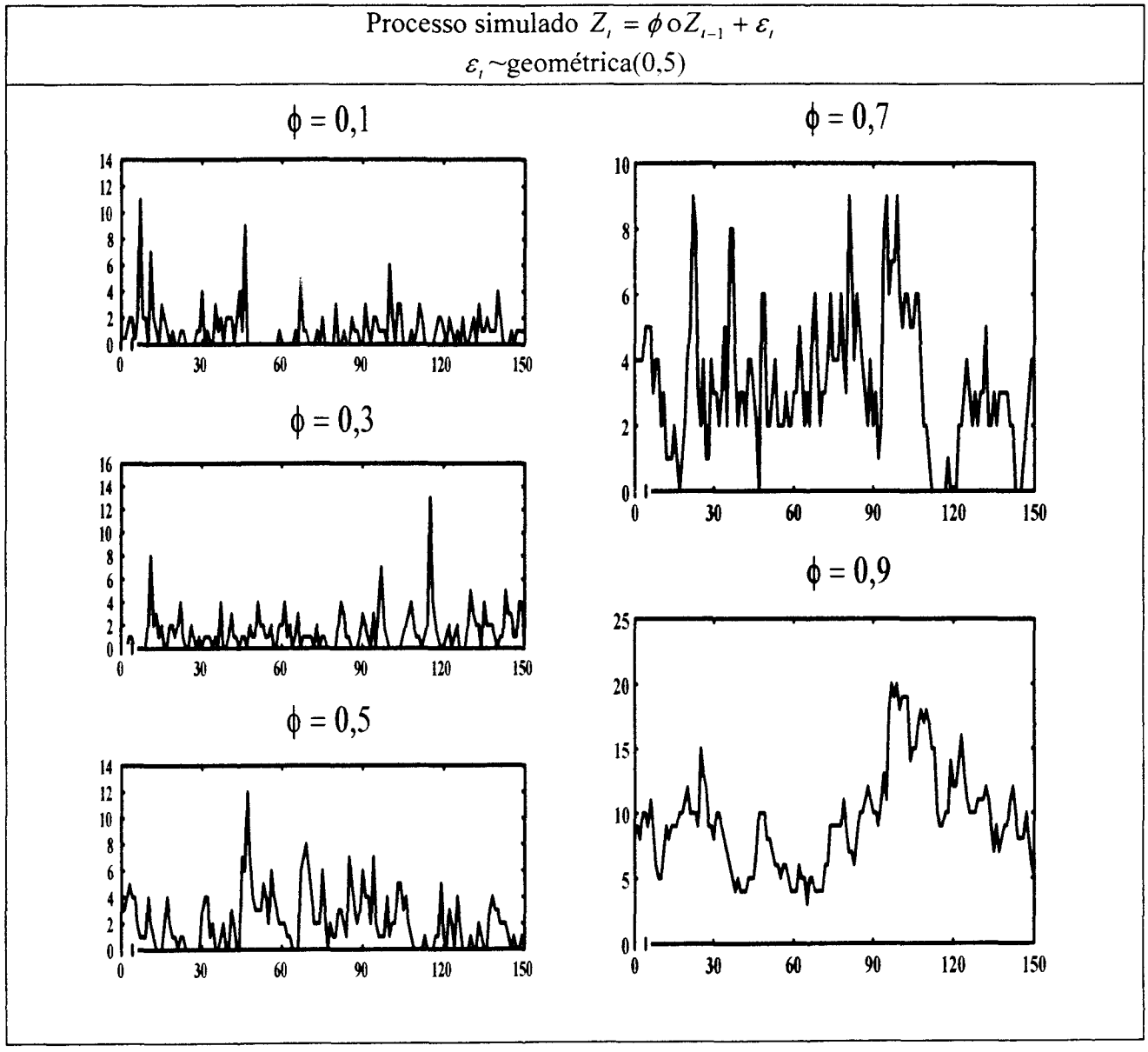
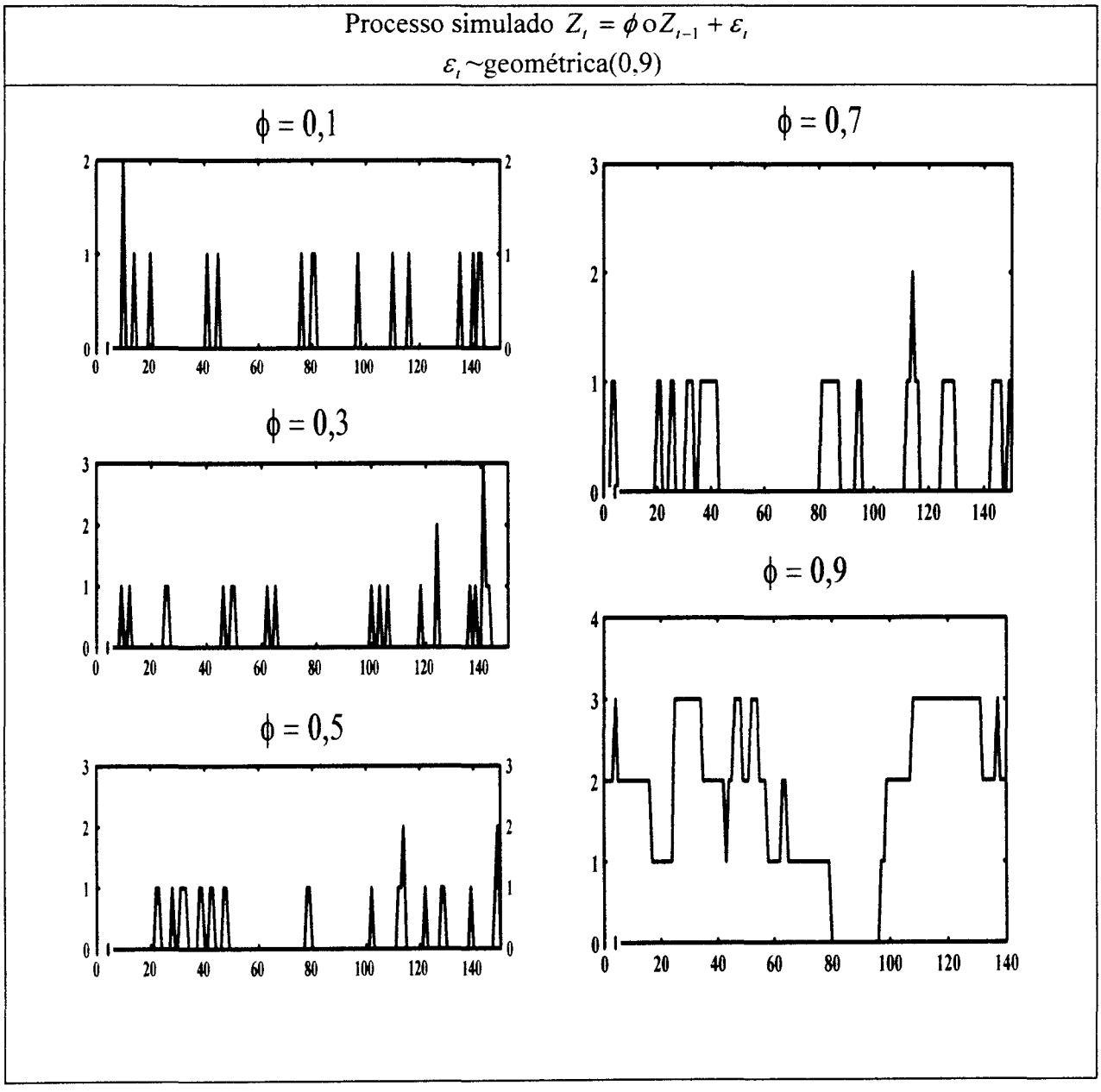


Figura 2.7.3 - Simulações de tamanho n=150 do processo INAR(1)-Geométrica com valor fixo do parâmetro de locação $\theta = 0{,}9$.



A média e variância do processo INAR(1)-Geométrica, ao utilizarmos os valores iniciais em (2.2.3) e (2.2.4) que possibilitam a estacionariedade de segunda ordem a partir do valor inicial , são respectivamente :

$$E(Z_t) = \frac{(1-\theta)}{\theta(1-\phi)}$$

$$Var(Z_t) = \frac{(1-\theta)(1-\phi\theta)}{\theta^2(1-\phi^2)}$$

Uma das principais diferenças deste processo em relação ao processo INAR(1)-Poisson, é que este modelo apresenta sobredispersão .

A escolha dos 3 valores para o parâmetro de locação do modelo INAR(1)-Geomética , θ = 0,1 , 0,5 e 0,9 , faz com que consigamos 3 situações bem distintas; na primeira delas, para θ = 0,1, a probabilidade de obter um sucesso na prova de Bernoulli é pequena e como conseqüência , a probabilidade da entrada de inovações passa a ser alta . Logo, a Figura 2.7.1 mostra processos cuja magnitude dos valores simulados é alta, propiciando a aproximação para a continuidade, em especial quando ϕ = 0,9 e a simulação gera valores que estão entre 60 e 180 , que não são interessantes no contexto deste trabalho. A segunda situação na qual θ = 0,5 . o processo INAR(1)-Geométrica assemelha-se mais aos processos INAR(1)-Poisson mostrados nas Figuras 2.6.1 , 2.6.2 e 2.6.3, principalmente para valores de ϕ = 0,5, 0,7 e 0,9. A principal diferença existente para valores de ϕ = 0,1 e 0,3, são picos que aparecem esporádicamente (Figura 2.7.2) na série. Uma situação particular acontece quando simulamos o processo INAR(1)-Geométrica com parâmetro θ = 0,9 que representa uma probabilidade que pouco favorece a entrada de inovações. As consequências , são processos de magnitude muito pequena, com muitos zeros e conseqüentemente com uma variância pequena. Em casos mais extremos como por exemplo ϕ = 0,1 , a série simulada na Figura 2.7.3 está próxima de uma situação envolvendo dados binários.

2.7.1 - Distribuição Preditiva 1-passo-à-frente do modelo INAR(1)-Geométrica.

A convolução das distribuições binomial e geométrica nos leva ao seguinte resultado para a probabilidade de transição :

$$P_t(z_t) = \sum_{i=0}^{min(z_t, z_{t-1})} (\theta)(1-\theta)^{z_t - i} \binom{z_{t-1}}{i} \phi^i (1-\phi)^{z_{t-1}-i}$$

A média da distribuição preditiva pode ser utilizada como previsão pontual.

$$E(Z_t | Z_{t-1} = z_{t-1}) = \phi z_{t-1} + \frac{(1-\theta)}{\theta}$$

Abaixo, apresentamos o procedimento para a obtenção de estimadores de máxima para o vetor (ϕ, θ).

2.7.2 - Estimadores de máxima verossimilhança do modelo INAR(1)-Geométrica

Obtemos, derivando a função log-verossimilhança em relação aos parâmetros ϕ e θ , os seguintes resultados :

$$\partial \frac{\log L}{\partial \phi} = \sum_{t=1}^{n} \frac{G(i) - \phi Z_{t-1}}{P_t(z_t)\phi(1-\phi)} \tag{2.7.1}$$

$$\partial \frac{\log L}{\partial \theta} = \sum_{t=1}^{n} \frac{P_t(z_t)[1-\theta(z+1)] + \theta G(i)}{P_t(z)\theta(1-\theta)} \tag{2.7.2}$$

onde $G(i) = \sum_{i=0}^{\min(z, z_{t-1})} i(\theta)(1-\theta)^{z_t - i} \binom{z_{t-1}}{i} \phi^i (1-\phi)^{z_{t-1}-i}$

Igualando (2.7.1) e (2.7.2) a zero, encontramos a seguinte relação entre os parâmetros:

$$\theta = \frac{n}{\sum_{t=1}^{n} z_t + n - \phi \sum_{t=1}^{n} z_{t-1}}$$

Sustituindo o valor de θ em 2.7.2, conseguimos realizar o procedimento iterativo para encontrar o zero desta equação apenas. Neste trabalho, as soluções que maximizaram a função de verossimilhança, nos modelos INAR(1)-Poisson e INAR(1)-Geométrica, foram obtidas através da rotina *fzero* presente na pacote computacional MATLAB que possibilita encontrar o zero de uma função de um parâmetro somente

2.8 - O MODELO GINAR(1)

O modelo apresentado na seção anterior, ao contrário do modelo INAR(1)-Poisson, não apresenta distribuição marginal idêntica a distribuição das inovações. Uma seqüência de variáveis aleatórias dependentes com distribuição geométrica é apresentada em McKenzie(1986) sob o título de GINAR(1), idêntica referência é feita em Alzaid e Al-Osh (1988). O processo INAR(1) cuja distribuição marginal é geométrica é obtido com a introdução da seguinte modificação no modelo apresentado anteriormente

$$Z_t = \begin{cases} \alpha \circ Z_{t-1} & \text{com probabilidade } \alpha \\ \alpha \circ Z_{t-1} + \varepsilon_t & \text{com probabilidade } 1-\alpha \end{cases}$$

A distribuição marginal geométrica é obtida para o modelo INAR, através de uma aleatorização feita entre o próprio modelo INAR(1)-Geométrica e apenas o "resgate", com probabilidade α, das observações no instante anterior. Alternativamente, este modelo pode ser representado da seguinte forma :

$$Z_t = \alpha \circ Z_{t-1} + I_t \varepsilon_t$$

$$P(I_t = 0) = 1 - P(I_t = 1) = \alpha$$

I_t é uma sequência de variáveis aleatórias independentes com distribuição de Bernoulli(α). A inovação ε_t é o mesmo processo descrito anteriormente na apresentação do modelo INAR(1)-Geométrica

Seguem abaixo, algumas simulações do modelo GINAR(1), utilizando como valor inicial $Z_0 = \dfrac{(1-\theta)}{\theta(1-\phi)}$ que foi o mesmo utilizado nas simulações do modelo INAR(1)-Geométrica.

Figura 2.8.1 - Séries simuladas de tamanho n = 150 do processo GINAR(1) com valor fixo do parâmetro de locação θ=0,1.

Processo simulado $Z_t = \phi o Z_{t-1} + I_t \varepsilon_t$

$\varepsilon_t \sim$ geométrica(0.1)

$P(I_t = 0) = 1 - P(I_t = 1) = \phi$

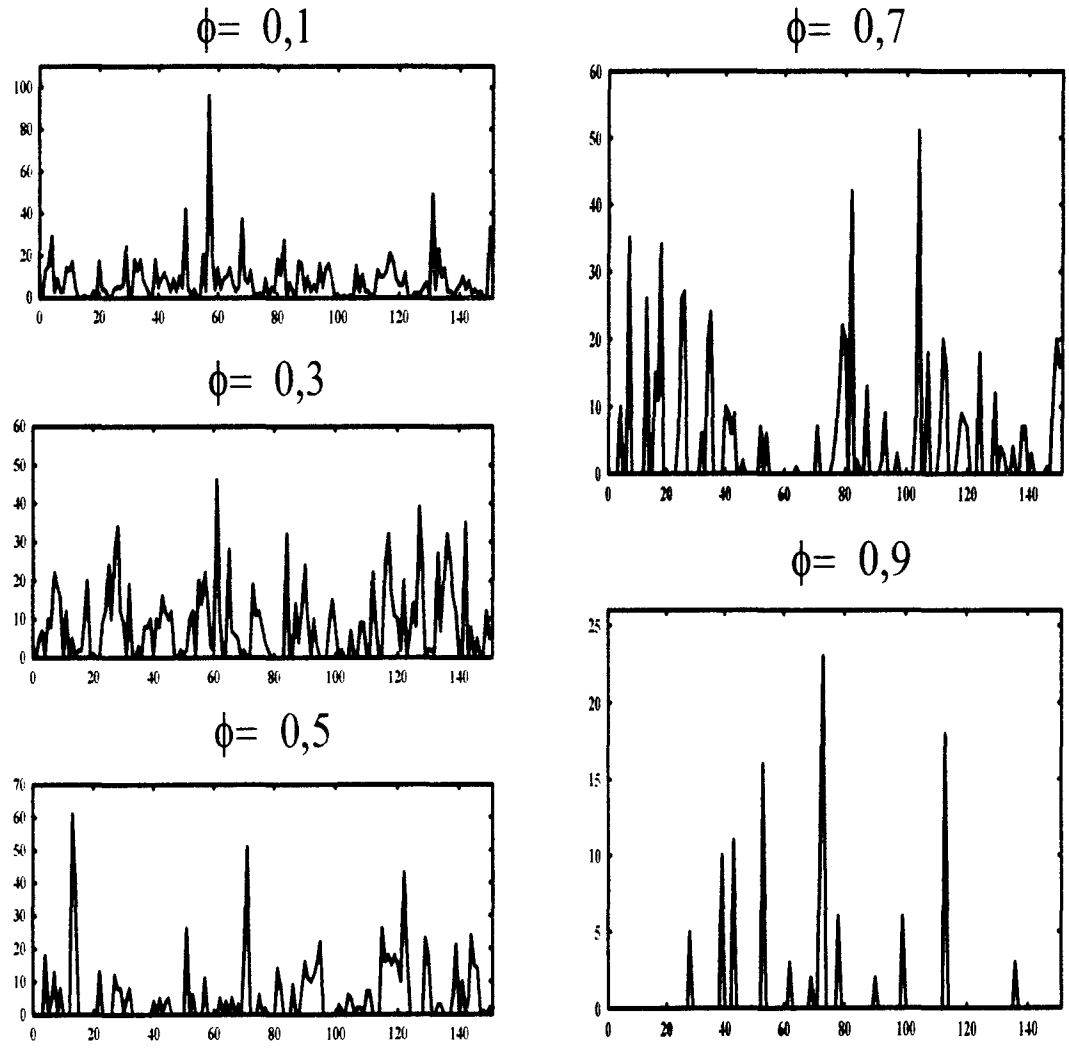


Figura 2.8.2 - Séries simuladas de tamanho n = 150 do processo GINAR(1), utilizando parâmetro de locação θ=0,5

Processo simulado $Z_t = \phi \circ Z_{t-1} + I_t \varepsilon_t$

$\varepsilon_t \sim$ geométrica(0,5)

$P(I_t = 0) = 1 - P(I_t = 1) = \phi$

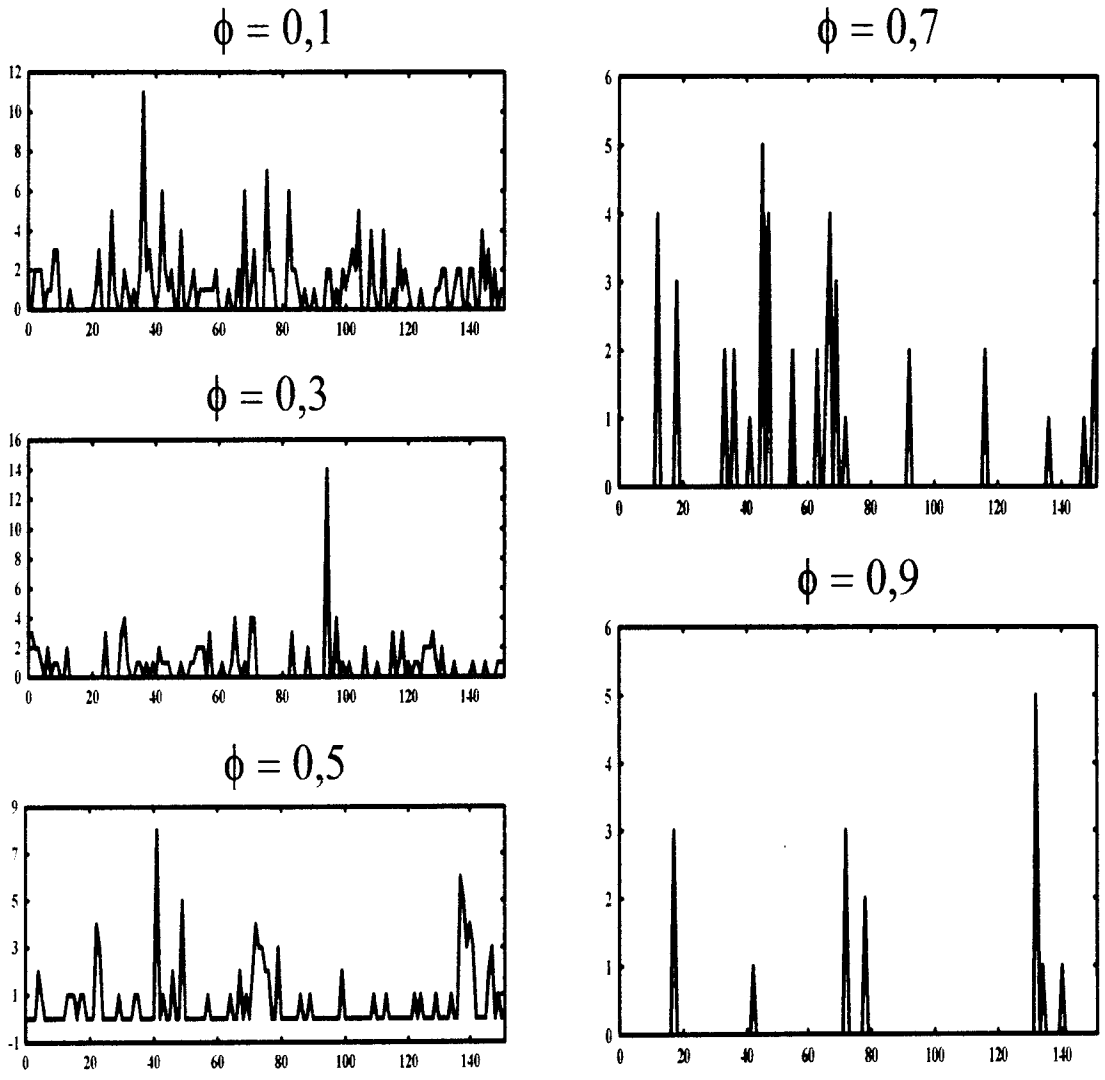
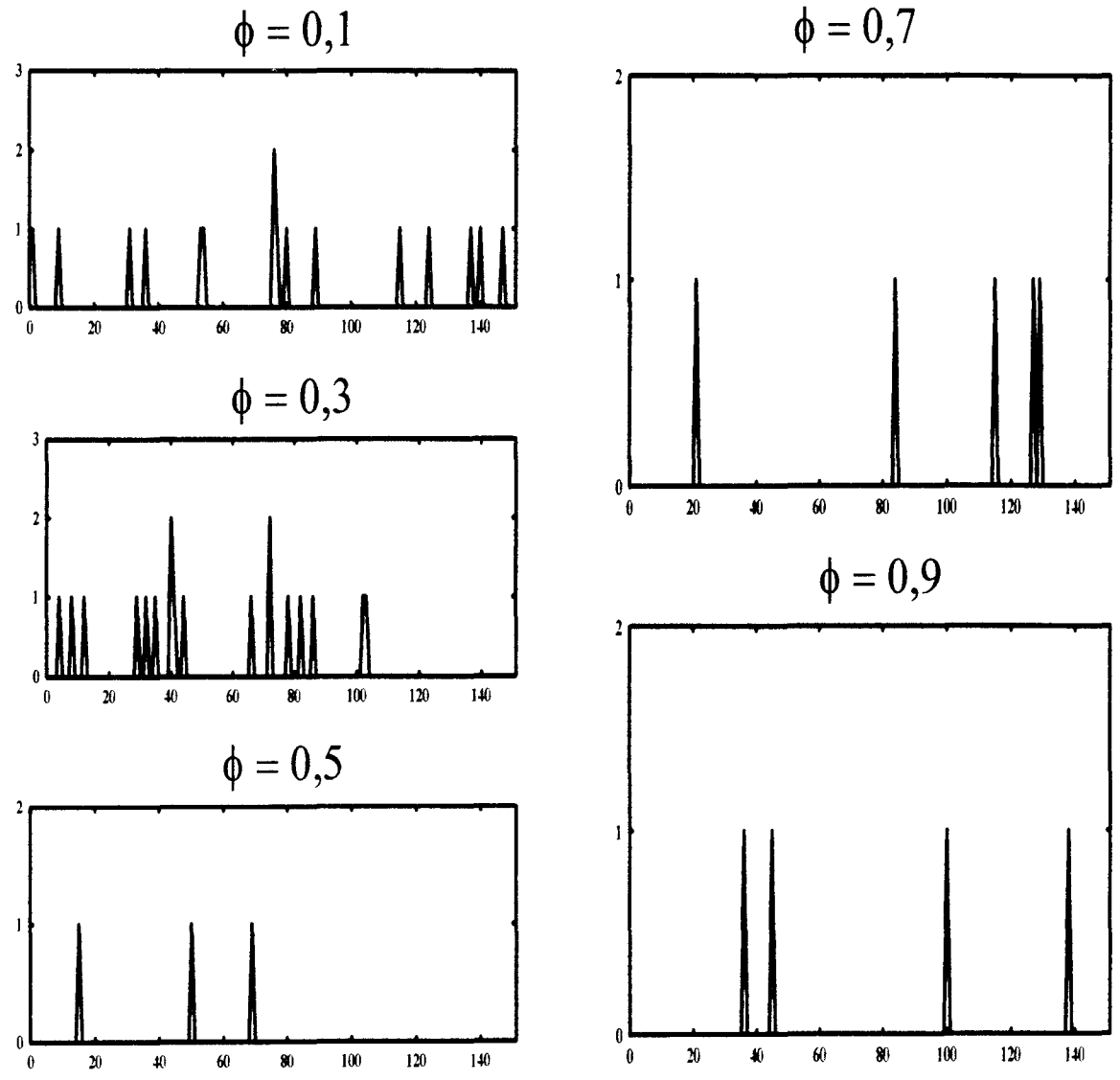


Figura 2.8.3 - Séries simuladas de tamanho n = 150 do processo GINAR(1), utilizando parâmetro de locação θ=0,9

Processo simulado $Z_t = \phi \circ Z_{t-1} + I_t \varepsilon_t$

$\varepsilon_t \sim$ geométrica(0,9)

$P(I_t = 0) = 1 - P(I_t = 1) = \phi$



Através das figuras 2.8.1, 2.8.2 e 2.8.3, podemos identificar nos modelos GINAR(1) aspectos que o diferenciam do modelo INAR(1)-Geométrica. O primeiro deles é que a aleatorização feita nos modelos GINAR(1) introduz no processo uma quantidade maior de zeros em conseqüência do "esquecimento ",com probabilidade ϕ, das inovações ocorridas no passado. Outro aspecto observado nas simulações é que o processo GINAR(1) quando utiliza os mesmos parâmetros do processo INAR(1)-Geométrica , produz observações de maior magnitude. A escolha de parâmetros de locação com valores altos, tal como a figura 2.8.3 que apresenta $\theta = 0{,}9$, faz com que as séries sejam repletas de zeros, dificilmente saindo deste estado para valores altos de ϕ, ou seja, séries com alta correlação serial. Em séries cuja a correlação serial é pequena, os processos simulados INAR(1)-Geométrica e GINAR(1) são semelhantes.

2.9 - MODELOS INAR(p)

A generalização dos modelos propostos por Al-Osh e Alzaid (1987) é encontrada em Jin -Guan & Yuan (1991) que extenderam a ordem auto-regressiva dos modelos INAR e demonstraram que a estrutura de autocorrelação destes modelos é a mesma dos modelos AR(p) propostos por Box & Jenkins. A formulação geral destes modelos é dada por :

$$Z_r = \alpha_1 \circ Z_{t-1} + \alpha_2 \circ Z_{t-2} + \ldots + \alpha_p Z_{t-p} + \varepsilon_t \qquad (2.9.1)$$

$$\alpha_j \in [0,1] \ ; j = 1,2,\ldots,p$$

ε_t são variáveis aleatórias i.i.d (independentes e identicamente distribuídas) com distribuição de probabilidade discreta. Além disto, Cov(ε_t, z_s)= 0 para s<t.

Apresentamos abaixo o teorema que enuncia sobre quais condições os parâmetros $\alpha_1, \alpha_2, \ldots, \alpha_p$ garantem a existência de um processo INAR(p) estacionário. A prova deste teorema será omitida, podendo ser encontrada em Jin -Guan & Yuan (1991).

TEOREMA 2.9.1

No modelo em (2.9.1), considerando uma sequência de inovações ε_t, i.i.d. assumindo valores no conjunto dos inteiros não negativos com $E(\varepsilon_t) = \mu$, $Var(\varepsilon_t) = \sigma^2$ e $\alpha_j \in [0,1]$; $j = 1,2,\ldots,p$. Se as raízes da equação :

$$a^k - a^{k-1}\alpha_1 - a^{k-2}\alpha_2 - \ldots - a\alpha_{p-1} - \alpha_p = 0$$

estão dentro do círculo unitário, existe uma única série estacionária Z_t satisfazendo :

$$Z_r = \alpha_1 \circ Z_{t-1} + \alpha_2 \circ Z_{t-2} + \ldots + \alpha_p \circ Z_{t-p} + \varepsilon_t$$

$Cov(\varepsilon_t, Z_s) = 0$ *para* $s<t$.

Denotando γ_j e ρ_j, $j = 0,1,2,\ldots$, como sendo respectivamente as funções de autocovariância e autocorrelação do processo Z_t, os estimadores obtidos através do método dos momentos, com base na amostra $\mathbf{Z} = (z_1, z_2, \ldots, z_n)$ são :

Função de autocovariância amostral

$$\hat{\gamma}_j = \frac{1}{n}\sum_{t=1}^{n-j}(z_t - \bar{z})(z_{t+j} - \bar{z})$$

Função de autocorrelação amostral

$$\hat{\rho}_j = \frac{\hat{\gamma}_j}{\hat{\gamma}_0}$$

Estes estimadores são utilizados nas equações de Yule-Walker para a obtenção de uma estimativa para o vetor de parâmetros $\Theta = (\mu, \alpha_1, \alpha_2, \ldots, \alpha_p, \sigma^2)$

Estimadores de Yule-Walker

Os estimadores de Yule-Walker são obtidos através da solução do seguinte sistema de equações :

$$\alpha_1 + \alpha_2\rho_1 + \ldots + \alpha_p\rho_{p-1} = \rho_1$$

$$\alpha_1\rho_1 + \alpha_2 + \ldots + \alpha_p\rho_{p-2} = \rho_2$$

.

$$\alpha_1\rho_{p-1} + \alpha_2\rho_{p-2} + \ldots + \alpha_p = \rho_p$$

Substituindo ρ_j pelos estimadores $\hat{\rho}_j$ são obtidas estimativas para α_j que satisfaçam o sistema de equações acima. A estimativa para μ é obtida através de :

$$\hat{\mu} = (1 - \hat{\alpha}_1 - \hat{\alpha}_2 - \ldots - \hat{\alpha}_p)\bar{z}$$

A estimativa para $\hat{\sigma}^2$ é obtida em :

$$\hat{\sigma}^2 = \frac{1}{n-p}\sum_{t=p+1}^{n}(\varepsilon_t - \bar{\varepsilon}_n)^2$$

$$\bar{\varepsilon}_n = \frac{1}{n-p}\sum_{t=p+1}^{n}\hat{\varepsilon}_t$$

Pode ser demonstrado que os estimadores $(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_p, \hat{\mu}, \hat{\sigma}^2)$ são consistentes. A previsão 1-passo-à-frente para este modelo pode ser encontrada através de:

$$E(Z_{t+1}|\mathbf{Z}_t) = \alpha_1 Z_t + \alpha_2 Z_{t-1} + \ldots + \alpha_p Z_{t+1-p} + \mu$$

Analogamente podemos obter a previsão m-passos-à-frente

$$E(Z_{t+m}|\mathbf{Z}_t) = \sum_{j=1}^{p} E(\alpha_j \circ Z_{t+m-j}|\mathbf{Z}_t) + \mu$$

CAPÍTULO 3

MODELOS ESTRUTURAIS PARA DADOS DE CONTAGEM

3.1 - INTRODUÇÃO

Os modelos estuturais para dados de contagem constituem uma adaptação dos modelos estruturais propostos por Harvey (1984), nos quais uma série temporal é representada em termos das suas componentes tais como a tendência, sazonalidade e ciclo. Nesta referência, toda a teoria foi desenvolvida, supondo as observações provenientes da distribuição normal, fato que ajudava a tornar explicita a forma de evolução das componentes ao longo do tempo. Ao contrário dos pressupostos dos modelos Gaussianos, a presente situação nos traz observações de uma distribuição discreta, aumentando a dificuldade em explicitar o comportamento dinâmico dos parâmetros. Utilizando nesta abordagem, distribuições tais como a Poisson, binomial negativa, binomial e multinomial, entre outras, os modelos estruturais para dados de contagem são construídos de forma que fique implícita, através da média e variância, a evolução dos parâmetros.

O mecanismo de evolução dos parâmetros destas distribuições, uma vez identificado, permite o acesso a distribuição preditiva e a estimação por máxima verossimilhança, através de procedimentos recursivos semelhantes aos que são utilizados no Filtro de Kalman (Kalman,1960). A solução para a identificação da evolução estocástica dos parâmetros é similar àquela da inferência Bayesiana onde distribuições conjugadas naturais são atribuídas aos parâmetros que evoluem ao longo do tempo.

Dentre as distribuições discretas utilizadas para a formulação de um modelo estrutural, duas receberão enfoque neste trabalho ; as distribuições de Poisson e a binomial negativa. Através destas distribuições, são encontrados os modelos Poisson-Gamma e Binomial Negativa-Beta, assim intitulados por Harvey e Fernandes (1989), que se aplicam a situações envolvendo dados de contagem.

A formulação dos modelos estruturais pode ser ilustrada a partir de um dos mecanismos mais simples para descrever a evolução de uma série temporal que é o passeio aleatório :

$$z_t = \mu + \varepsilon_t \tag{3.1.1}$$

Na equação (3.1.1) , o comportamento da serie Z_t é descrito por uma média global perturbada por um ruído aleatório cuja média pode ser diferente de zero, segundo alguma distribuição de probabilidade. Permitindo que a média evolua ao longo do tempo, segundo um passeio ao acaso, temos :

Equação de observação: $$z_t = \mu_t + \varepsilon_t \tag{3.1.2}$$

Equação de transição : $$\mu_t = \mu_{t-1} + \eta_t \tag{3.1.3}$$

Nesta situação, z_t é um processo estocástico cujo nível μ_t evolui estocasticamente segundo a equação de transição (3.1.3). O nível do processo no instante t, μ_t , é , em geral, uma variável não observável, enquanto η_t e ε_t são variáveis não correlacionadas, com variâncias constantes σ_η^2 e σ_ε^2 , respectivamente, para qualquer escolha de t. Quanto mais elevadas for a variância σ_η^2 destas perturbações, maiores serão os movimentos estocásticos do nível da série.

Definimos m_{t-1} e p_{t-1} como sendo respectivamente a média e variância do nível da série condicionado ao vetor de observações $\mathbf{Z}_{t-1} = (z_1, z_2, \ldots, z_{t-1})$. Denotamos a distribuição de probabilidade condicionada por DP, isto é :

$$(\mu_{t-1}|\mathbf{Z}_{t-1}) \sim DP(m_{t-1}, p_{t-1}) \tag{3.1.4}$$

$$m_{t-1} = E(\mu_{t-1}|\mathbf{Z}_{t-1}) \tag{3.1.5}$$

$$p_{t-1} = Var(\mu_{t-1}|\mathbf{Z}_{t-1}) \tag{3.1.6}$$

A evolução do nível na transição do instante t-1 para o instante t, em termos da média e variância, é representada através das equações (3.1.7) e (3.1.8) :

$$E(\mu_t|\mathbf{Z}_{t-1}) = E(\mu_{t-1} + \eta_t|\mathbf{Z}_{t-1})$$
$$= E(\mu_t|\mathbf{Z}_{t-1}) + E(\eta_t|\mathbf{Z}_{t-1}) \qquad (3.1.7)$$
$$= m_{t-1}$$

$$Var(\mu_t|\mathbf{Z}_{t-1}) = Var(\mu_{t-1} + \eta_t|\mathbf{Z}_{t-1})$$
$$= Var(\mu_{t-1}|\mathbf{Z}_{t-1}) + Var(\eta_t|\mathbf{Z}_{t-1}) \qquad (3.1.8)$$
$$= p_{t-1} + \sigma_\eta^2 > Var(\mu_{t-1}|\mathbf{Z}_{t-1}) \quad ; \; p_{t-1}, \sigma_\eta^2 > 0$$

Ou seja, durante a transição, o nível esperado do processo permanece constante, enquanto a sua variância é acrescida da variância do ruído.

Considerando nas equações (3.1.2) e (3.1.3) os ruídos η_t e ε_t com médias iguais a zero e ambos com distribuição Gaussiana , o processo fica definido da seguinte forma :

$$(Z_t|\mu_t) \sim Normal(\mu_t, \sigma_\varepsilon^2) \qquad (3.1.9)$$

$$(\mu_t|\mu_{t-1}) \sim Normal(\mu_{t-1}, \sigma_\eta^2) \qquad (3.1.10)$$

3.2 - O MODELO POISSON-GAMA

No modelo Poisson-Gama, as observações, condicionadas a λ_t, seguem a distribuição Poisson com média λ_t cuja função massa de probabilidade é dada por :

$$P_{Z_t|\lambda_t}(z) = \frac{e^{-\lambda_t}\lambda_t^z}{z!} \qquad z = 0,1,2,... \tag{3.2.1}$$

Adicionalmente, é adotada a suposição de que o nível do processo, condicionado ao vetor de observações $\mathbf{Z}_{t-1} = (z_1, z_2, \ldots, z_{t-1})$, apresenta a distribuição gama com parâmetros a_{t-1} e b_{t-1} , sendo a sua função densidade de probabilidade representada abaixo:

$$f_{\lambda_{t-1}|\mathbf{Z}_{t-1}}(\lambda) = \frac{e^{-b_{t-1}\lambda}\lambda^{a_{t-1}}}{\Gamma(a_{t-1})b_{t-1}^{a_{t-1}}}, \quad a_{t-1}, b_{t-1} > 0 \tag{3.2.2}$$

Nesta abordagem, a_{t-1} e b_{t-1} são calculados em função do vetor de observações $\mathbf{Z}_{t-1}$ e assim são estabelecidas as equações abaixo :

$$E(\lambda_{t-1}|\mathbf{Z}_{t-1}) = E(\lambda_{t-1}|\, a_{t-1}, b_{t-1}) = \frac{a_{t-1}}{b_{t-1}} \tag{3.2.3}$$

$$Var(\lambda_{t-1}|\mathbf{Z}_{t-1}) = Var(\lambda_{t-1}|\, a_{t-1}, b_{t-1}) = \frac{a_{t-1}}{b_{t-1}^2} \tag{3.2.4}$$

No modelo Poisson-Gama, a equação de transição não é colocada de forma explícita tal como é feito em (3.1.2), embora esta possa ser explicitada de forma multiplicativa (ver seção 3.4), mas é construída de forma a induzir os mesmos resultados obtidos em (3.1.7) e (3.1.8), ou seja, a média e a variância do nível do processo sejam tais que :

$$E(\lambda_t|\mathbf{Z}_{t-1}) = E(\lambda_{t-1}|\mathbf{Z}_{t-1}) \tag{3.2.5}$$

$$Var(\lambda_t|\mathbf{Z}_{t-1}) > Var(\lambda_{t-1}|\mathbf{Z}_{t-1}) \tag{3.2.6}$$

Supondo adicionalmente que a distribuição na transição permaneça sendo a Gama com parâmetros $a_{t|t-1}$ e $b_{t|t-1}$ como resultados temos as seguintes relações :

$$a_{t|t-1} = \omega a_{t-1} \tag{3.2.7}$$

$$b_{t|t-1} = \omega b_{t-1} \tag{3.2.8}$$

Os valores de $\omega > 0$ garantem que os parâmetros sejam positivos, enquanto que as relações (3.2.7) e (3.2.8) são necessárias e suficientes para que o valor esperado do nível do processo permaneça constante na transição, conforme mostra a equação abaixo :

$$E(\lambda_t \mid Z_{t-1}) = \frac{\omega a_{t-1}}{\omega b_{t-1}} = \frac{a_{t-1}}{b_{t-1}} = E(\lambda_{t-1} \mid Z_{t-1}) \tag{3.2.9}$$

Se considerarmos também valores de $\omega < 1$, alcançamos o mesmo resultado alcançado em (3.2.6), ou seja, o aumento da variância na transição, mostrado a seguir :

$$Var(\lambda_t | \mathbf{Z}_{t-1}) = \frac{a_{t-1}}{b^2_{t|t-1}} = \frac{Var(\lambda_{t-1} | \mathbf{Z}_{t-1})}{\omega} \tag{3.2.10}$$

Por analogia à abordagem Bayesiana, chamamos a partir de agora, ω de hiperparâmetro do modelo Poisson-Gama. É direta a verificação em (3.2.10) de que quanto mais próximo de zero estiver o valor do hiperparâmetro, maior será o aumento da variância do nível do processo, na transição.

Quando a observação no instante t passa a estar disponível, os parâmetros da distribuição de $\lambda_t | \mathbf{Z}_{t-1}$ são atualizados de forma a incorporar no sistema o ganho de informação ocorrido. Através do Teorema de Bayes e novamente recorrendo a uma analogia ao enfoque apresentado pela inferência Bayesiana, a atualização dos parâmetros é feita através da seguinte relação de proporcionalidade :

$$\text{posteriori} \propto \text{priori x verossimilhança} \tag{3.2.11}$$

No presente enfoque, as distribuições de $\lambda_t|\mathbf{Z}_t$ e $\lambda_t|\mathbf{Z}_{t-1}$ representam respectivamente as distribuições a posteriori e a priori na abordagem Bayesiana .

$$p(\lambda_t|\mathbf{Z}_t) \propto p(\lambda_t|\mathbf{Z}_{t-1})p(Z_t|\lambda_t,\mathbf{Z}_{t-1}) \tag{3.2.12}$$

Procedendo as devidas substituições em (3.2.12) e utilizando a verossimilhança com base na observação feita no instante t , teremos

$$f_{\lambda_t|\mathbf{Z}_{t-1}} \propto \frac{e^{-\lambda_t}\lambda_t^{z_t}}{z_t!}\frac{e^{b\lambda_t}\lambda_t^{a}b^{a}}{\Gamma(a)} \tag{3.2.13}$$

$$= \frac{1}{c}\frac{e^{-\lambda_t(b+1)}\lambda_t^{a+z_t-1}}{\Gamma(a+z)(1+b)^{-(a+z_t)-1}}$$

$$a = a_{t|t-1}$$
$$b = b_{t|t-1}$$

onde 1/c é a constante de proporcionalidade dada por :

$$\frac{1}{c} = \frac{\Gamma(a+z_t)(1+b)^{-(a+z_t)-1}}{z_t!\Gamma(a)b^{-a}}$$

Fora a constante de proporcionalidade, temos que a distribuição obtida para $(\lambda_t|Z_t)$ é gama com os parâmetros :

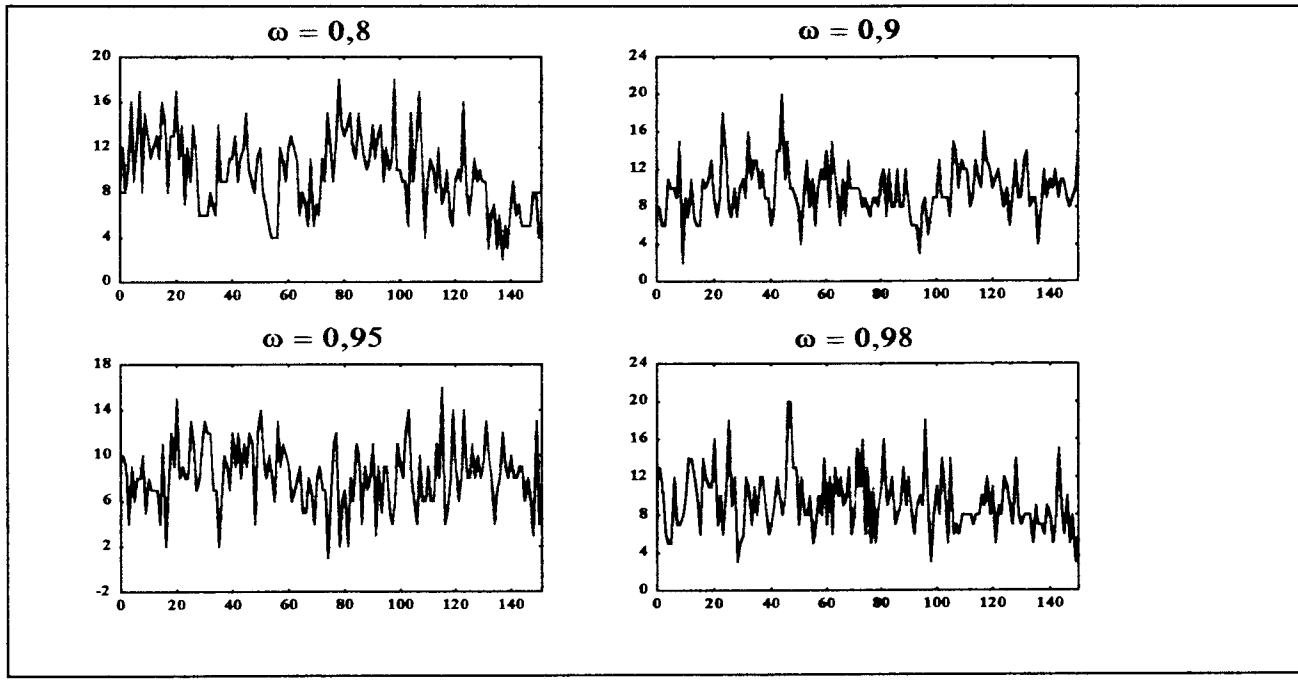
$$a_t = a_{t|t-1} + z_t$$
$$b_t = b_{t|t-1} + 1 \tag{3.2.14}$$

A partir da atualização ocorrida nos parâmetros da distribuição gama após a observação no instante t, a distribuição na transição de t para $t+1$, $(\lambda_{t+1} | Z_t)$, passa a

ser gama com parâmetros ωa_t e ωb_t, finalizando desta forma o filtro recursivo para a estimação dos parâmetros que caracterizam o estado. A inicialização deste algoritmo requer valores iniciais para a_o e b_o que podem ser atribuídos de forma a caracterizar uma alta variância para a transição a partir do estado inicial, como por exemplo se tomarmos $a_0 = ½$ e $b_0 = 0$. Com estes valores, a distribuição inicial é conhecida como densidade a priori não informativa de Jeffrey que é amplamente utilizada na inferência Bayesiana quando se deseja mostrar ignorância sobre o estado inicial do processo.

Partindo dos valores iniciais $a_o = 10$ e $b_o = 1$, consequentemente supondo que o nível esperado da série no instante $t = 0$ é 10, apresentamos abaixo algumas simulações do processo Poisson-Gamma, com 150 observações.

Figura 3.2.1 - Simulações de 150 valores do processo Poisson-Gama com valores iniciais $a_0 = 10$ e $b_0 = 1$



3.2.1 - Propriedades do modelo Poisson-Gama

Realizando sucessivas substituições em (3.2.7) e (3.2.8), os parâmetros de $(\lambda_t \mid Z_{t-1})$ passam a ser escritos em função das observações :

$$a_{t/t-1} = \sum_{j=1}^{t-1} \omega^j z_{t-j}$$

$$b_{t/t-1} = \sum_{j=1}^{t-1} \omega^j$$

Ao tomarmos a esperança de $(\lambda_t \mid Z_{t-1})$ como valor predito do nível da série no instante t , podemos escrevê-lo em função das observações.

$$E(\lambda_t \mid Z_{t-1}) = \frac{a_{t|t-1}}{b_{t|t-1}} = \frac{\sum_{j=1}^{t-1} \omega^j Z_{t-j}}{\sum_{j=1}^{t-1} \omega^j} = \sum_{j=1}^{t-1} \beta_j Z_{t-j} \qquad (3.2.14)$$

$$\beta_j = \frac{\omega^j}{\sum_{j=1}^{t-1} \omega^j}$$

Para séries grandes, ou seja, a medida que $t \to \infty$, o valor de $\sum_{j=1}^{t-1} \omega^j \to \omega / (1-\omega)$ e podemos encarar o modelo Poisson-Gama como um tipo de alisamento exponencial dos valores da série com fator de desconto $\beta_j = \omega^j (1-\omega)$. Neste enfoque, a medida que o valor de ω se aproxima de 1, menor é o desconto dado as observações passadas e o nível predito 1-passo-à-frente está próximo da média aritmética das observações.

3.2.2 - Distribuição Preditiva 1-passo-à-frente

Embora tenhamos caracterizado a evolução do parâmetro, o interesse principal pode estar na previsão dos valores de Z. A distribuição preditiva 1-passo-à-frente é resultante da aplicação de :

$$P_{Z_t|Z_{t-1}}(z)=\int_{\Lambda} P_{Z_t|\lambda_t Z_{t-1}}(z) f_{\lambda_t|Z_{t-1}}(\lambda)d\lambda_t \qquad (3.2.15)$$

As distribuições Poisson e gama em (3.2.15) resultam na distribuição binomial negativa como a preditiva 1-passo-à-frente.

$$P_{Z_t|Z_{t-1}}(z)=\binom{a+z-1}{z}\left(\frac{b}{1+b}\right)^a\left(\frac{1}{1+b}\right)^z \qquad (3.2.16)$$

Conforme as características da distribuição binomial negativa, a previsão 1-passo-à-frente mede a probabilidade de obtenção de z fracassos até a obtenção do a-ésimo sucesso em uma seqüência de experimentos de bernoulli, independentes, com probabilidade de sucesso igual a $b/(1+b)$.

Em decorrência das propriedades da distribuição binomial negativa, a média e variância da distribuição preditiva 1-passo-à-frente são respectivamente :

$$E(z_t|Z_{t-1}) = \frac{a_{t|t-1}}{b_{t|t-1}} = \frac{a_{t-1}}{b_{t-1}}$$

$$Var(z_t|Z_{t-1}) = \frac{a_{t|t-1}}{b^2_{t|t-1}}(1+b_{t|t-1})$$

Como verificamos acima, o valor esperado da série para o instante t, condicionado às observações até o instante t-1, é dado pelo valor esperado do nível da série, enquanto a variância, ao contrário da distribuição Poisson, é superior à média. Conforme (3.2.14), em séries de grande tamanho, o valor predito da série um passo a frente pode ser escrito como uma média ponderada das observações até t-1.

$$E(Z_t|\mathbf{Z}_{t-1}) = \sum_{j=1}^{t-1} \beta_j Z_{t-j}\,, \qquad 0 < \beta_j < 1$$

A previsão um passo a frente assume a forma de uma suavização dos valores da série onde os pesos β_j decaem exponencialmente com a defasagem j de tempo. Na situação onde $\omega = 1$, o modelo assume que o nível é estático e $E(Z_t|\mathbf{Z}_{t-1})$ é fornecida pela média aritmética dos valores observados até o instante t-1.

Utilizar um horizonte maior para previsão implica em uma maior dificuldade de estabelecimento da distribuição preditiva. Entretanto, é possível encontrar analiticamente as expressões para o primeiro e segundo momento. Em Harvey e Fernandes (1989), podemos encontrar este desenvolvimento e o resultado para a previsão pontual k-passos-à-frente é o mesmo encontrado no método de alisamento exponencial simples.

$$E(Z_{t+k}\,|Z_t) = a_t\,/b_t = \sum_{j=1}^{t-1} \beta_j Z_{t-j}\ , \quad 0 < \beta_j < 1$$

3.2.3 - Estimação por máxima verossimilhança

Utilizando a distribuição preditiva em (3.2.16), a função de máxima verossimilhança é calculada com base no vetor $Z = (z_\tau, \ldots, z_n)$ no qual τ é o índice da primeira observação diferente de zero. Desta forma o procedimento de maximização pode ser realizado, atribuindo, sem maiores problemas, os valores iniciais $a_0 = 0$ e $b_0 = 0$, já que a distribuição inicial para λ_t é obtida a partir do instante t = τ, quando é observado na série o primeiro valor diferente de zero.

$$L(\omega) = \prod_{t=\tau}^{n} P_{Z_t|Z_{t-1}}(z)$$

A função de log-verossimilhança para o parâmetro ω é obtida em :

$$\log L(\omega) = \sum_{\tau+1}^{n} \{\log\Gamma(a_{t|t-1} + z_t) - \log z_t! - \log\Gamma(a_{t|t-1}) + a_{t|t-1}\log b_{t|t-1} -$$

$$-(a_{t|t-1} + z_t)\log(1 + b_{t|t-1})\}.$$

A estimativa de máxima verossimilhança do parâmetro ω é encontrada maximizando a função logL(ω), o que é feito através da rotina de maximização *fmin* presente no programa computacional MATLAB. Esta rotina é adequada para a maximização de uma função de um parâmetro sob restrições, sem que haja necessidade de transformações. Nesta situação, a restrição é de que o parâmetro assume valores entre 0 e 1. As propriedades deste estimador foram estudadas através de simulações de Monte Carlo , em Fernandes (1990), que mostraram em grandes amostras, que este é não viciado, consistente e apresenta distribuição normal assintótica. Entretanto, em amostras de tamanhos pequeno e médio, o estimador se comportou de forma pouco usual, especialmente para valores de ω próximos a 1.

3.2.4 - Correção de Shephard para o modelo Poisson-Gama

Conforme demonstrado em Smith & Miller (1990), a equação de transição do modelo Poisson-Gamma pode ser escrita de forma explícita e multiplicativa, com uso da distribuição Beta, tomando a seguinte forma :

$$\lambda_t = \omega^{-1}\lambda_{t-1}\eta_t$$
$$\eta_t \sim Beta(\alpha,\beta);\ \alpha = \omega a_{t-1}\ ;\ \beta = (1-\omega)a_{t-1}$$

A função densidade de probabilidade para a distribuição Beta é escrita da forma :

$$f_{\eta_t|\mathbf{z}_{t-1}}(\eta) = \frac{\eta^{\alpha-1}(1-\eta)^{\beta-1}}{B(\alpha,\beta)}$$

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

No modelo Poisson-Gamma, ω <1 faz com que o parâmetro λ_t convirja quase certamente para 0 quando $t \to \infty$. A explicação para este resultado pode ser encontrada, se considerarmos que $\log \lambda_t/\lambda_{t-1}$ representa uma taxa de crescimento e através da desigualdade de Jensen podemos verificar que esta é , em média, negativa. Shephard (1990) propõe uma pequena modificação na equação de transição multiplicativa para contornar este problema de forma que a esperança desta taxa de crescimento seja igual a zero. A proposta de Shephard é proceder a substituição de ω por e^{r_t} fazendo com que o valor esperado de $\log \lambda_t/\lambda_{t-1}$ seja igual a zero. Desta forma, a equação multiplicativa de transição passa a ser escria da seguinte forma :

$$\lambda_t = e^{r_t}\lambda_{t-1}\eta_t \quad \text{onde} \quad r_t = -E[\log \eta_t]$$

Apesar de não existir uma forma fechada para r_t , há em Shephard (1994) um simples algoritmo para o cálculo desta quantidade.

A correção de Shephard é extremamente útil para a situação na qual pretende-se simular o processo Poisson-Gama com valores pequenos para o hiperparâmetro ω , evitando que λ_t convirja para zero e conseqüentemente o processo seja absorvido pelo estado zero.

3.3 - O MODELO BINOMIAL NEGATIVA-BETA

A distribuição binomial negativa (Johnson e Kotz, 1969), é utilizada quando desejamos obter a probabilidade da realização de k experimentos de Bernoulli independentes com probabilidade p de sucesso até a obtenção do r-ésimo sucesso. A função massa de probabilidade para uma variável aleatória Z com distribuição binomial negativa é dada por :

$$P(Z = k) = \binom{r+k-1}{k} p^r (1-p)^k \qquad 0 < p < 1$$

$r > 0$, $k = 0,1,2,3...$

Quando r =1, temos a distribuição geométrica como um caso particular da distribuição acima. Embora possa ser de interesse prático trabalhar-se com valores inteiros de r, pela definição da distribuição, nada impede que este valor pertença ao conjunto dos reais.

No modelo Binomial Negativa-Beta, as observações seguem a distribuição binomial negativa com parâmetros π e ν, sendo a função massa de probabilidade escrita da seguinte forma :

$$P_{Z_t|\pi_t,\nu}(z) = \binom{\nu+z-1}{z} \pi_t^{\nu} (1-\pi_t)^z \qquad (3.3.1)$$

$\nu > 0$

$0 < \pi_t < 1$

Para $\nu = 1$,. Temos o modelo Geométrica-Beta como caso particular deste modelo. Em (3.3.1), a suposição é de que o parâmetro π_t segue a distribuição beta que é a distribuição conjugada natural da distribuição binomial negativa no enfoque Bayesiano. A função densidade de probabilidade do parâmetro, condicionada ao vetor de observações até o instante t-1 é dada por :

$$f_{\pi_{t-1}|Z_{t-1}}(\pi) = \frac{\pi^{a_{t-1}-1}(1-\pi)^{b_{t-1}-1}}{B(a_{t-1}, b_{t-1})} \tag{3.3.2}$$

$a_{t-1} > 0$

$b_{t-1} > 0$

$$B(a_{t-1}, b_{t-1}) = \frac{\Gamma(a_{t-1})\Gamma(b_{t-1})}{\Gamma(a_{t-1} + b_{t-1})}$$

A transição do estado neste modelo é caracterizada da mesma forma que no modelo Poisson-Gama, só que desta vez, o procedimento é induzir que a esperança de $(1-\pi_t)/\pi_t$, ao invés de π_t , permaneça constante e sua variância aumente na transição, de forma que fiquem caracterizadas as seguintes relações :

$$E(\frac{1-\pi_t}{\pi_t} \,|Z_{t-1}) = E(\frac{1-\pi_{t-1}}{\pi_{t-1}} \,|Z_{t-1}) \tag{3.3.3.a}$$

$$Var(\frac{1-\pi_t}{\pi_t} \,|Z_{t-1}) > Var(\frac{1-\pi_{t-1}}{\pi_{t-1}} \,|Z_{t-1}) \tag{3.3.3.b}$$

Supondo que a distribuição de π permaneça sendo beta na transição com parâmetros $a_{t|t-1}$ e $b_{t|t-1}$, garantimos que sejam satisfeitas as relações em (3.3.3.a ,b) através de :

$$a_{t|t-1} = \omega a_{t-1} + (1-\omega) \;, \tag{3.3.4.a}$$

$$b_{t|t-1} = \omega b_{t-1}, \tag{3.3.4.b}$$

$0 < \omega < 1.$

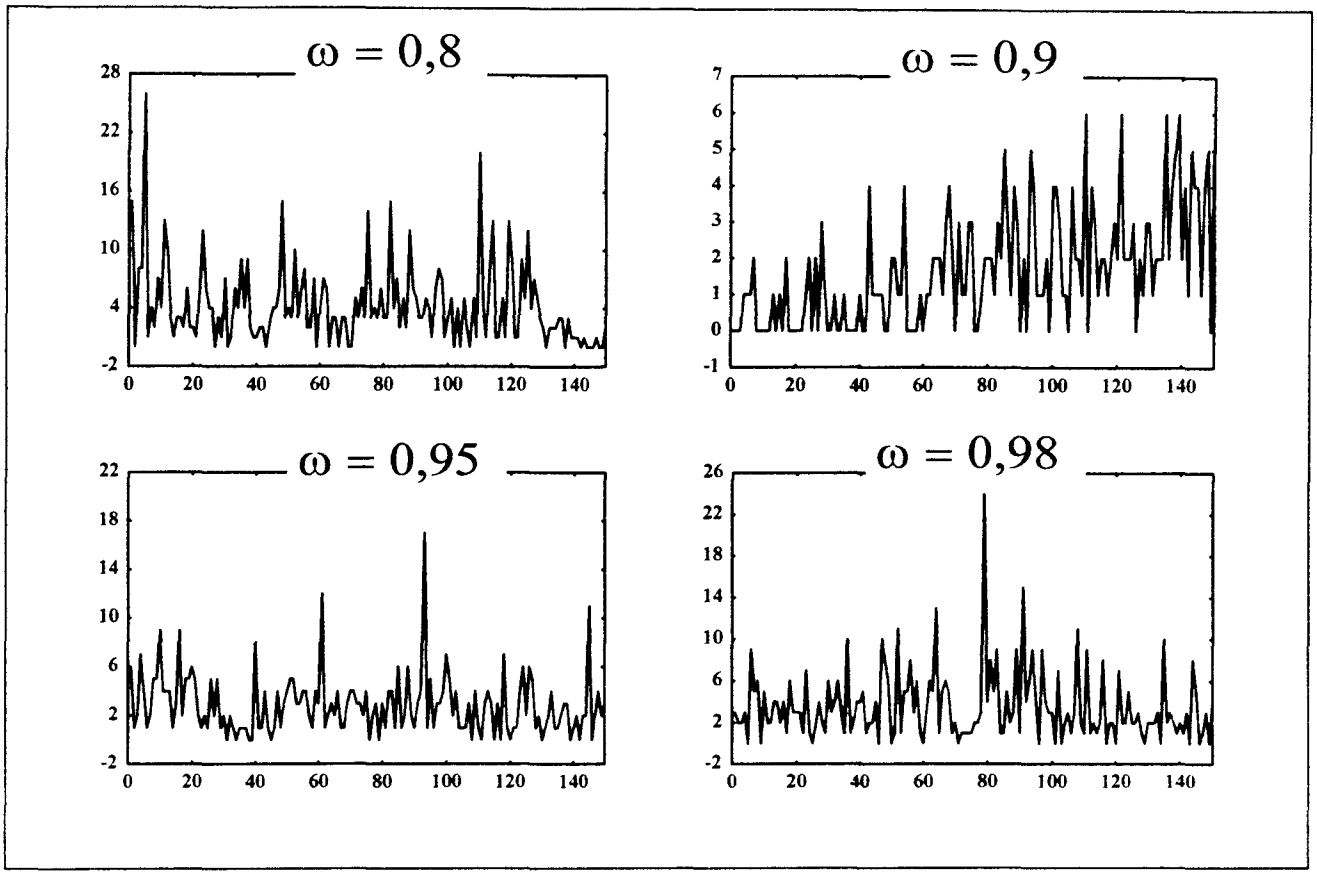
O hiperparâmetro ω é caracterizado da mesma forma que no modelo Poisson-Gama. Após ter observado z_t , a utilização direta do Teorema de Bayes leva as equações que caracterizam o estado a serem atualizadas conforme as equações abaixo :

$$a_t = a_{t|t-1} + v \tag{3.3.5.a}$$

$$b_t = b_{t|t-1} + z_t \tag{3.3.5.b}$$

Na Figura 3.3.1 apresentamos algumas simulações do processo Binomial Negativa-Beta , fixando o parâmetro $\nu = 3$ e utilizando os valores iniais para os parâmetros da distribuição Beta $a_0 = 9$ e $b_0 = 10$.

Figura 3.3.1 - Simulações do modelo Binomial Negativa-Beta para o valor fixo $\nu = 3$ e distribuição do parâmetro de estado no instante $t = 0$, BETA(9,10).



A distribuição binomial negativa apresenta sobredispersão, ao contrário da distribuição poisson. Para ter uma idéia da diferença do processo Binomial Negativa-Beta em relação ao Poisson-Gama, podemos comparar a figura acima com a Figura 3.2.1 e perceberemos que na situação atual há trechos na série simulada que apresentam uma variação maior, com a presença de picos que se destacam, ao contrário do processo Poisson-Gama no qual as mudanças são percebidas no nível da série e não na variância.

3.3.1 - Propriedades do modelo Binomial Negativa-Beta

O momento de n-ésima ordem do parâmetro π é obtido conforme :

$$E(\pi_t^n|\mathbf{Z}_{t-1}) = \frac{\Gamma(a_{t|t-1} + b_{t|t-1})\Gamma(a_{t|t-1} + n)}{\Gamma(a_{t|t-1})\Gamma(a_{t|t-1} + b_{t|t-1} + n)} \tag{3.3.6}$$

Através de sucessivas substituições em (3.3.4 a,b) , escrevemos os parâmetros em função dos valores passados de modo a obter:

$$a_{t|t-1} = v\sum_{j=1}^{t-1}\omega^j + (1-\omega)\sum_{j=1}^{t-1}\omega^j + (1-\omega) \tag{3.3.7a}$$

$$b_{t|t-1} = \sum_{j=1}^{t-1}\omega^j z_{t-j} \tag{3.3.7b}$$

A situação na qual $t \to \infty$, $a_{t|t-1} \to \dfrac{\omega(1-v)+1}{1-\omega}$

3.3.2 - Distribuição preditiva 1-passo-à-frente

Utilizando o mesmo procedimento feito em (3.2.15), desta feita com as distribuições binomial negativa e beta, encontramos como distribuição preditiva um-passo-a-frente a distribuição beta-Pascal, apresentada abaixo :

$$P_{Z_t|Z_{t-1}}(z) = \frac{1}{v+z}\frac{B(v+a_{t\ t-1}, z+b_{t\ t-1})}{B(v,z+1)B(a_{t\ t-1}, b_{t\ t-1})} \tag{3.3.8}$$

Em decorrência das propriedades da distribuição beta-Pascal, a média e variância da distribuicao preditiva sâo respectivamente :

$$E(Z_t|Z_{t-1}) = \frac{vb_{t-1}}{a_{t-1}-1} \tag{3.3.9 a}$$

$$Var(Z_t|\, Z_{t-1}) = E(Z_t|Z_{t-1})\frac{(a_{t\,|t-1} + b_{t\,|t-1} + 1)[\nu + \omega(a_{t\,|t-1})]}{(a_{t\,|t-1} - 1)[\omega(a_{t\,|t-1} - 1) - 1]} \qquad (3.3.9\ b)$$

3.3.3 - Estimação por máxima verossimilhança

Utilizando (3.3.8), construímos a função de verossimilhança e a estimativa para o vetor de parâmetros (ω,ν) é encontrada através da maximização da função de log-verossimilhança apresentada abaixo :

$$\log L(\omega,\nu) = \sum_{t=t+1}^{n} \{\log[\Gamma(\nu + a_{t|t-1})/\Gamma(a_{t|t-1})] + \log[\Gamma(\nu + z_t + 1)/\Gamma(\nu)] +$$
$$+ \log[\Gamma(z_t + b_{t|t-1})/\Gamma(b_{t|t-1})] - \log[\Gamma(z_t + \nu + d_{t|t-1})/d_{t|t-1}]\} \qquad (3.3.10)$$

$$d_{t|t-1} = a_{t|t-1} + b_{t|t-1}$$

A solução para (3.3.10) pode ser encontrada maximizando (3.3.10) em relação ao vetor de parâmetros (π, ν) ou , alternativamente, fixando o valor de ν e maximizando esta função em relação a π somente.

CAPÍTULO 4

COMPARAÇÃO ENTRE OS MODELOS

4.1 - INTRODUÇÃO

Neste capítulo, apresentamos a metodologia a ser utilizada para a comparação de alguns dos modelos apresentados nos capítulos anteriores. Serão analisados 3 modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama, possibilitando assim duas formas básicas de comparação : a primeira entre dois modelos INAR(1) com inovações de diferentes distribuições e a segunda entre o modelo INAR(1) e o modelo estrutural.

Na comparação, serão avaliados os seguintes aspectos : ajuste do modelo e previsão 1-passo-à-frente , utilizando séries simuladas e séries reais. Ao utilizarmos séries reais (ver seção 4.2) a comparação será basicamente qualitativa pois estas são relativamente curtas. Ao trabalharmos com séries simuladas, os resultados poderão ser mais detalhados pois teremos controle sobre o processo gerador dos dados (PGD). Ao simularmos , por exemplo, o processo INAR(1)-Poisson (seção 4.3) e ajustarmos os 3 modelos em questão, podemos verificar se uma classe de modelos diferente daquela utilizada para a geração se ajusta aos dados. Desta forma, estaremos acessando a robustez destes modelos em relação ao PGD. Em seguida, nas seções 4.4 e 4.5, simulamos processos gerados pelos modelos INAR(1)-Geométrica e Poisson-Gama e repetimos a comparação. Os pontos escolhidos para a simulação estão descritos na seção 4.1.1.

Para estudar a qualidade dos ajustes dos modelos em séries simuladas, ou seja, o quanto uma família de modelos pode aproximar modelos gerados por outra estrutura, na seção 4.6, vamos utilizar os critérios AIC (Akaike,1973) e BIC (Schwarz ,1978) para realizar a seleção do melhor ajuste nas séries simuladas. Em uma série gerada, por exemplo, pelo modelo INAR(1)-Poisson, vamos escolher, através destes critérios, o melhor ajuste entre ele e uma das alternativas; o modelo INAR(1)-Geométrica e o Poisson-Gama. Quanto menor for a frequência de escolhas corretas, ou seja, o número de vezes que o modelo INAR(1)-Poisson for identificado como o correto, melhor será a aproximação deste PGD por outros modelos.

A comparação em relação a previsão 1-passo-à-frente será realizada através da distribuição preditiva. Inicialmente comparamos alguns de seus parâmetros tais como a média, variância e erro médio quadrático e posteriormente utilizamos a medida de informação não paramétrica de Kullback-Leibler, apresentada na seção 4.1.2, para medir o grau de afinidade entre a distribuição preditiva 1-passo-à-frente do modelo ajustado e a do modelo utilizado para a geração dos dados.

4.1.1- Escolha dos processos geradores para simulação das séries

Foram simuladas séries de tamanho $n=50,100$ *e* 300 de forma a representar pequenas, médias e grandes amostras. Os parâmetros do modelo INAR(1)-Poisson escolhidos para a simulação foram $\alpha=\{0,1;0,3;0,5;0,7;0,9\}$ e $\lambda=\{1;2;3\}$. Para o modelo INAR(1)-Geométrica, foram escolhidos $\phi=\{0,1;0,3;0,5;0,7;0,9\}$ e $\theta=\{0,1;0,3;0,5;0,7;0,9\}$ e para o modelo Poisson-Gama foram escolhidos os valores $\omega=\{0,8;0,9;0,95\}$. Para cada caso foram utilizadas 20 replicações.

O critério principal na escolha destes valores foi a geração de séries temporais com valores de pequena magnitude de forma que a utilização de modelos para dados contínuos fosse totalmente descartada. Desta forma, da escolha inicial feita para os parâmetros, foi retirada a combinação $\phi = 0,9$ e $\theta = 0,1$ pois os processos INAR(1)-Geométrica obtidos nestas simulações apresentaram valores de alta magnitude, dificultando os procedimentos de estimação de parâmetros. Como valores inciais na simulação dos processos INAR(1) foram utilizados $Z_0 = \dfrac{\lambda}{1-\mu}$ na utilização da distribuição Poisson como inovação e $Z_0 = \dfrac{1-\theta}{\theta(1-\phi)}$ na situação envolvendo a distribuição geométrica. Na simulação do processo Poisson-Gama, não foi utilizada a correção de Shephard, de modo que os parâmetros foram escolhidos de forma que a série não convergisse rapidamente para zero, não utilizando assim, valores de ω <0,8. Os valores iniciais escolhidos para caracterização do estado no instante t=0 foram os mesmos sugeridos por Fernandes (1990), $a_0 = 10$ e $b_0 = 1$; desta forma, partimos de um nível inicial igual a 10. Após a simulação, cada uma das séries geradas foi submetida ao procedimentos de estimação por máxima verossimilhança dos 3 modelos : INAR(1)-Poisson , INAR(1)-Geométrica e Poisson-Gama. Para os modelos INAR(1), foi utilizada a função de verossimilhança condicionada ao conhecimento do valor

UNICAMP
BIBLIOTECA CENTRAL

inicial enquanto no modelo Poisson-Gama o procedimento só foi inicializado após a primeira observação diferente de zero.

4.1.2 - Afinidade entre as distribuições preditivas 1-passo-à-frente

Com o intuito de medir a afinidade entre a distribuição preditiva um-passo-a-frente do modelo simulado - aquele cujos parâmetros foram utilizados para a simulação da série - e a distribuição preditiva dos demais modelos ajustados, foi utilizada a medida de informação não paramétrica de Kullback-Leibler.

Considerando X e Y duas variáveis aleatórias discretas assumindo valores nos inteiros não negativos, com respectivas distribuições de probabilidades P e Q, a divergência direta de Kullback-Leibler (Kullback , 1959) é uma medida da distância da distribuição de Y em relação a distribuição de X, calculada da seguinte forma :

$$D(P:Q) = \sum_{i=1}^{n} p_i \, ln \frac{p_i}{q_i} \qquad (4.1.1)$$

$p_i = P(X = i)$
$q_i = P(Y = i)$

$\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} q_i = 1 \quad , \quad p_i, q_i > 0$

Observando a divergência (4.1.1) em cada ponto i, verificamos que quanto maior for a diferença $lnp_i - lnq_i$ maior será contribuição deste ponto para o aumento em D(P:Q). A situação extrema na qual $p_i = q_i$, para todo valor de i, a divergência calculada é igual a zero. A medida que o valor de D(P:Q) aumenta, menor é a afinidade entre as distribuições de X e Y[1]. Em nossa aplicação, a distribuição P é a distribuição preditiva 1-passo-à-frente verdadeira , que é encontrada utilizando os parâmetros escolhidos para a simulação da série, iremos medir a distâncias das demais em relação a esta.

[1] Existem outras medidas de divergências tais como a de Havrda Charvat e Csiszer sendo que a medida definida em (4.1.1) considera P como sendo a distribuição de referência para a comparação.

4.2 - COMPARAÇÃO UTILIZANDO SÉRIES REAIS

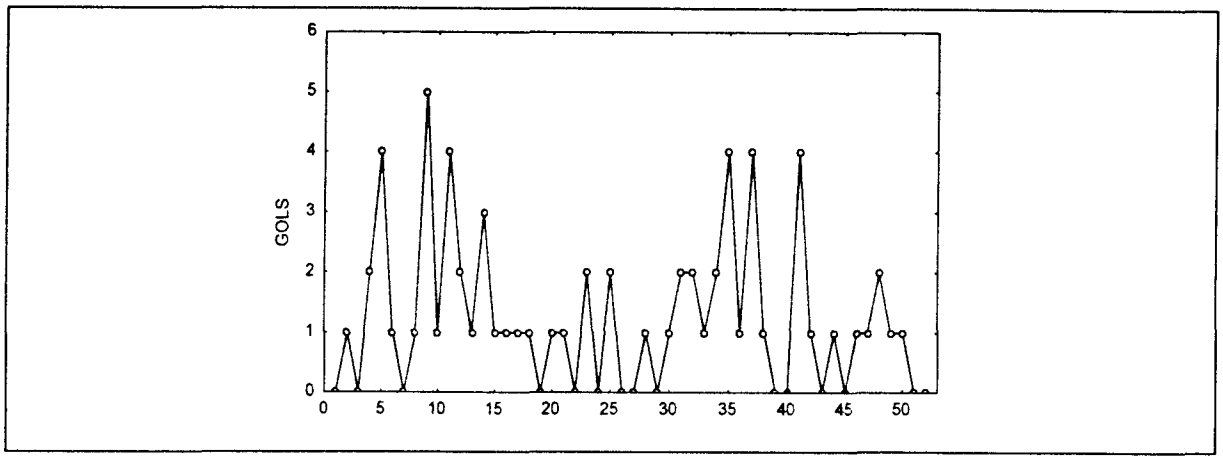
A comparação entre os modelos INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama será feita utilizando duas séries de contagens de pequena magnitude já citadas na literatura. A primeira série é retirada do artigo de Harvey e Fernandes (1989) e se refere ao número de gols marcados pela Inglaterra contra a Escócia (Figura 4.2.1) em partidas de futebol disputadas em Glasgow no período compreendido entre os anos de 1872 e 1987. Trataremos as observações desta série como se estas fossem equispaçadas, embora isto não aconteça pois elas sofrem interrupções devido a guerras e outros problemas políticos. A segunda série real a ser analisada é retirada do livro texto de McCleary e Hay (1980) e trata-se do número de bolsas furtadas no Hyde Park em Chicago em um período de 28 dias, sendo os resultados registrados a partir de janeiro de 1968.

O interesse no ajuste dos 3 modelos é compará-los qualitativamente, não realizando nehum esforço adicional de modelagem tal como buscar intervenções ou identificar fatores sazonais. Podemos encontrar em Harvey e Fernandes (1989) algumas tentativas para se chegar ao melhor modelo, utilizando estas séries.

4.2.1 - Série de gols marcados pela Inglaterra em partidas internacionais contra a Escócia.

A série de gols da Inglaterra (Figura 4.2.1) apresenta 52 observações e possui algumas características que dificultam a utilização de modelos gaussianos. Em primeiro lugar, a magnitude dos valores da série é extremamente pequena, visto que a série assume, no período das observações, apenas 6 estados. Adicionalmente, a série apresenta uma grande quantidade de valores iguais a zero que correspondem a aproximadamente 30 % das observações.

Figura 4.2.1 - Número de gols marcados pela Inglaterra contra a Escócia em partidas intrnacionais no período compreendido entre 1872 e 1987.



Nesta série, obtemos os seguintes resultados para as principais estatísticas descritivas :

Tabela 4.2.1 - Principais estatísticas descritivas da série de gols

série de gols	n	Média	Mediana	Moda	Variância	Desvio padrão	Coeficiente de variação (%)
	52	1,27	1,00	1,00	1,652	1,285	101,25 %

Nas 52 partidas realizadas entre Inglaterra e Escócia, ocorreram , em média, 1,27 gols por partida, com um coeficiente de variação um pouco maior que 100%. Em uma ocasião apenas, houve 5 gols na partida e em 29 % das partidas, a Inglaterra marcou mais de um gol. Para ter uma idéia do comportamento do número de gols em partidas consecutivas e desta forma acessar algum tipo de informação sobre a dependência temporal entre os valores , montamos a tabela 4.2.2 .

Tabela 4.2.2 - Frequência absoluta e relativa (percentual em relação ao valor anterior) de Gols em partidas consecutivas.

Z(t-1)	Z(t)	0	1	2	≥3	Totais
0		3 21,4 %	9 64,3 %	2 14,3 %	0 0%	14 27,5 %
1		7 31,8%	5 25,4 %	3 13,6-%	6 27,2 %	22 43,12 %
2		3 37,5%	3 37,5%	1 12,5%	1 12,5%	8 15,7 %
≥3		1 14,3%	4 57,1 %	2 28,9%	0 0%	7 13,7 %
					Total	51

Z(t) : número de gols no instante t

Verificamos na tabela 4.2.2. que exceto nas 6 situações em que a Inglaterra marcou 3 ou mais gols após ter marcado 1 gol no jogo anterior, o número de gols marcados na partida anterior não aparenta, isoladamente, causar muita influência no valor observado, embora na Figura 4.2.1 podemos até suspeitar que ocorreram algumas mudanças de nível no período das observações.

Tabela 4.2.3 - Ajuste dos modelos à séries de gols marcados pela Inglaterra contra a Escócia

Modelos	Estimativas para os parâmetros	AIC[2]	BIC	SQR	QMR
INAR(1)-Poisson	$\hat{\alpha}$ =0,074	158,61	162,51	83,33	1,70
	$\hat{\lambda}$ = 1,284				
INAR(1)-Geométrica	$\hat{\phi}$ = 0,178	165,43	169,33	85,75	1,75
	$\hat{\theta}$ = 0,485				
Poisson-Gama	$\hat{\omega}$ = 0,844	158,68	160,63	92,94	1,86

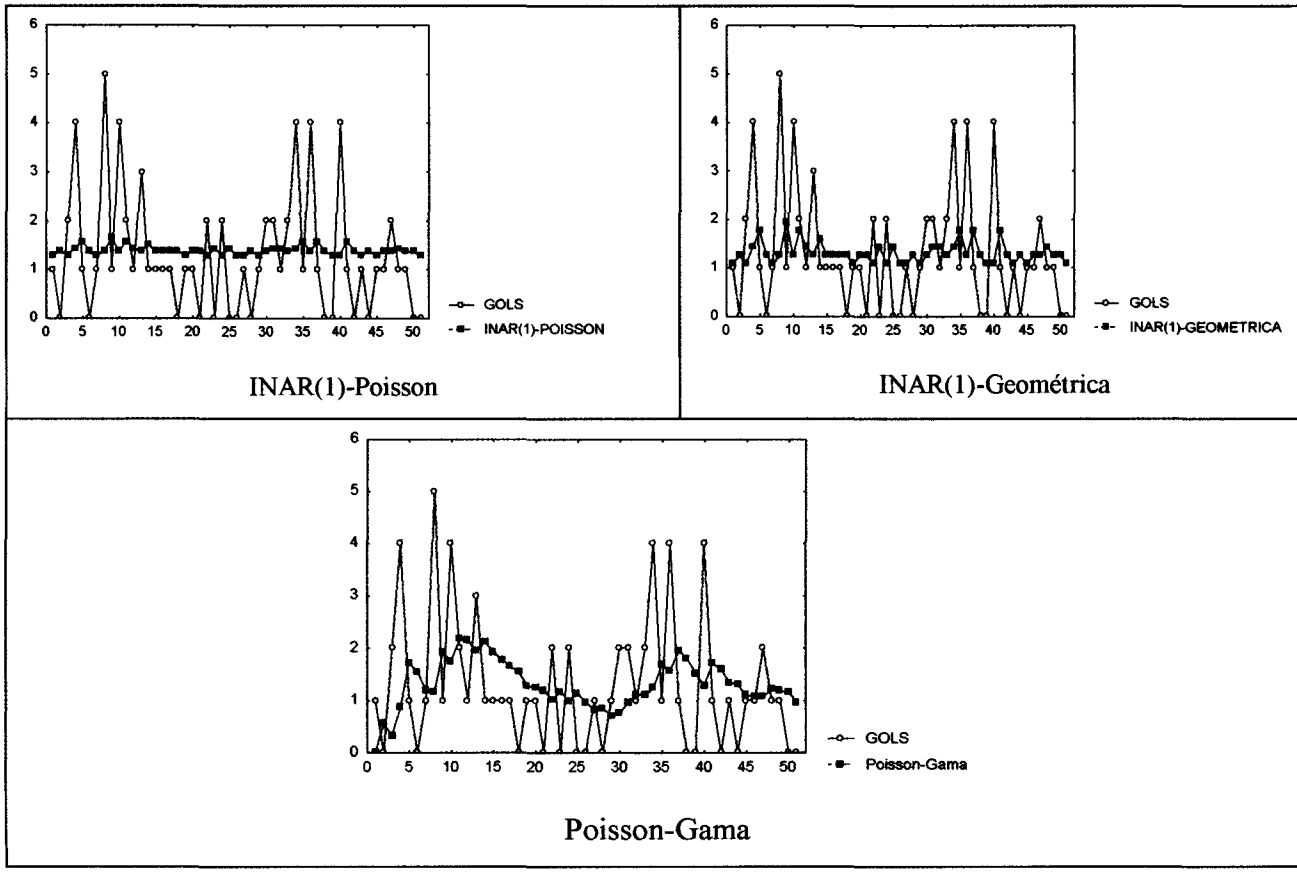
SQR - Soma dos Quadrados dos Resíduos.

QMR - Quadrado Médio dos Resíduos.

[2]Utilizamos os seguintes critérios AIC = -2log(ML)+2p , BIC = -2log(ML)+plog(n) p: número de parâmetros independentes do modelo ajustado , n é o tamanho da série observada e ML é a função de máxima verossimilhança .

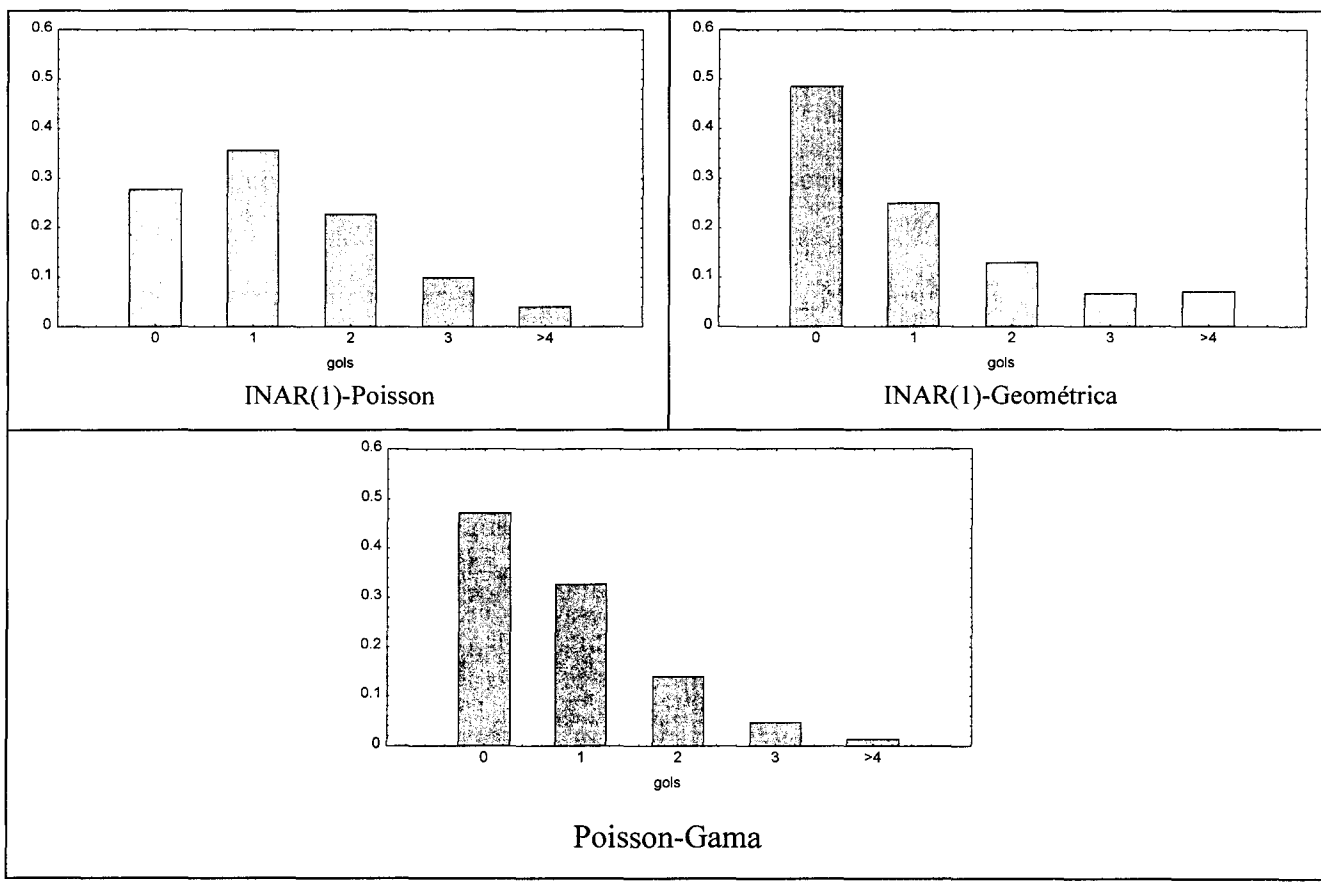
O modelo INAR(1)-Poisson ajustado, descreve o comportamento da série como sendo praticamente uma sequência de inovações de Poisson com média $\hat{\lambda}$ = 1,284, visto que a estimativa para parâmetro α, que caracteriza a dependência temporal na série, esteve bem próxima de zero. Por sua vez, o modelo INAR(1)-Geométrica descreve a série como se esta tivesse uma correlação serial ($\hat{\phi}$=0,178) superior a identificada no modelo anterior, porém com uma média inferior para as inovações. Se pensarmos em termos de um alisamento exponencial, o valor alto para a estimativa do hiperparâmetro do modelo Poisson-Gama demonstra que o ponderação dada aos últimos valores é pequena, o que nesta situação é bastante razoável pois a série não apresenta fortes oscilações no seu nível. Pelo critério AIC, o modelo selecionado foi o INAR(1)-Poisson enquanto pelo critério BIC, o modelo selecionado dentre os 3 é o modelo Poisson-Gama. Em ambos os critérios, o modelo INAR(1)-Geométrica foi a pior alternativa para descrever o comportamento da série.

Figura 4.2.2 - Comportamento da previsão pontual 1-passo-à-frente dos modelos ajustados à série de gols marcados pela Inglaterra.



A Figura 4.2.2 apresenta a previsão pontual 1-passo-à-frente obtida a partir do ajuste dos 3 modelos. Podemos observar que as previsões feitas pelos modelos INAR(1) são praticamente constantes pelo fato das observações apresentarem baixa correlação serial , enquanto o ajuste do modelo Poisson-Gama indica algumas possíveis mudanças no nível da série, principalmente em dois períodos das observações onde houve um número maior de gols marcados pela Inglaterra. Embora isto tivesse acontecido, verificamos na tabela 4.2.3 que a média dos quadrados de resíduos foi maior no modelo Poisson-Gama. Podemos encontrar uma explicação para este fato, observando que quando o nível descrito pelo modelo aumenta em função de alguns picos, este demora a retornar ao patamar usual da série, que é aquele descrito pelos modelos INAR(1), e em função disto, os erros de previsão 1-passo-à-frente no ajuste são grandes nestes pontos. A Figura 4.2.3 nos traz uma idéia das distribuições preditivas 1-passo-à-frente dos 3 modelos após a última observação.

Figura 4.2.3 - Distribuições preditivas 1-passo-à-frente dos modelos ajustados para a série de gols da Inglaterra



Podemos notar na Figura 4.2.3 a semelhança entre as distribuições preditivas dos modelos INAR(1)-Geométrica e Poisson-Gama. Nestes modelos , a probabilidade de não haver gols nas próximas partidas é muito alta, próxima de 0,5, enquanto no histórico da série, em aproximadamente 28% dos jogos, a Inglaterra não marcou gols. Como na última partida a Inglaterra não marcou gols, a distribuição preditiva para a próxima partida nos modelos INAR(1)-Poisson e INAR(1)-Geométrica são as próprias distribuições Poisson e geométrica. Segundo a distribuição preditiva do modelo INAR(1)-Poisson , na próxima partida , a probabilidade de haver gols é de 72,3 %. De uma forma geral, as 3 distribuições preditivas se comportaram de acordo com o histórico da série mostrado na Tabela 4.2.2 onde foi constatado que a ocorrência maior na série foi de partidas consecutivas com poucos gols (0 ou 1).

4.2.2 - Série de furtos em Hyde Park

A segunda série a ser analisada, é referente ao número de furtos de bolsas em um bairro de Chicago (Figura 4.2.4) . A série apresenta os furtos totais a cada período de 28 dias entre janeiro de 1968 até novembro de 1973 em um total de 71 observações. Ao contrário da série de gols, esta série não apresenta uma magnitude tão pequena,tanto que em McCleary e Hay (1980) foi possível ajustar um modelo AR(2) sem recorrer a transformações. Na tabela 4.2.4, apresentamos as estimativas dos parâmetros dos modelos ajustados a esta série.

Figura 4.2.4 - Série de furtos em Hyde Park em Chicago no período compreendido entre janeiro de 1968 e novembro de 1973.

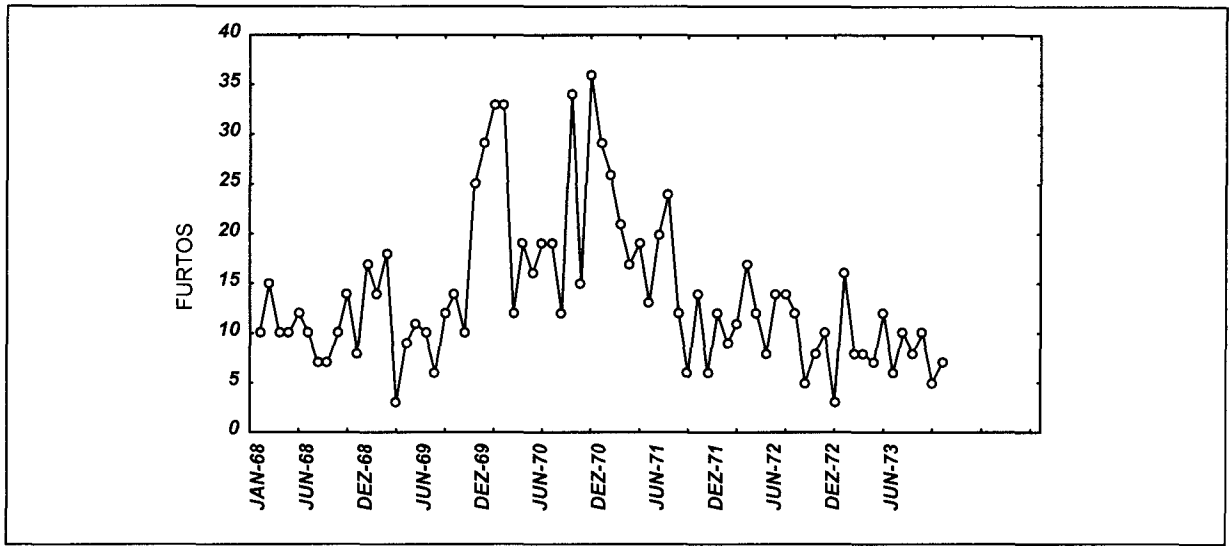


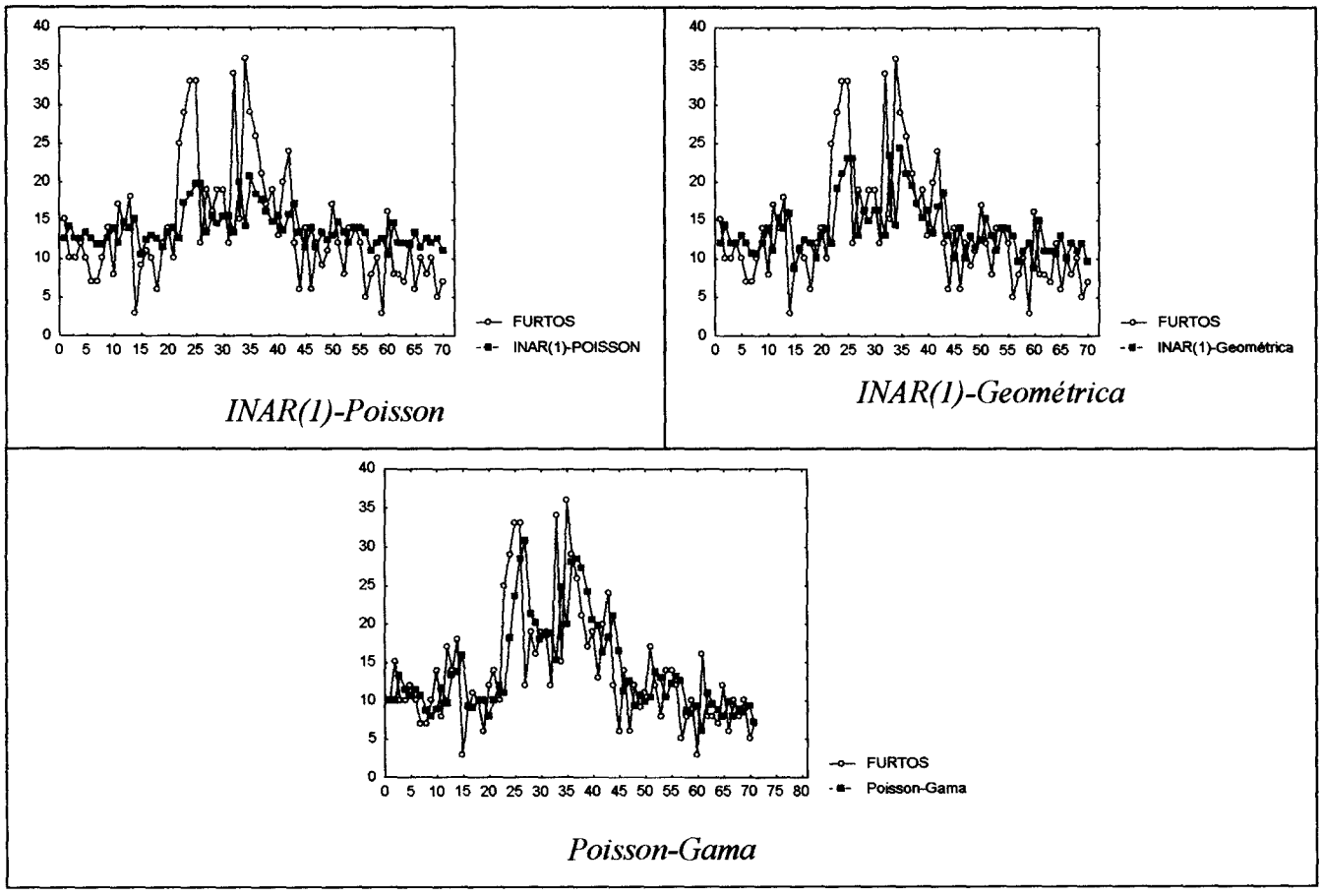
Tabela 4.2.4 - Ajuste dos modelos a séries de furtos em Hyde Park

Modelos	Estimativas para os parâmetros	AIC	BIC	SQR	QMR
INAR(1)-Poisson	$\hat{\alpha} = 0{,}3056$	526,70	531,23	3088,90	44,77
	$\hat{\lambda} = 9{,}5897$				
INAR(1)-Geométrica	$\hat{\phi} = 0{,}4764$	517,84	522,36	2913,10	42,22
	$\hat{\theta} = 0{,}1217$				
Poisson-Gama	$\hat{\omega} = 0{,}4943$	444,75	447,02	2749,20	39,84

Conforme os resultados encontrados em McCleary e Hay (1980), já era esperado encontrar valores nos modelos INAR(1) que indicassem a presença de alguma autocorrelação de primeira ordem entre os valores. Tal como aconteceu na série de gols , o modelo INAR(1)-Geométrica identificou uma correlação serial maior do que o modelo INAR(1)-Poisson que pode ser verificada através da estimativa $\hat{\phi}$ = 0,476. O hiperparâmetro estimado ($\hat{\omega}$ = 0,494) para o modelo Poisson-Gama reflete a variabilidade da previsão 1-passo-à-frente deste modelo que é bem maior nesta série do que na situação envolvendo a série de gols. Ao pensarmos em termos de um alisamento exponencial, como

a estimativa do nível feita pelo modelo Poisson-Gama dá peso próximo a 0,5 para a última observação, esta acompanha a evolução do valor observado, e como ela é a previsão 1-passo-à-frente esta apresentará um valor próximo valor observado com 1 lag de defasagem.

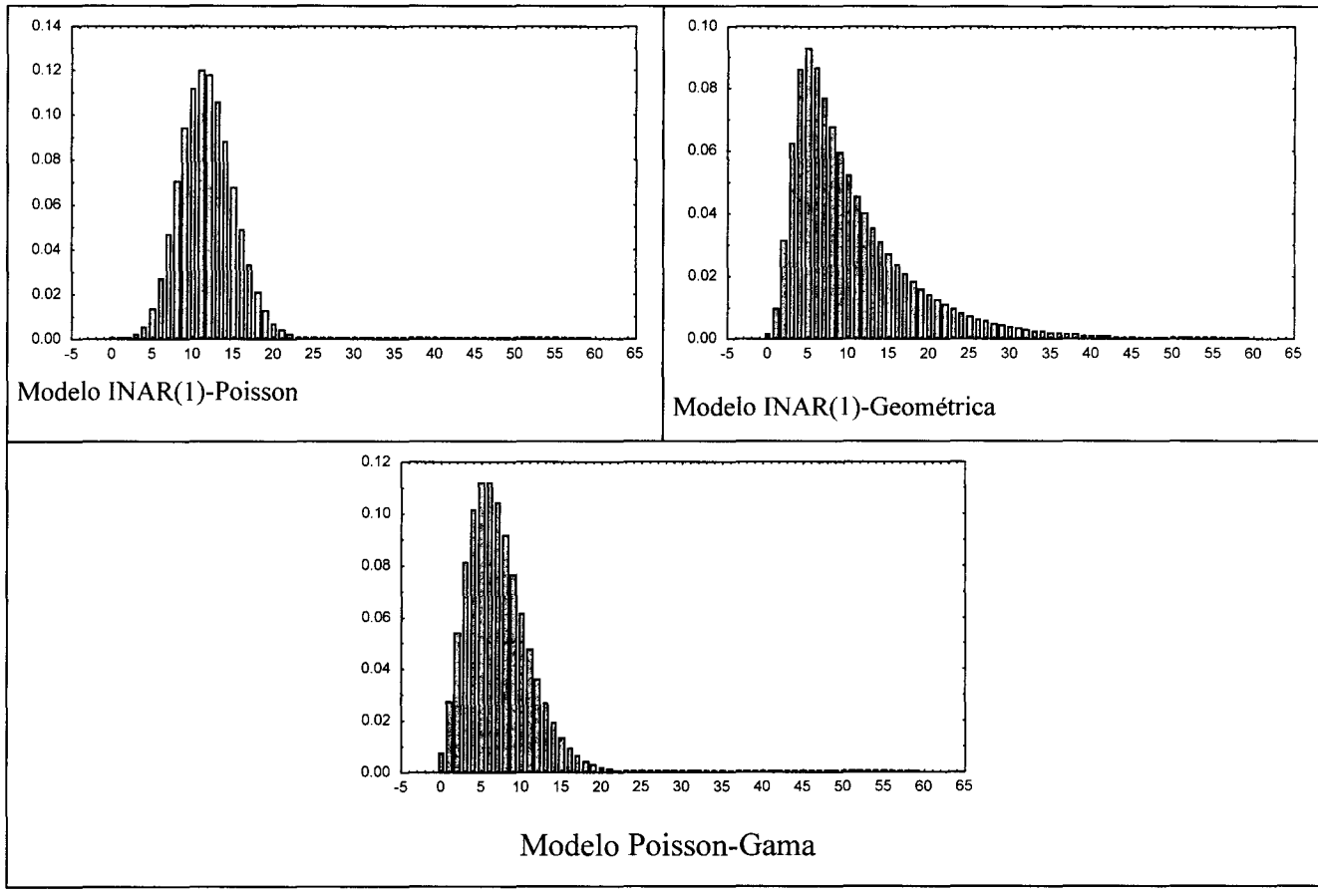
Figura 4.2.5 - Previsão pontual 1-passo-à-frente obtida pelos modelos na série de furtos em Hyde Park



Podemos verificar na Figura 4.2.5 a confirmação dos resultados obtidos na tabela 4.2.4 na qual o modelo Poisson-Gama apresentou um desempenho muito superior aos demais nos critérios AIC e BIC. De fato, percebemos que este modelo é o único dentre os 3 cujo o ajuste através da previsão 1-passo-à-frente consegue acompanhar as mudanças no decorrer da série. Por outro lado, dentre os modelos INAR(1), verificamos que o modelo

INAR(1)-Geométrica é o que reage melhor aos períodos de mudanças na série, fato que podemos também constatar na tabela 4.2.4 através do quadrado médio dos resíduos que é inferior a apresentada pelo modelo INAR(1)-Poisson. A Figura 4.2.6 apresenta as distribuições preditivas 1-passo-à-frente para que possamos ter uma idéia da incerteza das predições que serão feitas pelos 3 modelos para o próximo passo.

Figura 4.2.6 - Distribuição preditiva 1-passo-à-frente dos modelos ajustados à série de furtos em Hyde Park.



Podemos notar na Figura 4.2.6 que as distribuições preditivas 1-passo-à-frente dos modelos INAR(1)-Poisson e Poisson-Gama apresentam características semelhantes, havendo uma leve assimetria na distribuição do modelo Poisson-Gama, e ambas diferem da distribuição encontrada através do modelo INAR(1)-Geométrica que exibe clara assimetria e maior variância.

4.3 - RESULTADOS OBTIDOS A PARTIR DA SIMULAÇÃO DO PROCESSO INAR(1)-POISSON

4.3.1 - Ajuste de modelos

Nesta seção vamos estudar a aproximação dos modelos INAR(1)-Poisson pelos modelos INAR(1)-Geométrica e Poisson-Gama. Inicialmente, para melhor entender a relação entre os modelos e as estimativas de máxima verossimilhança, na presença de erros de especificação, iremos estudar as estimativas obtidas para os parâmetros dos 3 modelos, e em seguida estudaremos o ajuste destes modelos .

Primeiramente, consideremos a comparação envolvendo o modelo INAR(1)-Poisson e INAR(1)-Geométrica. No modelo INAR(1)-Poisson temos :

$$Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim \text{Poisson}\ (\lambda)$$

enquanto no modelo INAR(1)-Geométrica temos :

$$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$$

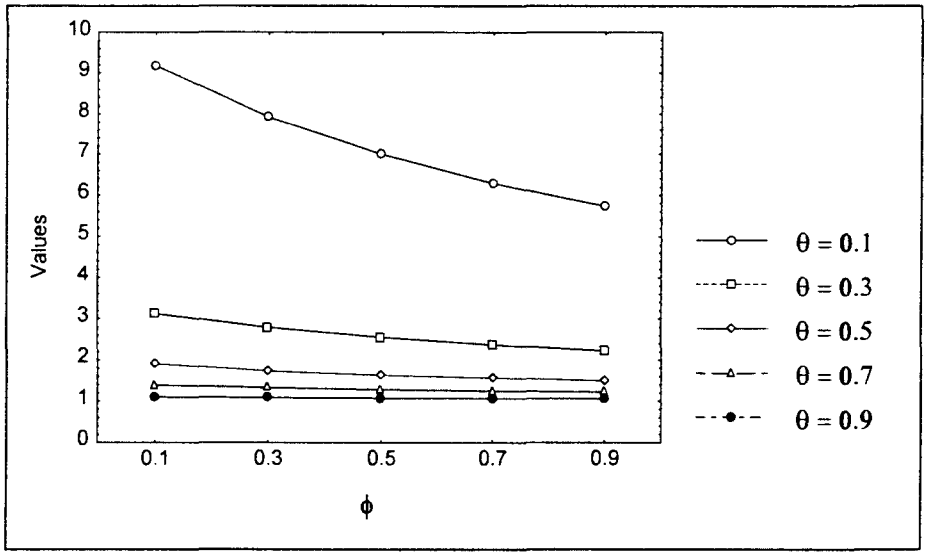
$$\varepsilon_t \sim \text{geométrica}\ (\theta)$$

Conforme visto na seção 2.2, a esperança e variância incondicionais destes processos são dadas por :

	INAR(1)-Poisson	INAR(1)-Geométrica
$E(Z_t)$	$\dfrac{\lambda}{(1-\alpha)}$	$\dfrac{(1-\theta)}{\theta(1-\phi)}$
$Var(Z_t)$	$\dfrac{\lambda}{(1-\alpha)}$	$\dfrac{(1-\theta)(1+\theta\phi)}{\theta^2(1-\phi^2)}$

A utilização dos estimadores de Yule-Walker, descritos na seção 2.3.1, significa encontrar estimativas tais que $\hat{\alpha} = \hat{\phi}$. Entretanto, não é possível igualar simultaneamente a função de autocorrelação, média e variância dos dois processos, pois ao contrário da distribuição Poisson, as inovações da distribuição geométrica e conseqüentemente o processo INAR(1)-Geométrica exibem sobredispersão (ver Figura 4.3.1) , salvo situações nas quais $\theta \approx 1$ e como conseqüência as inovações estão próximas de zero.

Figura 4.3.1 - Razão entre a variância e a média do processo INAR(1)-Geométrica



Conforme a Figura 4.3.1. verificamos que quando θ assume valores altos. o processo INAR(1)-Geométrica se aproxima do comportamento do processo INAR(1)-Poisson no que diz respeito a razão entre a variância e a média. Podemos verificar o comportamento destes processos através dos gráficos apresentados na seção 2.7.

Ao utilizarmos o método dos momentos, podemos igualar duas das características do processo, como por exemplo a função de autocorrelação e a esperança, obtendo :

$$\phi = \alpha$$

$$\theta = \frac{1}{1+\lambda}$$

ou podemos igualar a função de autocorrelação e a variância,

$$\phi = \alpha$$

$$\lambda = \frac{1-\theta}{\theta^2}$$

Ao simularmos o processo INAR(1)-Poisson e estimarmos os parâmetros do processo INAR(1)-Geométrica por máxima verossimilhança, esperamos que as estimativas sejam influenciadas pelos 3 fatores: as escolhas de α, λ e o tamanho da série, não sendo assim estabelecidas as relações acima.

Ao compararmos o modelo INAR(1)-Poisson com o modelo Poisson-Gama , a situação se torna mais complicada. Sabemos, por exemplo, que situações nas quais $\alpha \approx 0$, o processo INAR(1)-Poisson pode ser considerado uma seqüência de inovações com distribuição Poisson(λ). Podemos pensar que o processo Poisson-Gama com fator de desconto igual a zero, ou seja, quando o hiperparâmetro $\omega = 1$, tal situação também ocorre. Entretanto este argumento não é o bastante para afirmar que existe uma relação decrescente entre α e ω, pois devemos considerar os outros fatores : a média das inovações e o tamanho da amostra, principalmente este último pois os parâmetros que caracterizam o estado irão convergir em grandes amostras.

4.3.1.1 - Ajuste do modelo INAR(1)-Geométrica ao processo INAR(1)-Poisson simulado.

A Figuras 4.3.2 e 4.3.3 referem-se aos resultados obtidos na tentativa de ajustar o modelo INAR(1)-Geométrica aos dados simulados pelo processo INAR(1)-Poisson. Nelas encontramos a médias estimadas para os dois parâmetros destes modelos nos diferentes processos INAR(1)-Poisson. Uma tabela mais detalhada com estes resultados é apresentada no apêndice A.

Figura 4.3.2 - Estimativas médias em 20 replicações obtidas para o parâmetro ϕ do modelo INAR(1)-Geométrica a partir da simulação do processo INAR(1)-Poisson com parâmetros α e λ

processo simulado : $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$, $\varepsilon_t \sim$ Poisson(λ)

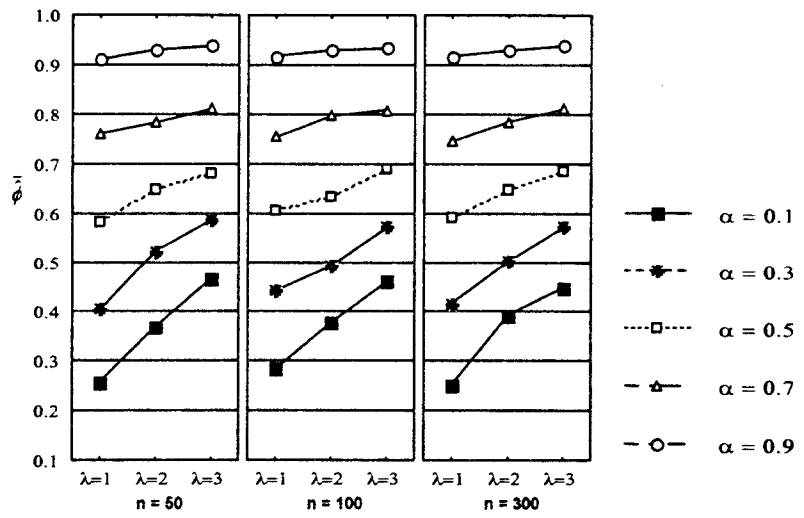
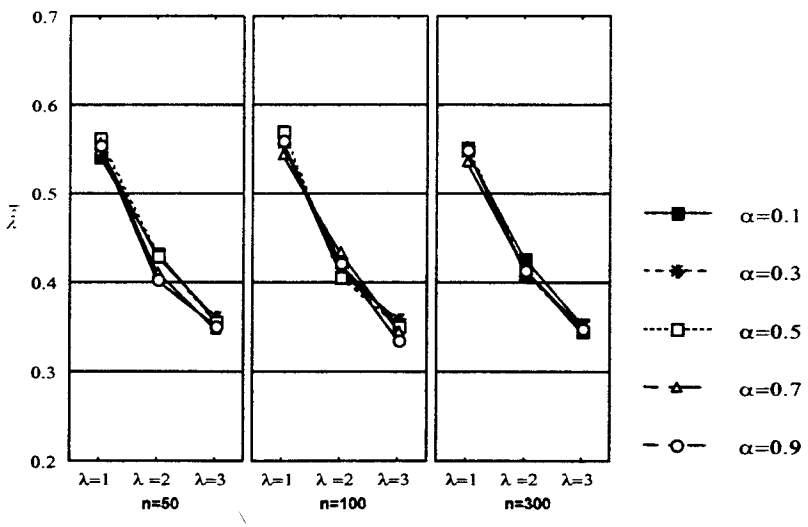


Figura 4.3.3 - Estimativas médias em 20 replicações obtidas para o parâmetro θ do modelo INAR(1)-Geométrica a partir da simulação do processo INAR(1)-Poisson com parâmetros α e λ.

processo simulado : $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$, $\varepsilon_t \sim$ Poisson(λ)

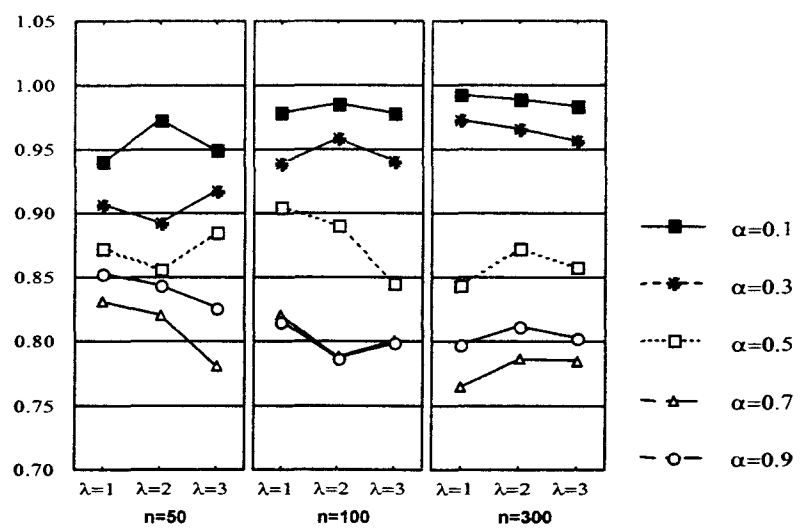


Ao utilizarmos o procedimento de máxima verossimilhança para ajustar o modelo INAR(1)-Geométrica aos dados simulados pelo processo INAR(1)-Poisson, os resultados obtidos procuraram aproximar a média do processo simulado e como conseqüência, as estimativas encontradas para o parâmetro ϕ não refletem a correlação serial dos valores simulados, assim como a média das inovações. Nas Figuras 4.3.2 e 4.3.3, podemos verificar que o modelo INAR(1)-Geométrica ajustado sugere uma correlação serial maior do que aquela utilizada para a simulação e uma média menor para as inovações , sendo que a medida que α se aproxima de 1 e λ é pequeno , esta diferença vai diminuindo e a estimação por máxima verossimilhança acaba por aproximar as duas características : autocorrelação e média das inovações. Na figura 4.3.3 observamos que a estimativa para o parâmetro da distribuição geométrica não depende da escolha de α e n, mas somente da média das inovações simuladas , representada por λ.

4.3.1.2 - Ajuste do modelo Poisson -Gama ao processo INAR(1)-Poisson simulado.

Figura 4.3.4 - Estimativas médias em 20 replicações obtidas para o hiperparâmetro ω do modelo Poisson-Gama a partir da simulação do processo INAR(1)-Poisson com parâmetros α e λ.

processo simulado $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$, $\varepsilon_t \sim$ Poisson(λ)



Conforme a discussão feita na seção 4.3 , os resultados indicam que quanto maior for a correlação serial dos valores simulados, menor o valor estimado para o hiperparâmetro do modelo Poisson-Gama. Entretanto, podemos constatar na figura 4.3.4 que os outros fatores controlados na simulação, média das inovações e tamanho da série também exercem algum tipo de influência nos resultados. A figura mostra que não há um padrão de comportamento bem definido de ω em função de λ, parâmetro de locação do processo INAR(1)-Poisson, mas indica que em simulações de séries grandes, os resultados se tornam mais extremos pois há um aumento na diferença entre os hiperparâmetros estimados para séries com correlação serial alta e aqueles estimados para séries com correlação baixa

4.3.2 - Estatísticas principais da distribuição preditiva 1-passo-à-frente

Nesta seção, desejamos comparar as distribuições preditivas dos 3 modelos ajustados com a distribuição preditiva real que é aquela obtida a partir do modelo com os parâmetros utilizados no PGD . Sendo o INAR(1)-Poisson o verdadeiro PGD ele é aqui denominado de modelo real, e esperamos que a sua distribuição preditiva esteja mais próxima da distribuição preditiva verdadeira. Verificaremos , primeiramente, a média e variância da distribuição preditiva 1-passo-à-frente e em seguida, medimos o erro médio quadrático obtido nas 20 replicações. Nesta seção, o modelo real ea distribuição preditiva real, são referências para comparação, pois estaremos verificando em que aspectos as distribuições preditivas estimadas pelos modelos incorretos estão se distanciando daquela que seria a distribuição preditiva do melhor modelo e a distribuição preditiva do modelo real.

4.3.2.1 - Média da distribuição preditiva 1-passo-à-frente

A primeira comparação a ser realizada envolve as médias das distribuições preditivas 1-passo-à-frente. A importância em estudá-las está no fato de que em muitas situações, estas são utilizadas como previsões pontuais 1-passo-à-frente.

$$\hat{Z}_{j,n+1} = E[Z_{j,n+1}|\mathbf{Z}_n]$$

$\hat{Z}_{j,n+1}$: previsão pontual 1-passo-à-frente obtida na j-ésima replicação.

Na Figura 4.3.3, apresentamos o diagrama de dispersão que possibilita visualizar o desvio de $\hat{Z}_{j,n+1}$ do modelo ajustado em relação ao real, ou seja, o quanto a média da distribuição preditiva de cada modelo se afastou da média verdadeira. Para uma maior compreensão destes resultados, apresentamos na tabela 4.3.2.1 os resultados para o erro relativo de cada modelo nas estimativas da média da distribuição preditiva 1-passo-à-frente.

$$\text{ERM} = \sum \frac{1}{20} \frac{|\hat{Z}_{j,n+1} - E(Z_{j,n+1})|}{E(Z_{j,n+1})} \text{ : erro relativo médio}$$

$E(Z_{j,n+1})$: média da distribuição preditiva do modelo real

Figura 4.3.5 - Diagrama de dispersão entre as médias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a média verdadeira na simulação do processo INAR(1)-Poisson.

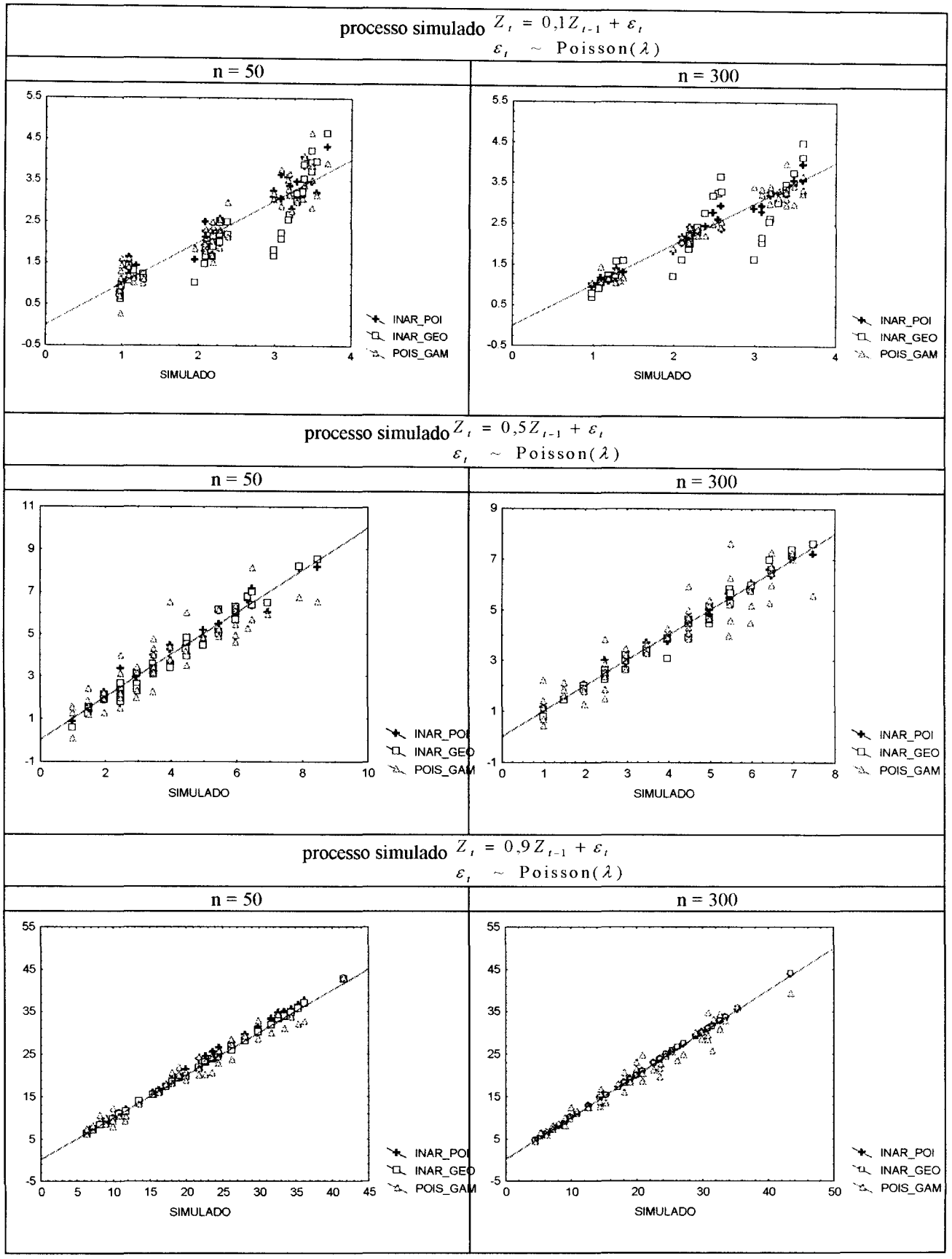


Tabela 4.3.1 - Erro relativo médio cometido ao utilizar a média da distribuição preditiva 1-passo-à-frente dos modelos ajustados ao invés da média da distribuição verdadeira.

α	λ	n	INAR(1)-Poisson	INAR(1)-Geométrica	Poisson-Gama
0,1	1	50	0,141	0,177	0,230
	1	100	0,102	0,145	0,160
	1	300	0,066	0,145	0,078
	2	50	0,113	0,174	0,123
	2	100	0,062	0,141	0,085
	2	300	0,050	0,170	0,041
	3	50	0,082	0,181	0,105
	3	100	0,061	0,126	0,073
	3	300	0,029	0,141	0,070
0,3	1	50	0,135	0,150	0,275
	1	100	0,099	0,113	0,218
	1	300	0,051	0,097	0,166
	2	50	0,067	0,115	0,108
	2	100	0,082	0,089	0,121
	2	300	0,038	0,102	0,158
	3	50	0,080	0,114	0,133
	3	100	0,050	0,094	0,116
	3	300	0,033	0,104	0,140
0,5	1	50	0,121	0,135	0,324
	1	100	0,064	0,085	0,263
	1	300	0,047	0,088	0,281
	2	50	0,083	0,093	0,182
	2	100	0,048	0,054	0,151
	2	300	0,027	0,055	0,132
	3	50	0,043	0,061	0,156
	3	100	0,039	0,090	0,205
	3	300	0,022	0,051	0,106
0,7	1	50	0,081	0,079	0,251
	1	100	0,036	0,031	0,188
	1	300	0,028	0,048	0,219
	2	50	0,038	0,030	0,131
	2	100	0,029	0,041	0,103
	2	300	0,017	0,041	0,159
	3	50	0,027	0,034	0,092
	3	100	0,020	0,032	0,102
	3	300	0,013	0,028	0,106
0,9	1	50	0,015	0,017	0,110
	1	100	0,012	0,009	0,104
	1	300	0,007	0,009	0,090
	2	50	0,022	0,009	0,056
	2	100	0,011	0,015	0,077
	2	300	0,006	0,010	0,090
	3	50	0,048	0,010	0,055
	3	100	0,010	0,007	0,064
	3	300	0,005	0,007	0,068

Conforme os resultados obtidos na seção 4.3.1 , quando foram exibidas as estimativas obtidas por máxima verossimilhança e o ajuste dos modelos, verificamos que este baseou-se principalmente na média do PGD (Processo Gerador de Dados). Logo, de encontro com estes resultados, as figuras mostram que, em média, as médias das distribuições preditivas se apresentam bem próximas e aleatoriamente dispersas em torno do "verdadeiro" valor , havendo no entanto diferenças nas variabilidades. Com a ajuda da tabela 4.3.1, notamos que as diferenças entre o modelo INAR(1)-Geométrica e o modelo real são pequenas para tamanhos de amostra n = 50 e valores pequenos de λ . A medida em que aumentamos a correlação serial, notamos que estes dois modelos se aproximam bastante, e em especial quando ϕ = 0,9 a diferença se torna quase imperceptível. Duas características da simulação tornam distantes os modelos Poisson-Gama do modelo real, em relação média da distribuição preditiva: valores intermediários de α e valores pequenos de λ . Além destes fatores, séries de tamanho grande também contribuem para que esta diferença seja maior. Com base nestes resultados, constatamos que as previsões pontuais dos 3 modelos, utilizando a média da distribuição preditiva, estarão mais próximas , quanto maior for a correlação serial dos valores.

4.3.2.2 -Variância da distribuição preditiva 1-passo-à-frente

Conforme a série do número de furtos analisada na seção 4.2, podemos constatar graficamente, naquela ocasião, a diferença entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos propostos. Dentre os modelos analisados, o INAR(1)-Geométrica é o único que pressupõe que os dados exibem sobredispersão no instante t. Portanto, neste item de comparação, é esperada alguma diferença deste modelo em relação aos demais.

Cabe salientar que a variância da distribuição preditiva fornece a incerteza do valor predito. Logo, ao encontramos em um modelo, uma variância maior do que a variância real, o procedimento não consegue utilizar de forma adequada o histórico da série. Em caso contrário, quando a variância do modelo ajustado é menor do que deveria ser, estamos subestimando a incerteza existente no predição.

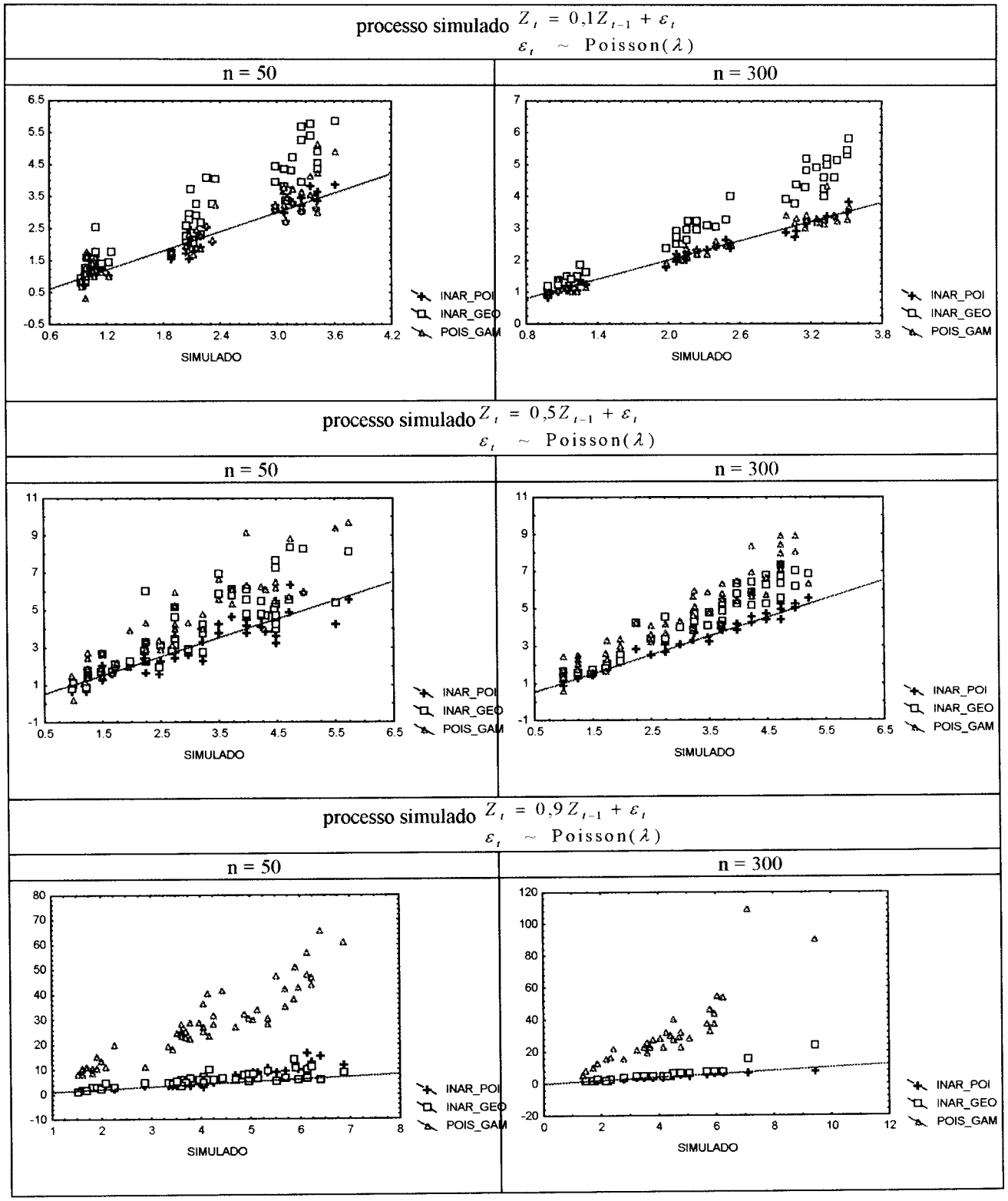
Verificamos nesta seção as médias e variâncias amostrais da variância das distribuições preditivas obtidas nos ajustes destes modelos às 20 séries simuladas de cada processo INAR(1)-Poisson, ou seja, estimamos com base nas simulações :

$$\mathrm{Var}(\hat{Z}_{n+1}) = \mathrm{Var}[Z_{n+1} | Z_n]$$

$S^2_{j,\hat{Z}_{n+1}}$: variância da distribuição preditiva na j-ésima replicação0.

Iremos utilizar o mesmo recurso gráfico que utilizamos anteriormente para a média com o intuito de verificar os desvios de $S^2_{j,\hat{Z}_{n+1}}$ em relação as variâncias do modelo considerado como o correto

Figura 4.3.6 - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Poisson.



A figura 4.3.6 mostra o comportamento diferenciado das distribuições preditivas em função da correlação serial utilizada na simulação do processo INAR(1)-Poisson. Podemos verificar que quando α é pequeno, isto é, há baixa correlação serial , a variância da distribuição preditiva do modelo INAR(1)-Geométrica é nítidamente maior que as demais e esta diferença aumenta a medida em que aumentamos o valor de λ na simulação. Os modelos INAR(1)-Poisson e Poisson-Gama apresentam resultados semelhantes nestas séries. O valor intermediário, $\alpha = 0.5$, faz com que as variâncias apresentadas pelas distribuições dos modelos INAR(1)-Geométrica e Poisson-Gama estejam próximas, porém ambas muito acima do modelo real. Neste caso, apenas o modelo real apresentou valores próximos aos corretos0, Séries com alta correlação serial fizeram com que os resultados dos modelos INAR(1)-Geométrica e INAR(1)-Poisson estivessem bem próximos, e ambos praticamente coincidentes com a reta que traduz a situação correta, principalmente para séries de tamanho grande. Em contrapartida, estas séries produziram um aumento considerável na variância da distribuição preditiva do modelo Poisson-Gama, principalmente para valores altos de λ. Os resultados mostraram de um modo geral que quando os modelos incorretos, INAR(1)-Geométrica e Poisson-Gama, não utilizam adequadamente as informações para construir a variância da distribuição preditiva, estes superestimam a variância verdadeira. O prejuízo em utilizar o modelo Poisson-Gama ao invés do modelo real é pequeno para séries simuladas com baixa correlação serial. Quando a correlação serial é alta e o tamanho da série e grande, quase não há prejuízo na variância da distribuição preditiva ao utilizar alternativamente o modelo INAR(1)-Geométrica.

Ao avaliarmos a média e a variância da distribuição preditiva 1-passo-à-frente dos modelos ajustados, encontramos uma forte indicação de que o modelo INAR(1)-Geométrica, dentre os modelos incorretamente utilizados, representa uma alternativa confiável para previsão 1-passo-à-frente em processos INAR(1)-Poisson com alta correlação serial. De forma menos contundente, o modelo Poisson-Gama pode ser utilizado alternativamente quando a correlação serial neste processo é pequena.

4.3.2.3 - Erro Quadrático Médio da Previsão 1-passo-à-frente

Outra estatística de interesse a ser analisada é o erro da previsão pontual 1-passo-à-frente utilizando a média da distribuição preditiva. Para que este pudesse ser obtido, utilizamos a semente armazenada na geração da série para simular o n+1-ésimo valor e assim subtraí-lo do valor predito.

$e_{j,n+1} = (\hat{Z}_{j,n+1} - Z_{j,n+1})^2$: erro de previsão na j-ésima replicação

Neste item comparamos o erro quadrático médio obtido por cada modelo nas 20 replicações do processo simulado, ou seja :

$$EQMP_{n+1} = \sum_{J=1}^{20} e_{j,n+1} \; / \; 20$$

Nas seções anteriores, foi possível verificar algumas deficiências no ajuste de modelos incorretos em relação a média e variância da distribuição preditiva. A importância de medir o EQMP é verificar o que acontecerá quando o valor que foi predito puder ser comparado com os valor observado. As Figuras 4.3.7a,b mostram o comportamento do EQMP em relação a correlação serial α utilizada na simulação . Nestes gráficos, ilustraremos apenas a média obtida para esta estatística nas 20 replicações de cada processo.

Figura 4.3.7a - Comportamento do erro quadrático - média em 20 replicações - obtido ao utilizar como previsão pontual 1-passo-à-frente a média das distribuições preditivas 1-passo-à-frente dos modelos ajustados ao processo INAR(1)-Poisson simulado.

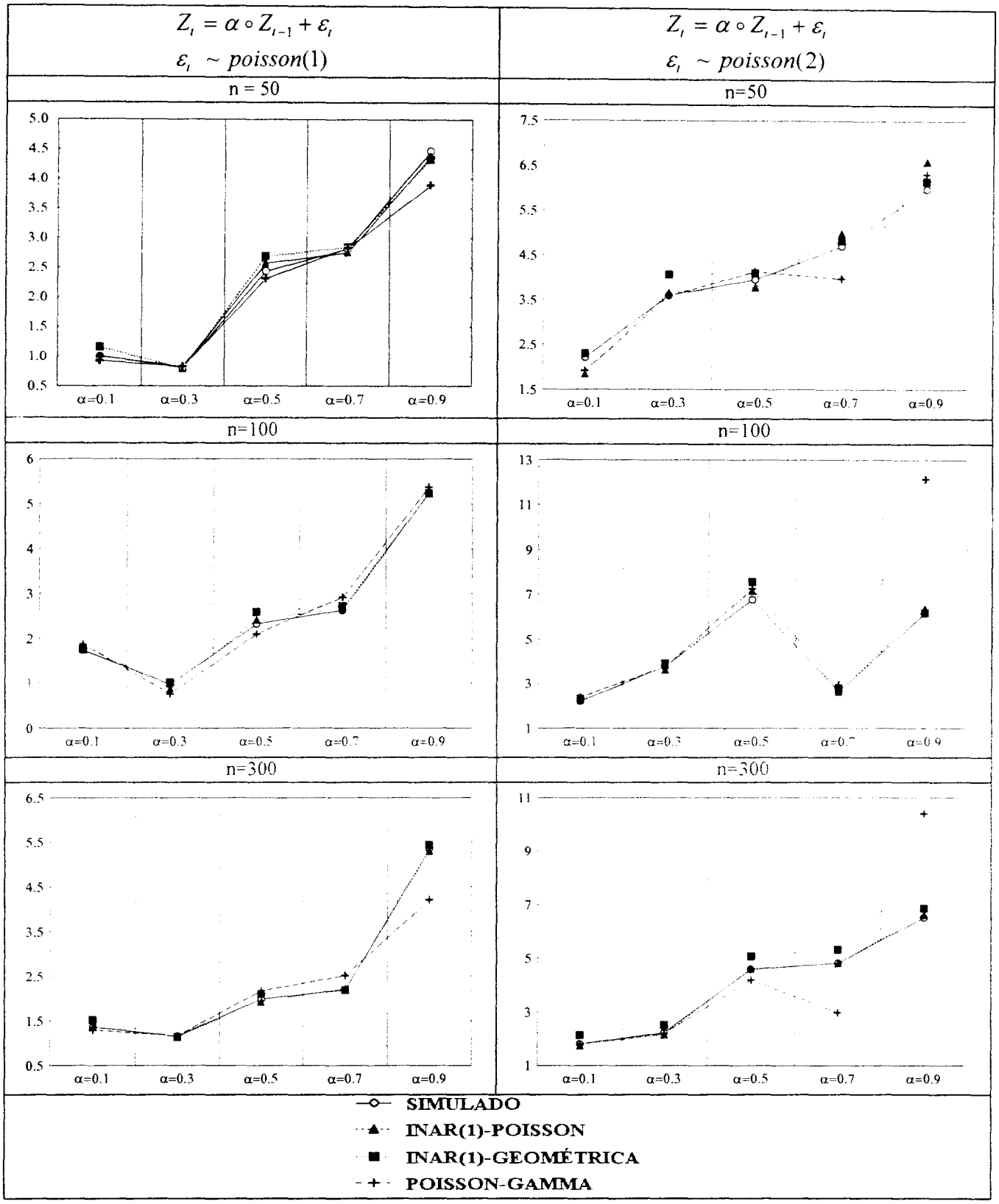
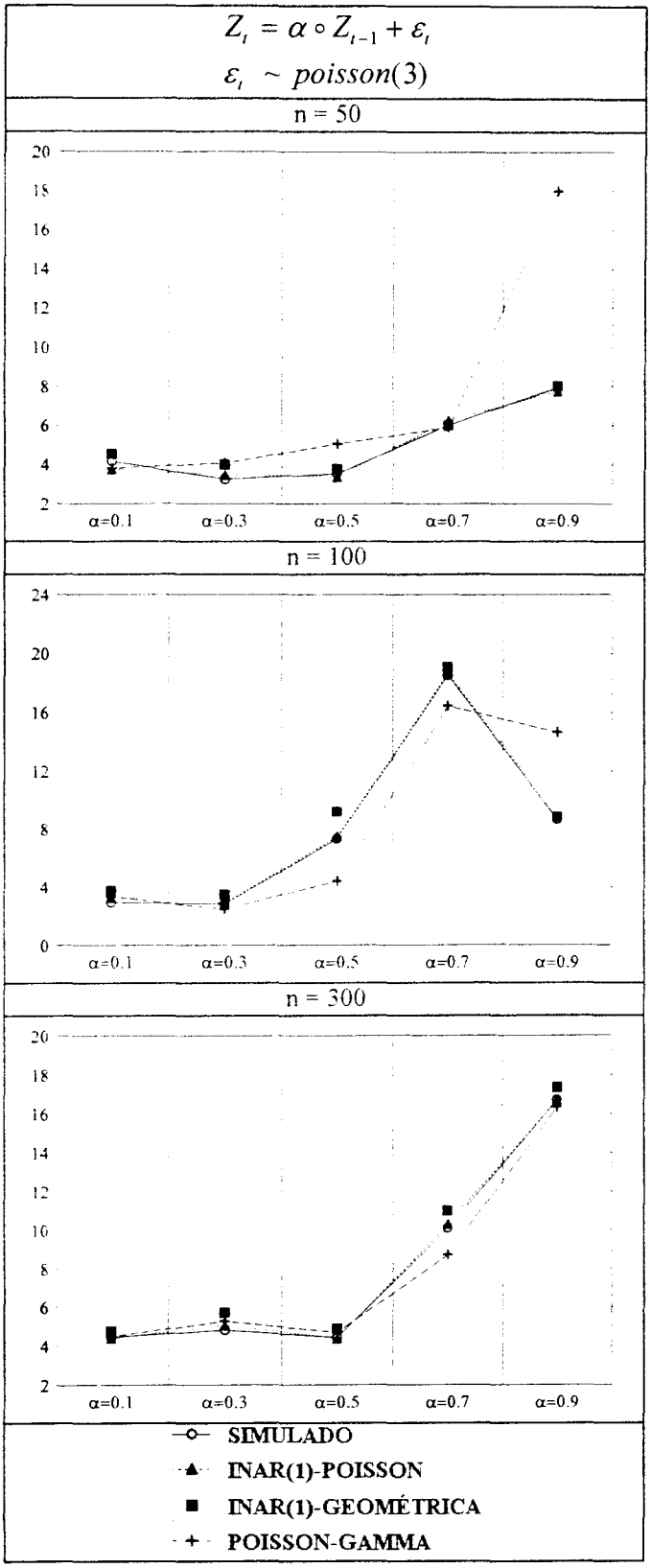


Figura 4.3.7 b- Comportamento do erro quadrático - média em 20 replicações - obtido ao utilizar como previsão pontual 1-passo-à-frente a média das distribuições preditivas 1-passo-à-frente dos modelos ajustados ao processo INAR(1)-Poisson simulado.



Ao analisarmos as diferenças entre os resultados obtidos para a previsão 1-passo-à-frente neste critério EQMP, podemos verificar que estas são pequenas, havendo alguns pontos que se distanciam da tendência geral aprsentada, mas que podem ter sido ocasionados pela variabilidade das estimativas dos parâmetros dos modelos. O processo INAR(1)-Geométrica, em boa parte dos pontos escolhidos para a simulação, é aquele que detém o maior EQMP, comparativamente com os demais: o aumento na locação dos processos INAR(1)-Poisson simulados faz com que os processos Poisson-Gama tenham um comportamento diferenciado em relação as estruturas INAR(1), principalmente para valores altos de α. Nas figuras podemos encontrar estas discrepâncias para valores de λ iguais a 2 e 3 , quando α igual a 0,9. Em sintonia com os resultados obtidos quando analisadas a média e variância da distribuição preditiva, os modelos INAR(1) voltaram a apresentar resultados muito próximos quando ajustados a processos com alta correlação serial.

4.3.3 - Afinidade entre as distribuições preditivas 1-passo-à-frente

Os resultados apresentados nas seções 4.3.2.1 e 4.3.2.2 mostraram que em termos da distribuição preditiva 1-passo-à-frente, as médias dos 3 modelos ajustados se equiparam enquanto a variância apresentada no modelo Poisson-Gama é consideravelmente maior quanto maior for a correlação serial do processo simulado. Para realizar uma comparação que não considere somente os dois primeiros momentos da distribuição mas que olhe de um modo geral a afinidade existente entre a distribuição preditiva real e a distribuição do modelo ajustado, utilizamos a medida de Kullback-Leibler.

As Figuras 4.3.8a,b mostram a afinidade da distribuição preditiva dos modelos ajustados com a distribuição preditiva real, porém isto é feito de forma mais geral do que a comparação entre as médias e variâncias, pois optamos em comparar somente as divergência médias, a cada 20 replicações, ao invés de utilizar todas as replicações.

$$\overline{D}(P{:}Q) = \frac{1}{20}\sum_{j=1}^{20} D_j(P{:}Q)$$

$D_j(P{:}Q)$: Divergência de Kullback-Leibler, na j-ésima replicação, entre a distribuição preditiva 1-passo-à-frente Q do modelo ajustado e a distribuição real P.

FIGURA 4.3.8a - Divergência de Kullback-Leibler - média em 20 replicações - entre as distribuições preditivas do modelo ajustado e modelo simulado - simulação do processo INAR(1)-Poisson.

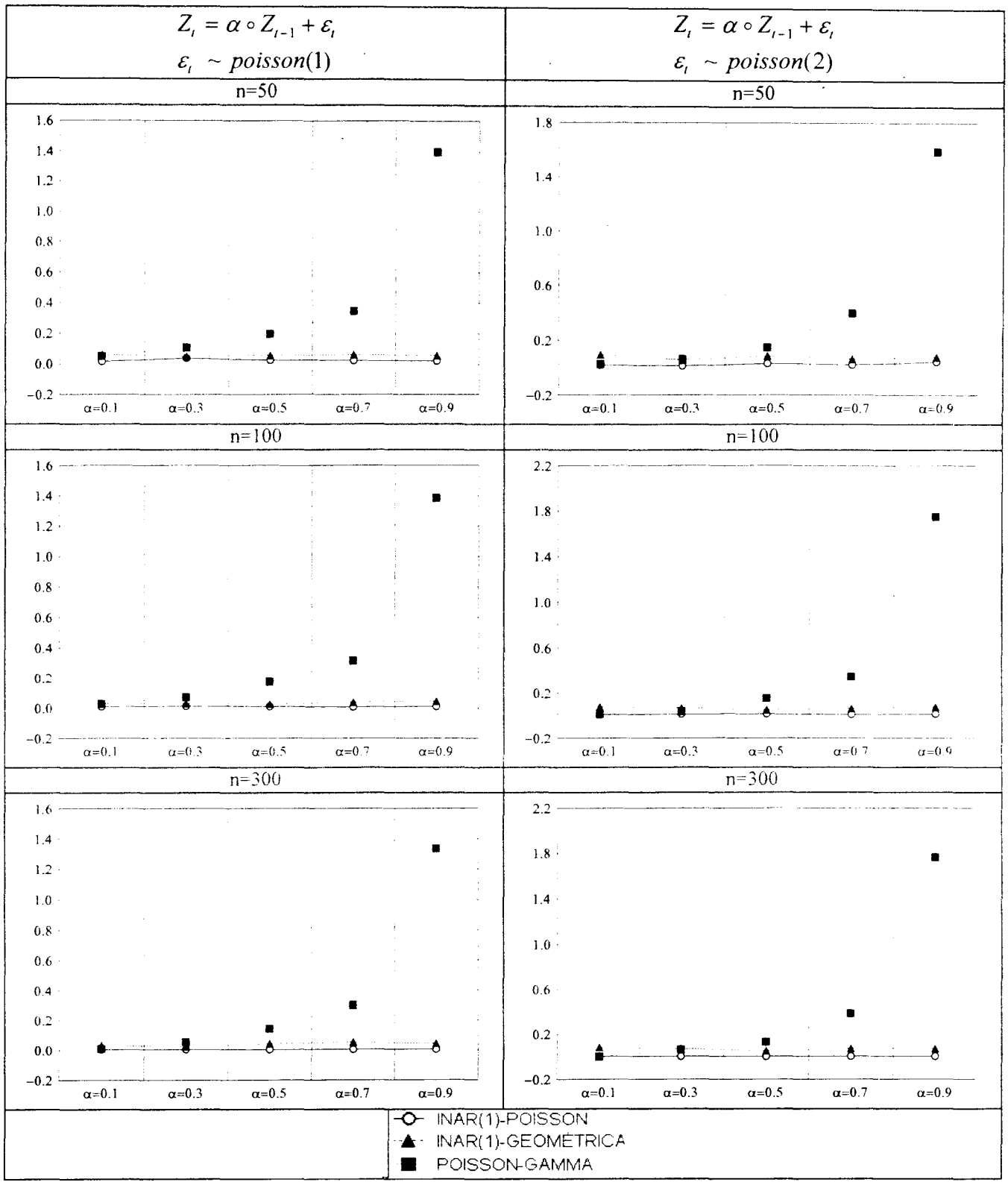
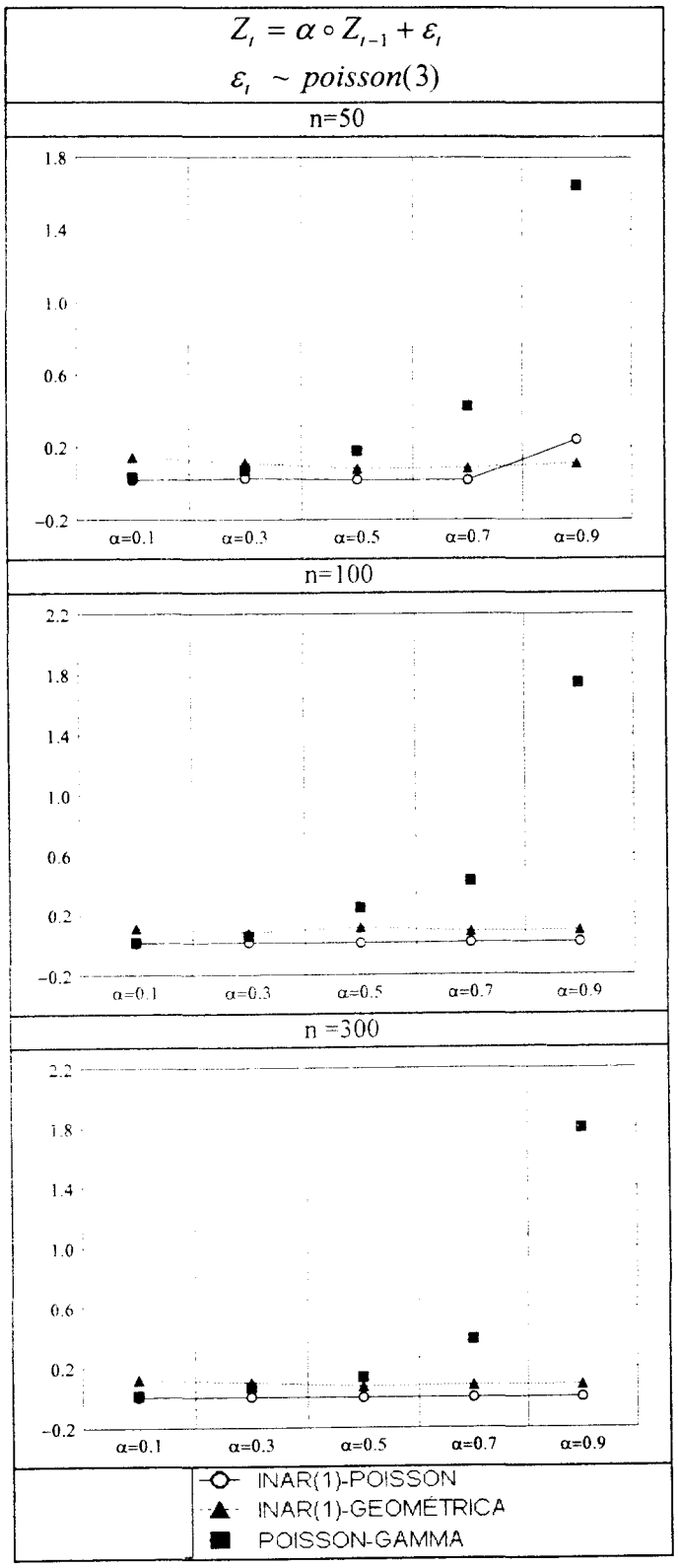


FIGURA 4.3.8b - Divergência de Kullback-Leibler - média em 20 replicações - entre as distribuições preditivas do modelo ajustado e modelo simulado - simulação do processo INAR(1)-Poisson.



As Figuras 4.3.8a,b caracterizam o seguinte comportamento para a divergência da distribuição preditiva 1-passo-à-frente dos 3 modelos ajustados em relação a distribuição preditiva real : o modelo Poisson-Gama é aquele que possui a menor afinidade com a distribuição preditiva real. Esta afinidade diminui a medida em que o valor de α . fixado para a simulação, aumenta. A comparação entre as estruturas INAR(1) mostra que para ambas inovações, Poisson e geométrica, a divergência de Kullback-Leibler é pequena, sendo esta levemente maior para o processo INAR(1)-Geométrica, principalmente em séries simuladas com pequena correlação serial. Podemos destacar nos gráficos acima que há poucos indícios de que divergência seja influenciada pela escolha para o nível das inovações, assim como pelo tamanho da série simulada.

4.4 – RESULTADOS OBTIDOS A PARTIR DA SIMULAÇÃO DO PROCESSO INAR(1)-GEOMÉTRICA.

4.4.1 - Ajuste dos modelos

Embora a estruturas de autocorrelação dos processos INAR(1)-Poisson e INAR(1)-Geométrica sejam as mesmas, uma marcante diferença entre os dois processos é a razão entre a variância e a média. Quando utilizamos as inovações de Poisson, esta razão é sempre igual a 1. Entretanto as inovações provenientes da distribuição geométrica propiciam um processo que apresenta sobredispersão cujo comportamento foi comentado na seção 4.3.1.

Processo INAR(1)-Geométrica : $Z_t = \phi \circ Z_{t-1} + \varepsilon_t$

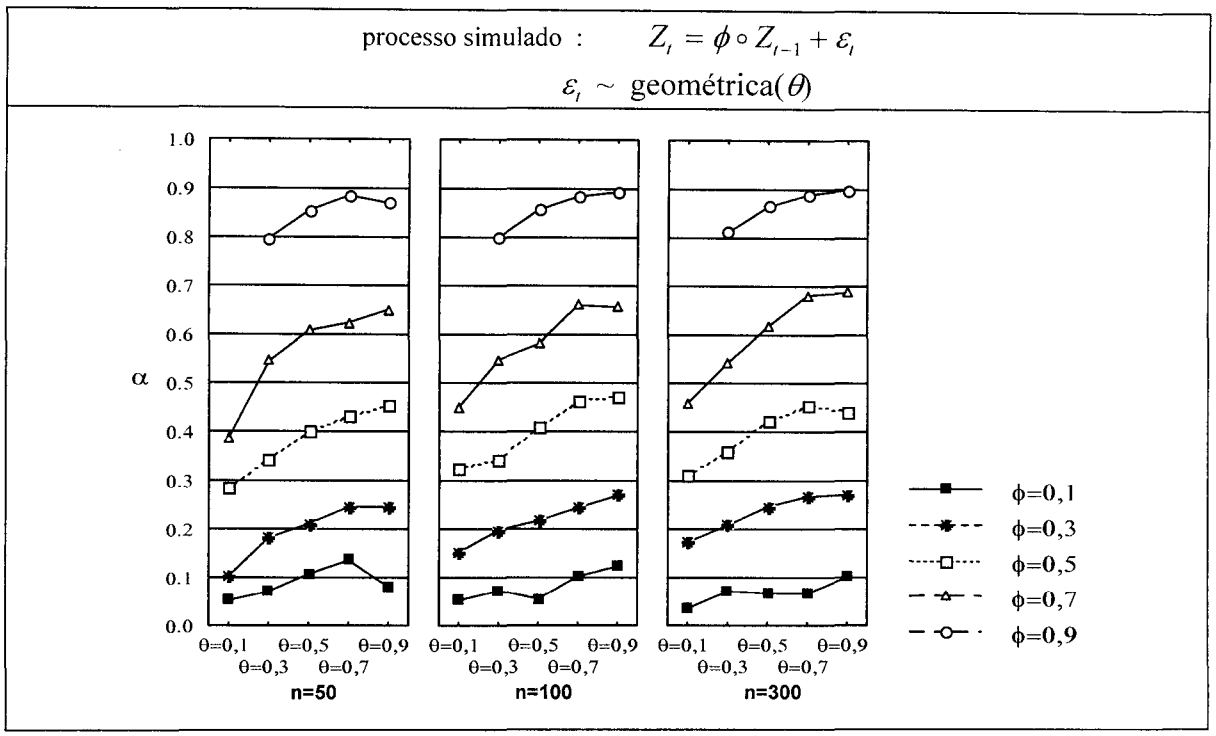
$\varepsilon_t \sim$ geométrica(θ)

Ao simularmos um processo INAR(1)-Geométrica, são válidas as mesmas considerações feitas na seção 4.3.1 no que diz respeito ao ajuste do modelo INAR(1)-Poisson caso utilizássemos o método dos momentos, igualando duas características do processo como a autocorrelação e média ou a autocorrelação e a variância. Entretanto, os resultados obtidos no ajuste do modelo INAR(1)-Geométrica na simulação do processo INAR(1)-Poisson mostraram que ao utilizar a estimação por máxima verossimilhança, obtivemos uma combinação de ϕ e θ que aproximou a média do processo gerado. Para que isto pudesse ser obtido, os valores estimados para ϕ estiveram bem acima da correlação serial que foi utilizada na simulação da série, com a exceção de valores de α =0,9. Invertendo-se a situação, iremos verificar qual o comportamento de α em função da correlação serial utilizada para simular o processo INAR(1)-Geométrica, só que desta feita, além da variação de ϕ, teremos que considerar a variação existente nas inovações da distribuição geométrica que abre um número maior de possibilidades no PGD se compararmos com os valores escolhidos para λ na simulação do processo INAR(1)-Poisson. Logo, os resultados agora estão mais sujeitos a efeitos causados pela interação entre ϕ e θ. Ao ajustarmos o modelo Poisson-Gama, as expectativas em relação as estimativas para o hiperparâmetro mudam em relação a

seção 4.3.1, pois esperamos que o valor escolhido para θ exerça algum tipo de influência já que em função deste parâmetro a sobredispersão dos valores no processo INAR(1)-Geométrica é maior do que no processo INAR(1)-Poisson.

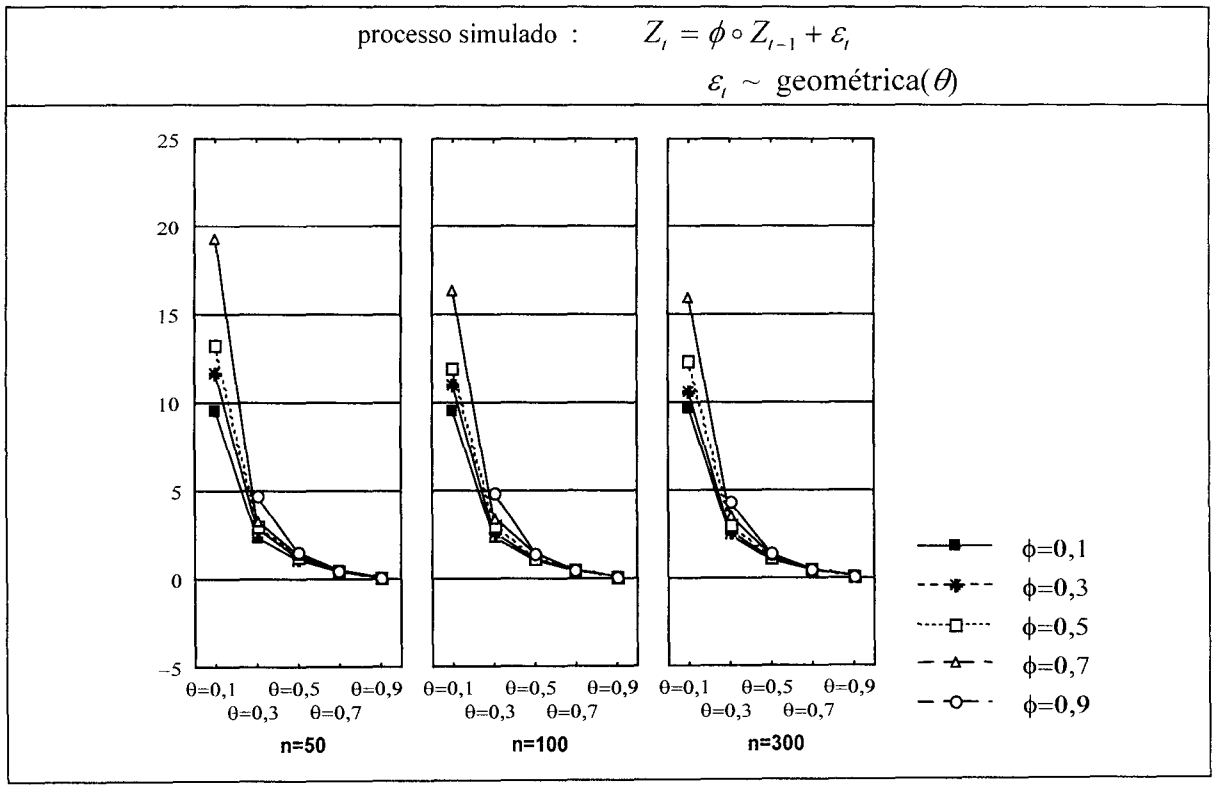
4.4.1.1 - Ajuste do modelo INAR(1)-Poisson ao processo INAR(1)-Geométrica simulado.

Figura 4.4.1 - Estimativas médias em 20 replicações obtidas para o parâmetro α do modelo INAR(1)-Poisson a partir da simulação do processo INAR(1)- Geométrica com parâmetros ϕ e θ.



A Figura 4.4.1 mostra que o parâmetro α do processo INAR(1)-Poisson, em sentido contrário da estimativa para ϕ na seção 4.3.1, indica uma correlação serial menor do que a utilizada para a simulação do processo. Entretanto, $\hat{\alpha}$ se aproxima de ϕ a medida em que os valores de ϕ e θ utilizados na simulação se aproximam de 0,9. Estes processos são aqueles simulados com alta correlação serial, porém com inovações próximas a zero e logicamente a variância nestes casos não se diferencia da média pois o processo assumirá poucos estados. Nestas situações, o procedimento de estimação por máxima verossimilhança aproxima simultaneamente a média e a autocorrelação do processo.. O tamanho da série escolhido para a simulação aparenta não exercer influência significativa sobre os resultados.

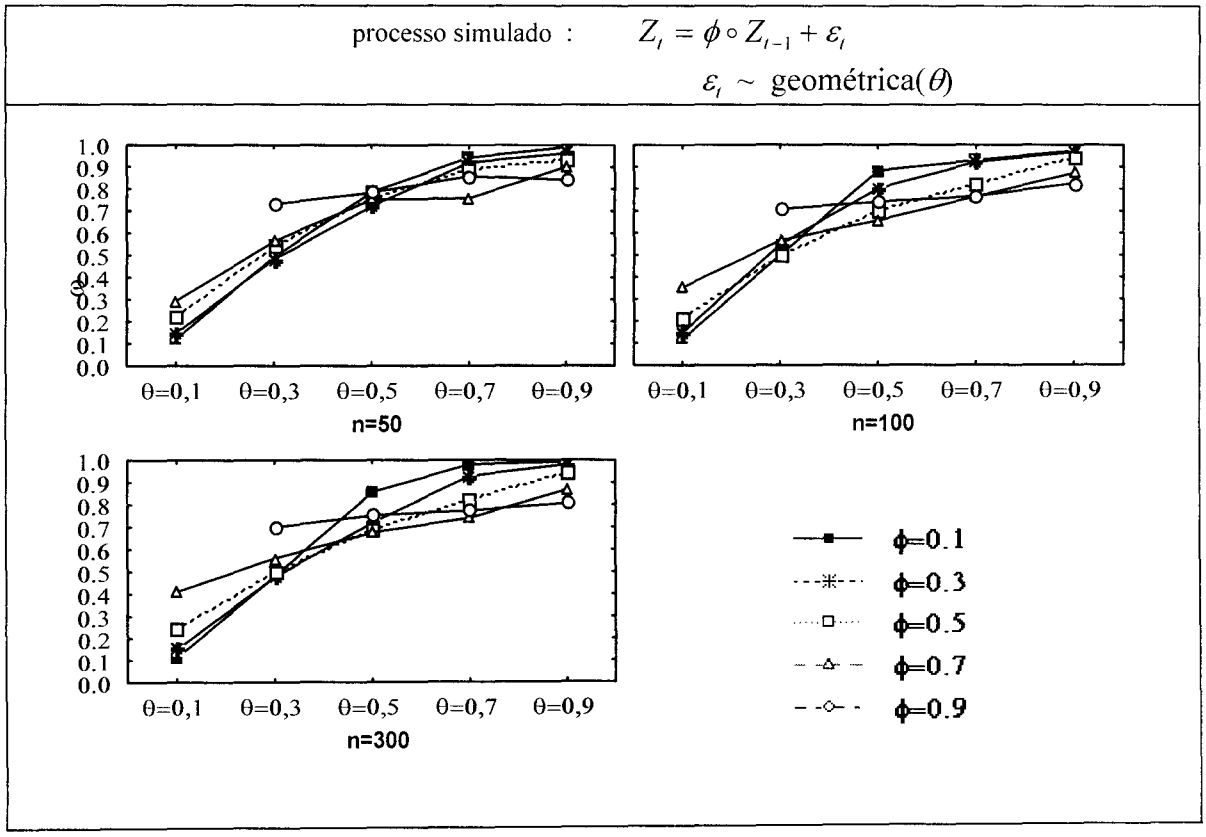
Figura 4.4.2 - Estimativas médias em 20 replicações obtidas para o parâmetro α do modelo INAR(1)-Poisson a partir da simulação do processo INAR(1)- Geométrica com parâmetros ϕ e θ



Na tentativa de aproximar a média, o procedimento de máxima verossimilhança encontrou estimativas para a média que foram influenciadas somente pela locação das inovações, representada pelo parâmetro θ. Entretanto, em séries que apresentaram associadamente alta correlação serial e grande magnitude, ou seja, valores altos de ϕ e pequenos de θ, podemos notar que a estimativa para a média das inovações de Poisson se comporta de forma crescente com os valores de ϕ.

4.4.1.2 -Ajuste do modelo Poisson-Gama ao processo INAR(1)-Geométrica simulado

Figura 4.4.3 - Estimativas médias em 20 replicações obtidas para o hiperparâmetro ω do modelo Poisson-Gama a partir da simulação do processo INAR(1)-Geométrica com parâmetros ϕ e θ.



A Figura 4.4.3 exibe o comportamento do hiperparâmetro do processo Poisson-Gama quando este é estimado baseado em observações provenientes do processo INAR(1)-Geométrica. A figura indica que os valores estimados de ω aumentam de acordo com o aumento na locação do processo simulado representada por θ. Em séries de pequena magnitude, como por exemplo , aquelas que utilizam θ=0,9, o valor estimado para o hiperparâmetro está próximo de 1, caracterizando a pequena variação existente nos valores destes processos. A medida em que a magnitude dos valores aumenta, o modelo Poisson-Gama retrata o aumento na variação dos valores ao apresentar estimativas baixas de $\hat{\omega}$ que caracterizam uma grande volatilidade no nível do processo . Um padrão de comportamento não fica tão bem definido quando observamos a variação de $\hat{\omega}$ em função de ϕ.

4.4.2 - Estatísticas principais da distribuição preditiva 1-passo-à-frente

4.4.2.1 - Média da distribuição preditiva 1-passo-à-frente

Repetiremos nesta seção o procedimento utilizado na seção 4.3.2.1 para verificarmos o quanto a média da distribuição preditiva de cada modelo se distancia da média do modelo real, que nesta situação é o modelo INAR(1)-Geométrica. Entretanto, como o número de simulações deste processo foi bem superior aos demais processos, apresentamos os resultados em três níveis: séries simuladas com alta magnitude, ou seja, com valores de θ=0,1, séries com média magnitude, θ=0,5, e séries de pequena magnitude com excesso de zeros que são aquelas simuladas com θ= 0,9. Em cada uma destas situações utilizamos os seguintes valores para ϕ: 0,1,0,5 e 0,9[1]. Os resultados para as outras combinações de parâmetros estão no apêndice B que contem todos os resultados obtidos para a simulação do processo INAR(1)-Geométrica. As figuras 4.4.4a,b,c mostram os resultados considerando séries pequenas (n=50) e séries grandes (n=300). Para auxiliar a interpretação dos resultados, apresentamos a tabela 4.4.1 contendo o erro relativo médio das médias dos modelos ajustados em relação a média real.

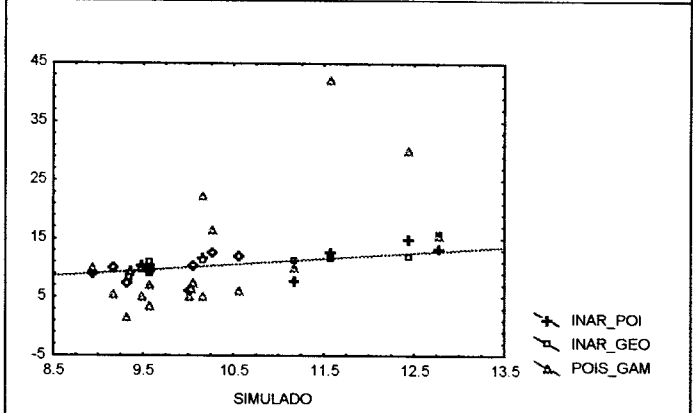
[1] Como a combinação ϕ =0,9 e θ =0,1 foi retirada do trabalho, utilizamos nesta situação ϕ =0,7

Figura 4.4.4a - Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo INAR(1)-Geométrica com θ =0,1

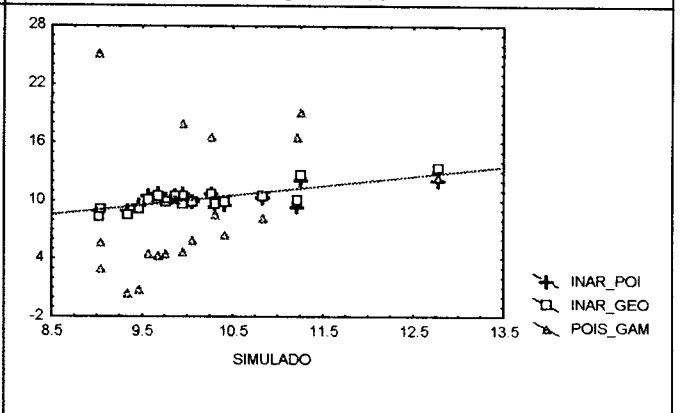
processo simulado $Z_t = 0{,}1Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ Geometrica(0,1)

SÉRIES DE TAMANHO n = 50



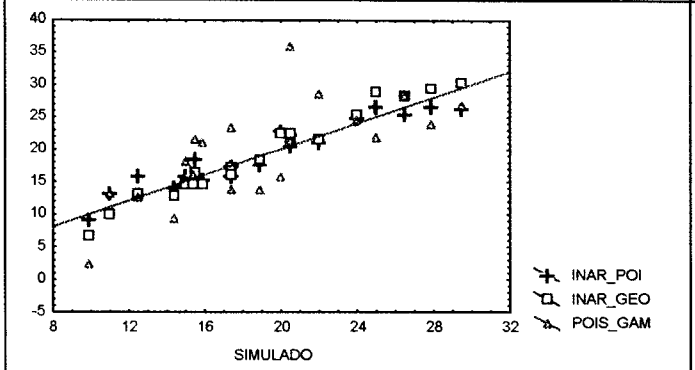
SÉRIES DE TAMANHO n = 300



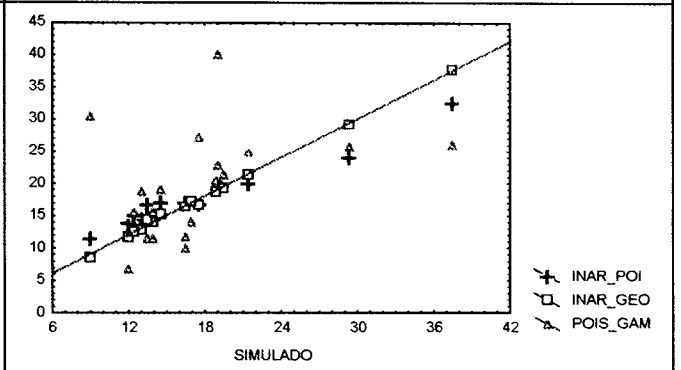
processo simulado $Z_t = 0{,}5Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ Geometrica(0,1)

SÉRIES DE TAMANHO n = 50



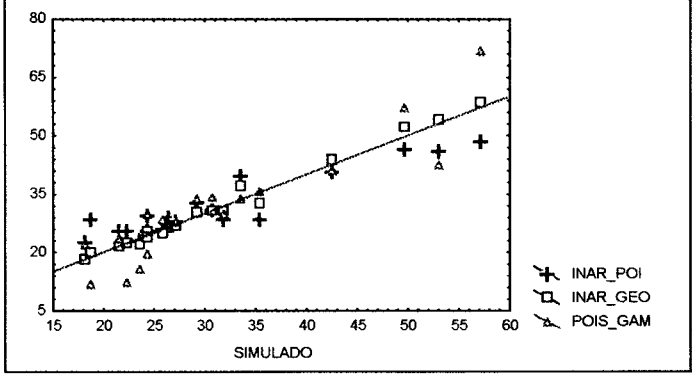
SÉRIES DE TAMANHO n = 300



processo simulado $Z_t = 0{,}7Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ Geometrica (0,1)

SÉRIES DE TAMANHO n = 50



SÉRIES DE TAMANHO n = 300

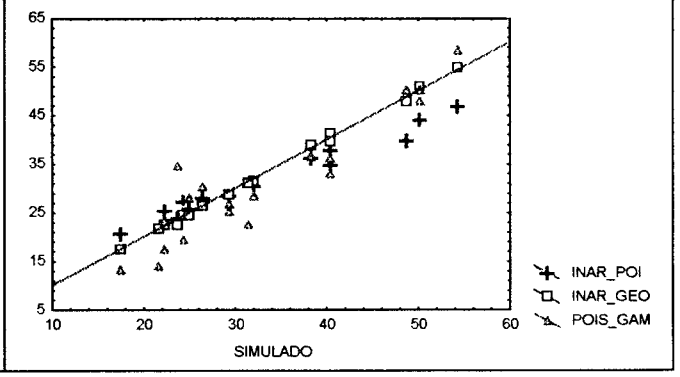


Figura 4.4.4b- Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo INAR(1)-Geométrica com θ =0,5.

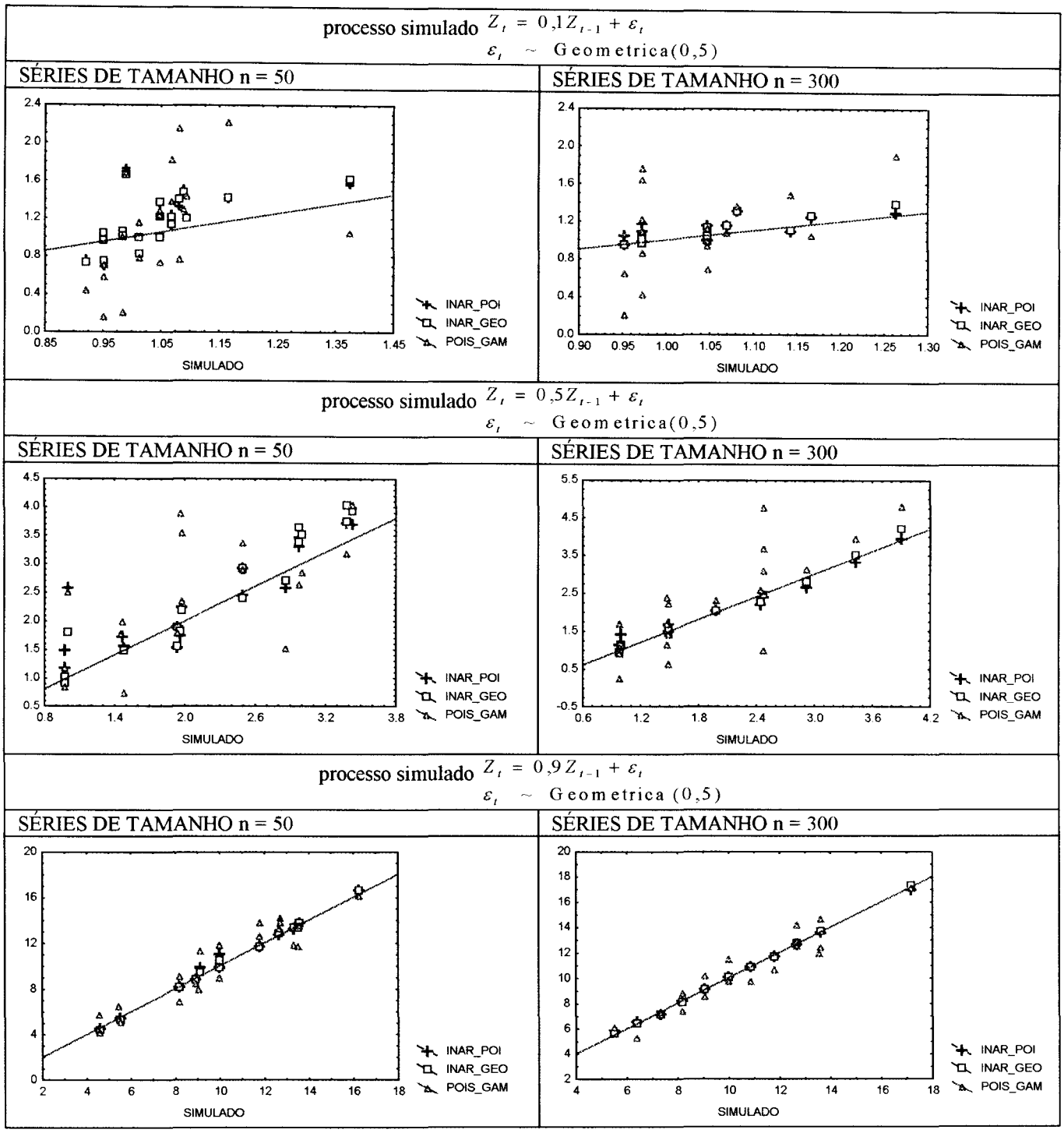
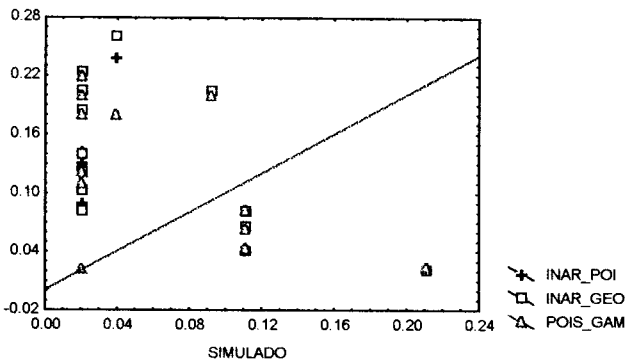


Figura 4.4.4c - Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo INAR(1)-Geométrica com θ =0,9.

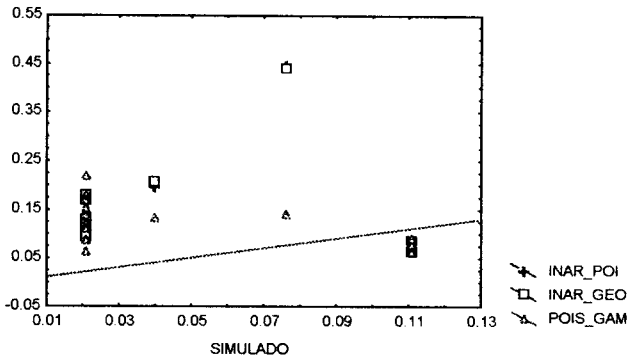
processo simulado $Z_t = 0{,}1Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ Geometrica(0,9)

SÉRIES DE TAMANHO n = 50



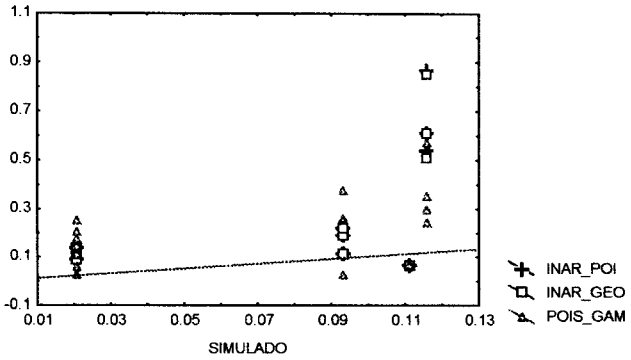
SÉRIES DE TAMANHO n = 300



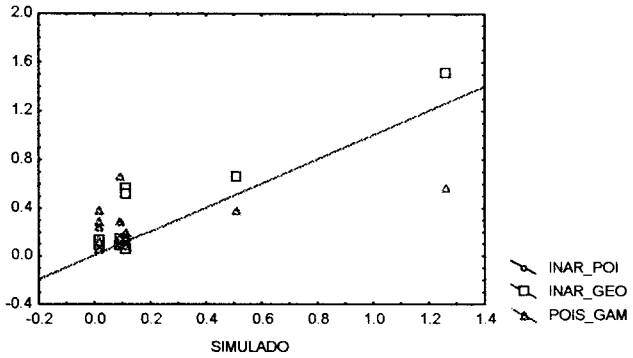
$Z_t = 0{,}5Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ Geometrica(0,9)

SÉRIES DE TAMANHO n = 50



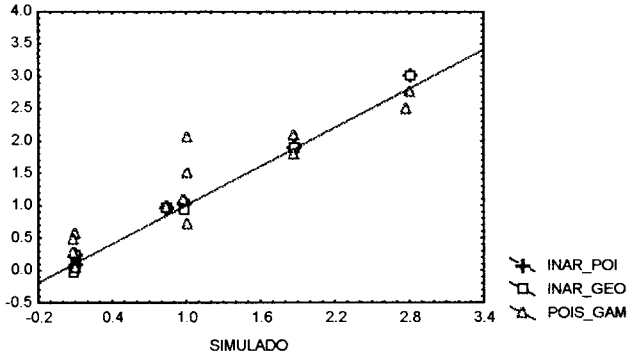
SÉRIES DE TAMANHO n = 300



$Z_t = 0{,}9Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ Geometrica (0,9)

SÉRIES DE TAMANHO n = 50



SÉRIES DE TAMANHO n = 300

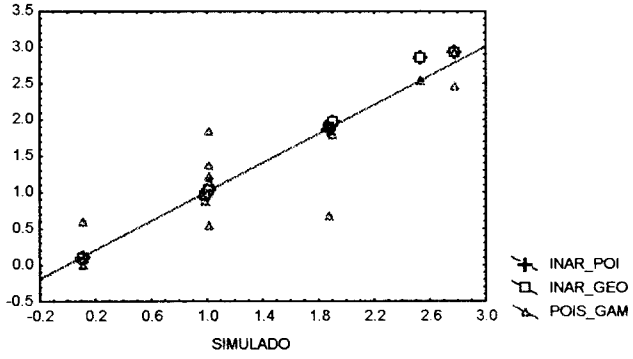


Tabela 4.4.1 - Erro relativo médio da média da distribuição preditiva 1-passo-à-frente dos modelos ajustados em relação à média da distribuição preditiva correta.

ϕ	θ	n	INAR(1)-Poisson	INAR(1)-Geométrica	Poisson-Gama
0,1	0,1	50	0,115	0,099	0,623
	0,1	100	0,090	0,066	0,673
	0,1	300	0,062	0,054	0,564
	0,5	50	0,179	0,181	0,425
	0,5	100	0,132	0,145	0,328
	0,5	300	0,067	0,060	0,299
	0,9	50	2,461	2,946	2,619
	0,9	100	2,448	2,643	2,594
	0,9	300	3,697	3,836	4,020
0,5	0,1	50	0,080	0,076	0,244
	0,1	100	0,071	0,032	0,339
	0,1	300	0,101	0,021	0,399
	0,5	50	0,210	0,132	0,329
	0,5	100	0,106	0,065	0,277
	0,5	300	0,067	0,052	0,364
	0,9	50	3,404	3,449	3,943
	0,9	100	2,365	2,361	3,596
	0,9	300	2,899	2,884	3,644
0,9	0,5	50	0,022	0,021	0,118
	0,5	100	0,019	0,014	0,116
	0,5	300	0,011	0,007	0,076
	0,9	50	0,200	0,253	0,946
	0,9	100	0,138	0,134	0,918
	0,9	300	0,064	0,065	0,707

Figura 4.4.4a mostra que em simulações séries de alta magnitude (θ=0,1), a média da distribuição preditiva do modelo Poisson-Gama é a que apresenta a maior variabilidade, principalmente para séries simuladas com baixa correlação serial. Nestas séries, as médias das distribuições preditivas 1-passo-à-frente dos modelos INAR(1)-Poisson e INAR(1)-Geométrica tem comportamento similar. A medida em que aumentamos a correlação serial, os resultados obtidos pelos 3 modelos ficaram próximos entre si e próximos da média da distribuição preditiva verdadeira. Utilizando um valor intermediário para a locação das inovações (θ=0,5), podemos observar na Figura 4.4.4b que os desvios em relação ao modelo correto aumentaram para as estruturas INAR(1) e continuaram a exibir o mesmo padrão para o modelo Poisson-Gama. Entretanto para ϕ igual a 0,9 , os resultados voltaram a convergir para aquele que seria o correto, porém com desvios levemente maiores do modelo Poisson-Gama. A Figura 4.4.4c traz uma situação muito particular que envolve processos simulados com magnitude muito pequena (θ = 0,9) e excesso de zeros, principalmente para ϕ =0,1. Nestes casos, todos os modelos, inclusive o real, apresentaram grandes desvios em relação a média verdadeira. Entretanto, devemos ter o cuidado de observar que muitos destes processos apresentaram variação muito pequena e consequentemente, qualquer desvio em relação a média verdadeira da distribuição preditiva fica sobressaltado.

4.4.2.2 - Variância da distribuição preditiva 1-passo-à-frente

Nesta sub-seção, trataremos de uma situação mais complexa do que a anterior que envolvia o modelo INAR(1)-Poisson. Conforme visto na Figura 4.3.1, o processo INAR(1)-Geométrica apresenta sobredispersão em todas as combinações de parâmetros utilizados para a simulação. O primeiro problema que surge é o ajuste do modelo INAR(1)-Poisson que pressupõe igualdade entre média e variância. O procedimento de estimação por máxima verossimilhança ao ajustar o modelo INAR(1)-Poisson, distorceu o valor da correlação serial utilizada na simulação e aproximou a média do PGD , conseqüentemente considerando esta igual a variância. Entretanto, esta característica distribucional do modelo INAR(1)-Poisson não é compatível com a sobredispersão do processo simulado. A expectativa em relação a variância da distribuição preditiva 1-passo-à-frente do modelo Poisson-Gama é que nas situações simuladas do processo INAR(1)-Geométrica nas quais o hiperparâmetro estimado foi pequeno (ver seção 4.4.1), a variância da distribuição preditiva deve ser grande em face a grande variância da distribuição que caracteriza o nível do processo na transição do instante n para o instante n+1. As Figuras 4.4.5a,b e c mostram a dispersão das variâncias dos modelos ajustados em relação a variância verdadeira.

Figura 4.4.5a - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Geométrica com $\theta = 0,1$.

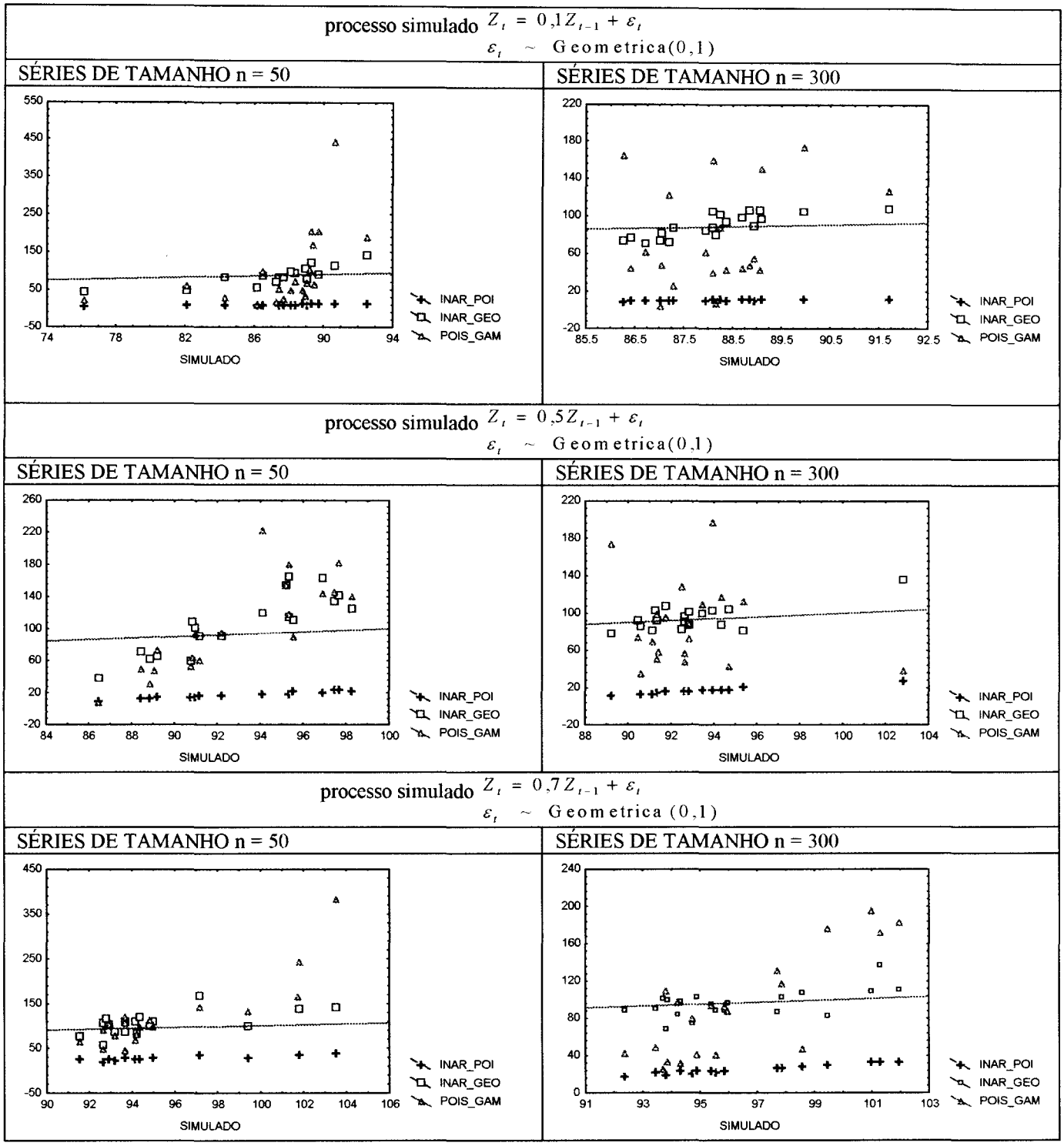
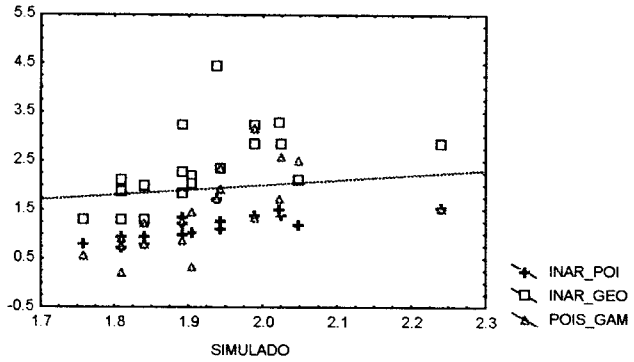


Figura 4.4.5b - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Geométrica com $\theta = 0,.5$.

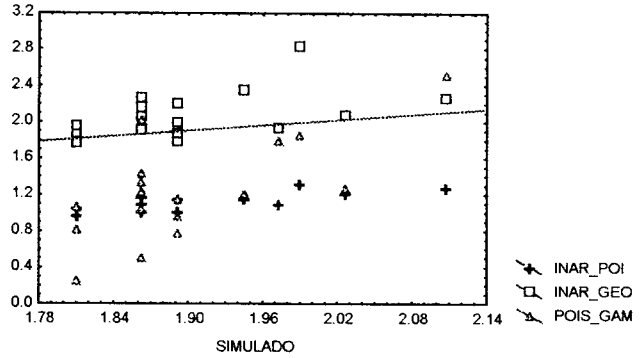
processo simulado $Z_t = 0,1Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ geométrica(0,5)

SÉRIES DE TAMANHO n = 50



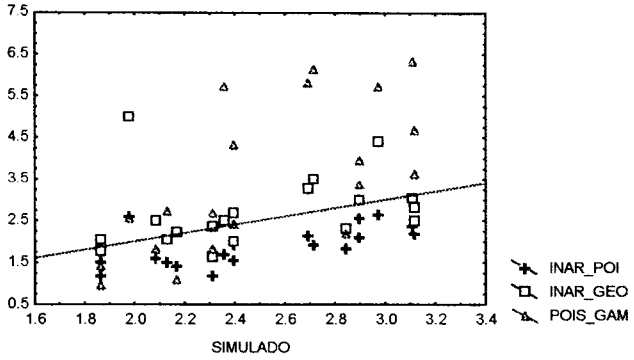
SÉRIES DE TAMANHO n = 300



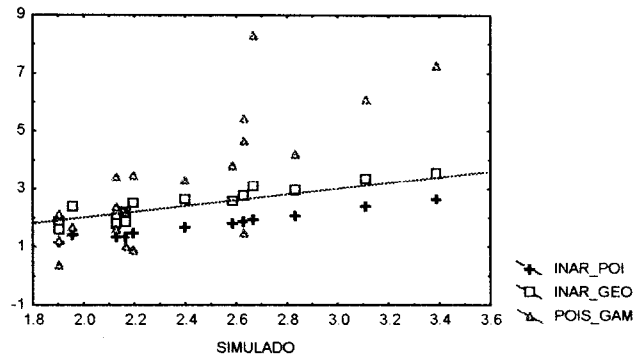
processo simulado $Z_t = 0,5Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ geométrica(0,5)

SÉRIES DE TAMANHO n = 50



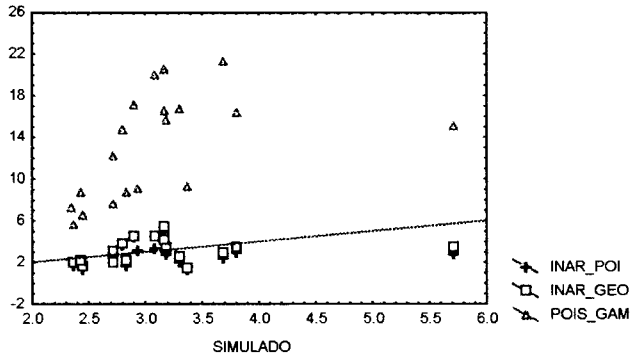
SÉRIES DE TAMANHO n = 300



processo simulado $Z_t = 0,9Z_{t-1} + \varepsilon_t$

$\varepsilon_t \sim$ geométrica(0,5)

SÉRIES DE TAMANHO n = 50



SÉRIES DE TAMANHO n = 300

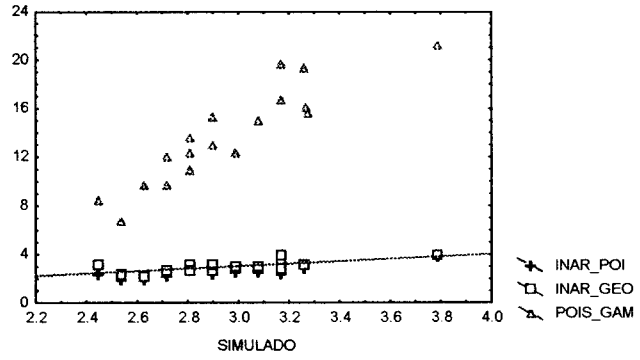
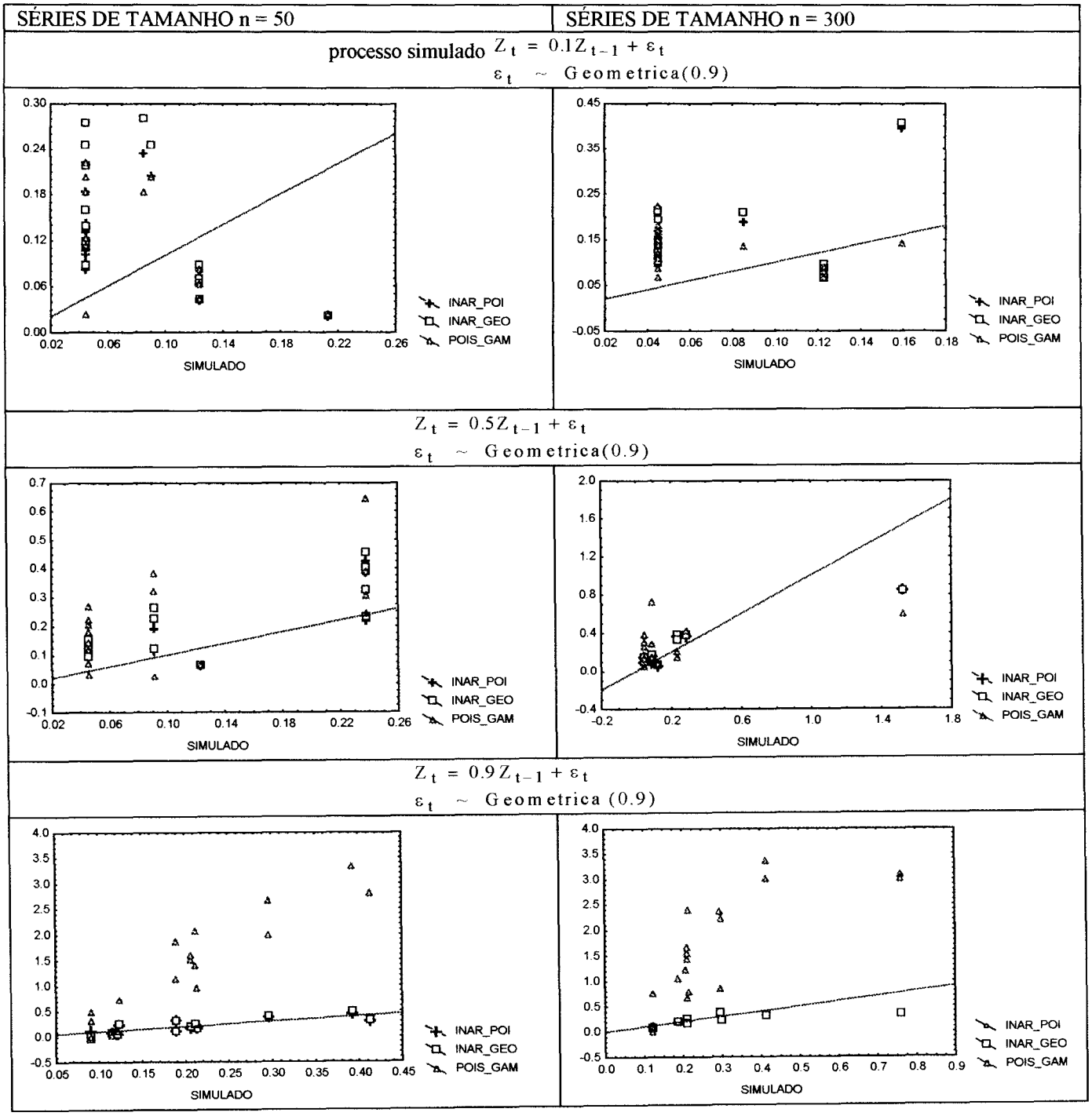


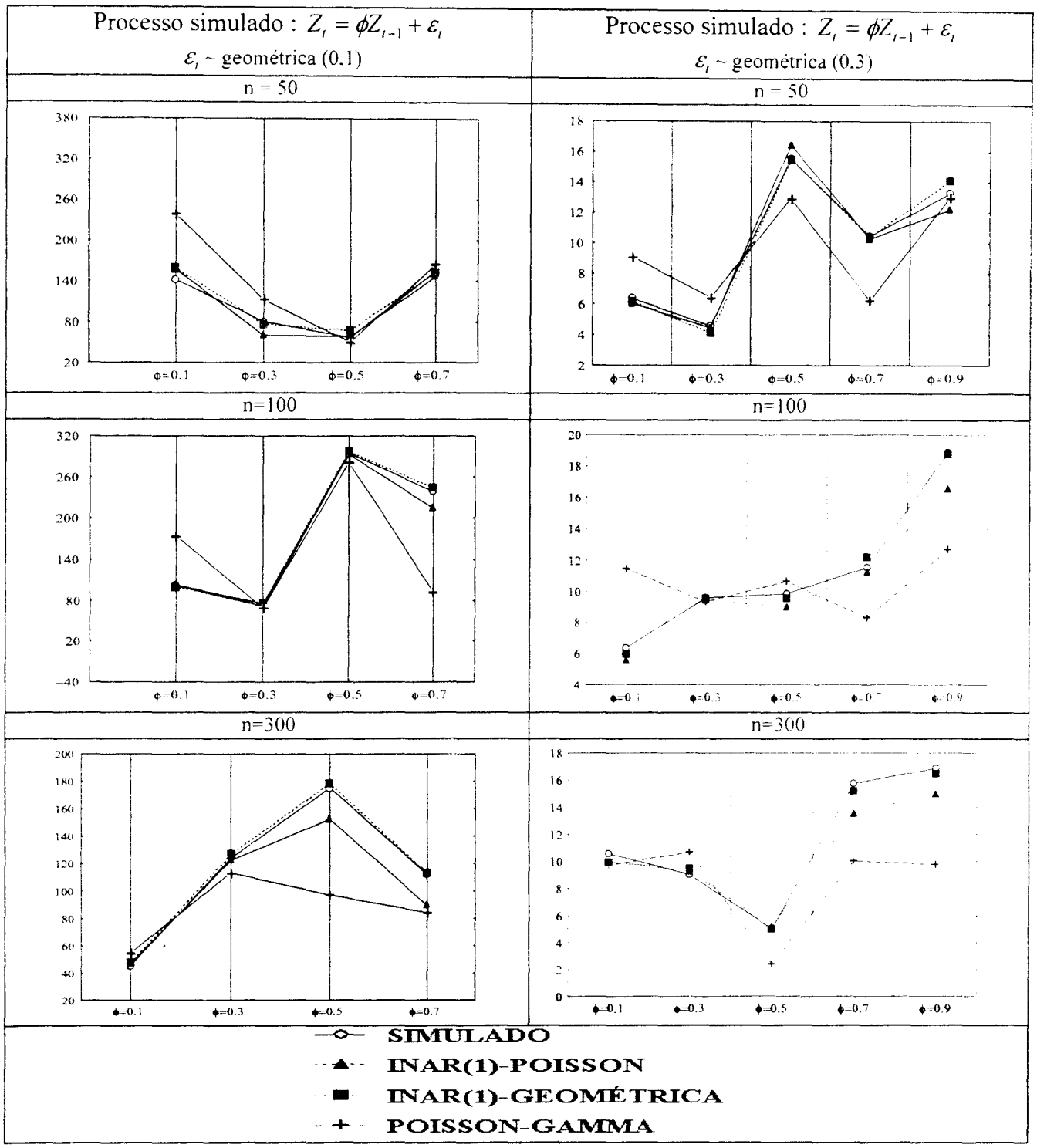
Figura 4.4.5c - Diagrama de dispersão entre as variâncias das distribuições preditivas 1-passo-à-frente dos 3 modelos ajustados e a variância verdadeira na simulação do processo INAR(1)-Geométrica com $\theta = 0,9$.



Na figura 4.4.5a que corresponde a simulações de processos com alta magnitude (θ=0,1), com grande a sobredispersão do processo INAR(1)-Geométrica, verificamos que o modelo INAR(1)-Poisson subestima os valores verdadeiros da variância. Este modelo ao ajustar a média do PGD que neste caso é o processo INAR(1)-Geométrica, ajustou também a variância pela característica da distribuição poisson. Logo, nos pontos em que a sobredispersão é grande, este modelo atua inadequadamente ao medir a incerteza do valor predito 1-passo-à-frente. O modelo Poisson-Gama não foi tendencioso como o modelo INAR(1)-Poisson, porém suas estimativas estiveram distantes do valor verdadeiro da variância, principalmente para séries de tamanho n= 300. Nestas séries, mesmo o modelo real apresentou grandes desvios em relação a variância verdadeira, principalmente para valores intermediários de ϕ. A Figura 4.4.5b mostra situações com um valor intermediário de θ . Nestas situações, o modelo INAR(1)-Poisson continua a subestimar a variância verdadeira , mas a subestimação decai com o aumento da correlação serial utilizada na simulação da série. O modelo Poisson-Gama nas situações em que ϕ=0,1 apresenta resultados que se aproximam bastante do modelo real , e para amostras pequenas, podemos considerá-los até melhores, mas a medida em que aumentamos a correlação serial, estes começaram a apresentar os maiores desvios em relação a variância verdadeira e uma clara tendência em superestimá-la. Os resultados para séries de pequena magnitude , θ=0,9, apresentam os mesmos problemas encontrados na seção anterior. Com correlações seriais pequenas e intermediárias (ϕ=0,1 e ϕ=0,5), as simulações são constituídas de zeros em sua quase totalidade e a variância fica próxima de zero e todos os modelos se equivalem nestas situações. Entretanto para correlação serial alta (ϕ=0,,9), os modelos real e INAR(1)-Poisson estimam a variância muito próximo do valor verdadeiro e fica clara a superestimação por parte do modelo Poisson-Gama.

4.4.2.3 - Erro Quadrático Médio da Previsão 1-passo-à-frente.

Figura 4.4.6a - Erro Quadrático de Previsão 1-passo-à-frente - média para os modelos ajustados em 20 replicações do processo INAR(1)-Geométrica simulado (ϕ = 0,1 e ϕ = 0,3).



Os resultados da Figura 4.4.6a foram calculados com base em séries simuladas com valores de alta magnitude ($\theta = 0{,}1$ e $\theta = 0.3$). Podemos notar que o comportamento do processo Poisson-Gama varia de acordo com a estrutura de autocorrelação da série simulada. Nas situações em que a autocorrelação é pequena, ou seja, valores de ϕ iguais a 0.1 e 0.3, o erro quadrático do processo Poisson-Gama é maior do que o calculado para as estruturas INAR(1), sendo que este quadro é invertido quando a autocorrelação da série simulada é alta. Estas situações correspondem as séries de maiores magnitudes.

A comparação entre as estruturas INAR(1) mostra um grande equilíbrio com resultados um pouco melhores para o erro quadrático médio do modelo INAR(1)-Poisson, principalmente para séries de tamanho n=300. Entretanto, ao exibirmos apenas a média desta estatística em 20 replicações, perdemos a informação de sua variabilidade que pode ser encontrada no apêndice B que contem os resultados obtidos para as simulações do processo INAR(1)-Geométrica.

Figura 4.4.6b - Erro Quadrático de Previsão 1-passo-à-frente - média para os modelos ajustados em 20 replicações do processo INAR(1)-Geométrica(ϕ = 0,5 e ϕ = 0,7).

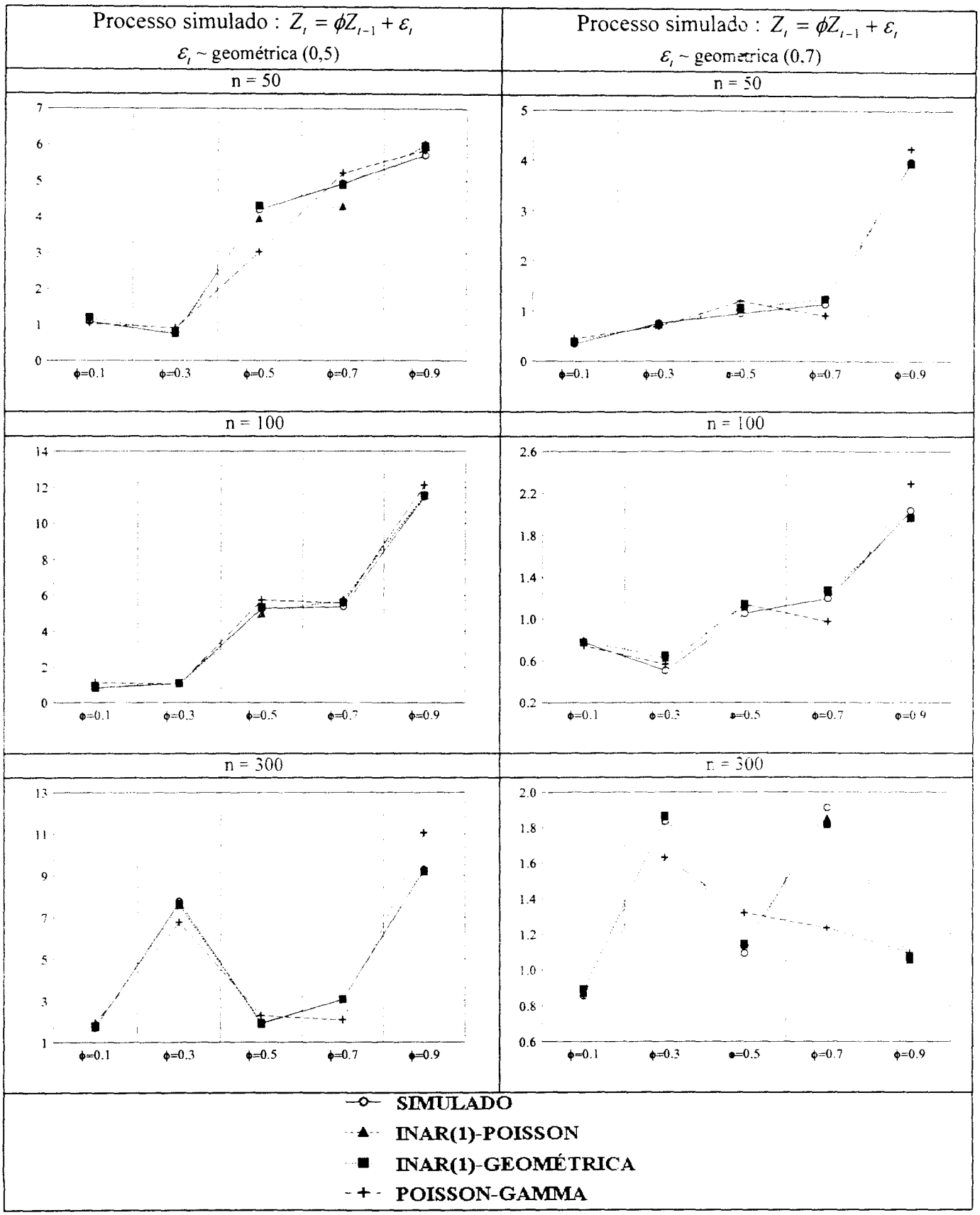
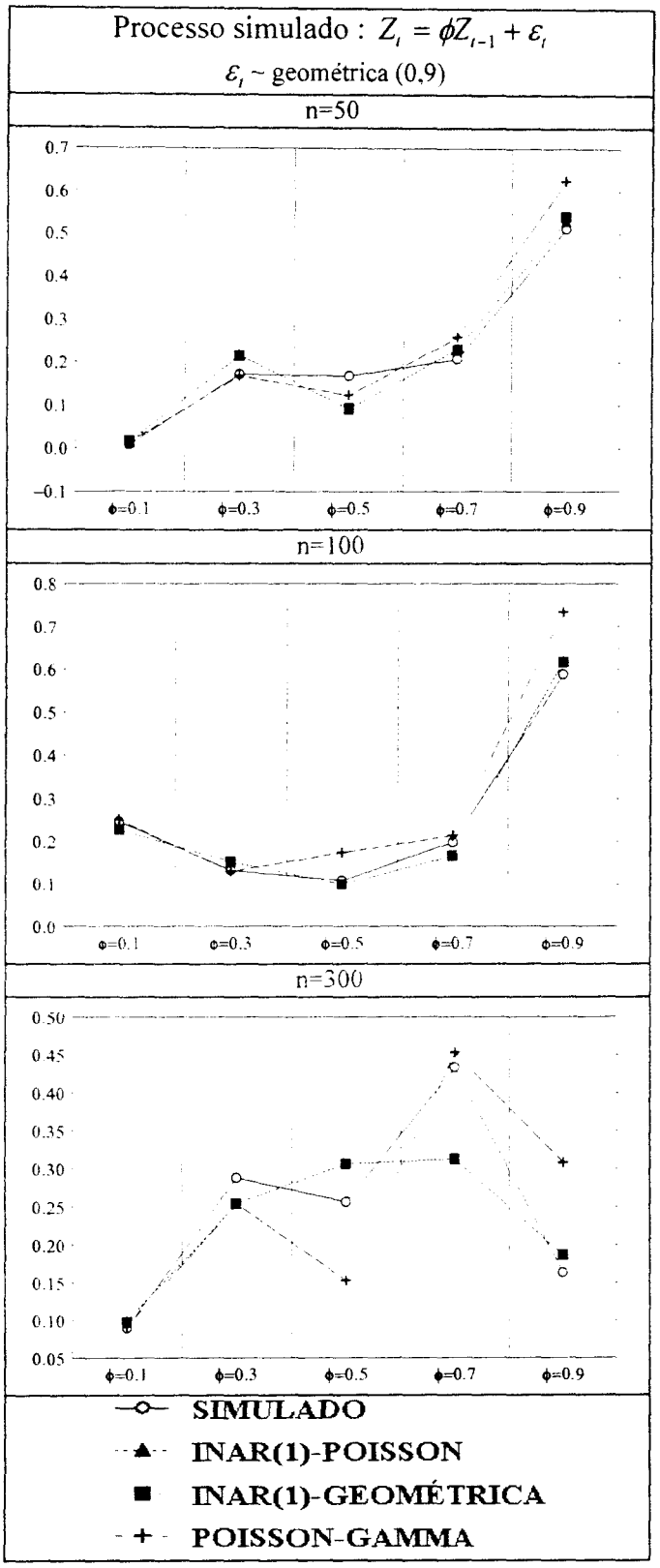


Figura 4.4.6c - Erro Quadrático de Previsão 1-passo-à-frente - média para os modelos ajustados em 20 replicações do processo INAR(1)-Geométrica.



O aumento na autocorrelação de primeira ordem das séries simuladas, torna os resultados apresentados pelos 3 modelos , mais semelhantes. De uma forma geral, não fica evidenciada, em nenhuma situação, uma clara superioridade de alguns dos 3 modelos. Se recordarmos os resultados encontrados na seção 4.4.2, iremos verificar que em termos da média da distribuição preditiva, os desvios apresentados pelos 3 modelos em relação a média verdadeira estiveram, aparentemente, aleatoriamente espalhados em torno do valor verdadeiro, não havendo nenhuma tendência de subestimação ou superestimação. A Figura 4.4.6c apresenta situações nas quais os modelos se comportaram de forma bastante semelhante , nesta estatística; em séries que apresentaram variância muito pequena e vários valores iguais a zero.

4.4.3 - Afinidade entre as distribuições preditivas.

A seguir, apresentamos os resultados para a medida de afinidade de Kullback-Leibler entre as distribuições preditivas dos modelos ajustados e a distribuição preditiva verdadeira. Considerando os resultados obtidos nas sub-seção 4.4.2, podemos esperar que sejam encontradas divergências altas entre as distribuições preditivas dos modelos ajustados e a do modelo correto, principalmente para o ajuste do modelo Poisson-Gama, pois este apresentou grandes desvios em relação ao valor verdadeiro, tanto na média como na variância da distribuição preditiva 1-passo-à-frente.

FIGURA 4.4.7a - Divergência de Kullback-Leibler entre a distribuição preditiva real e as distribuições preditivas dos modelos ajustados.

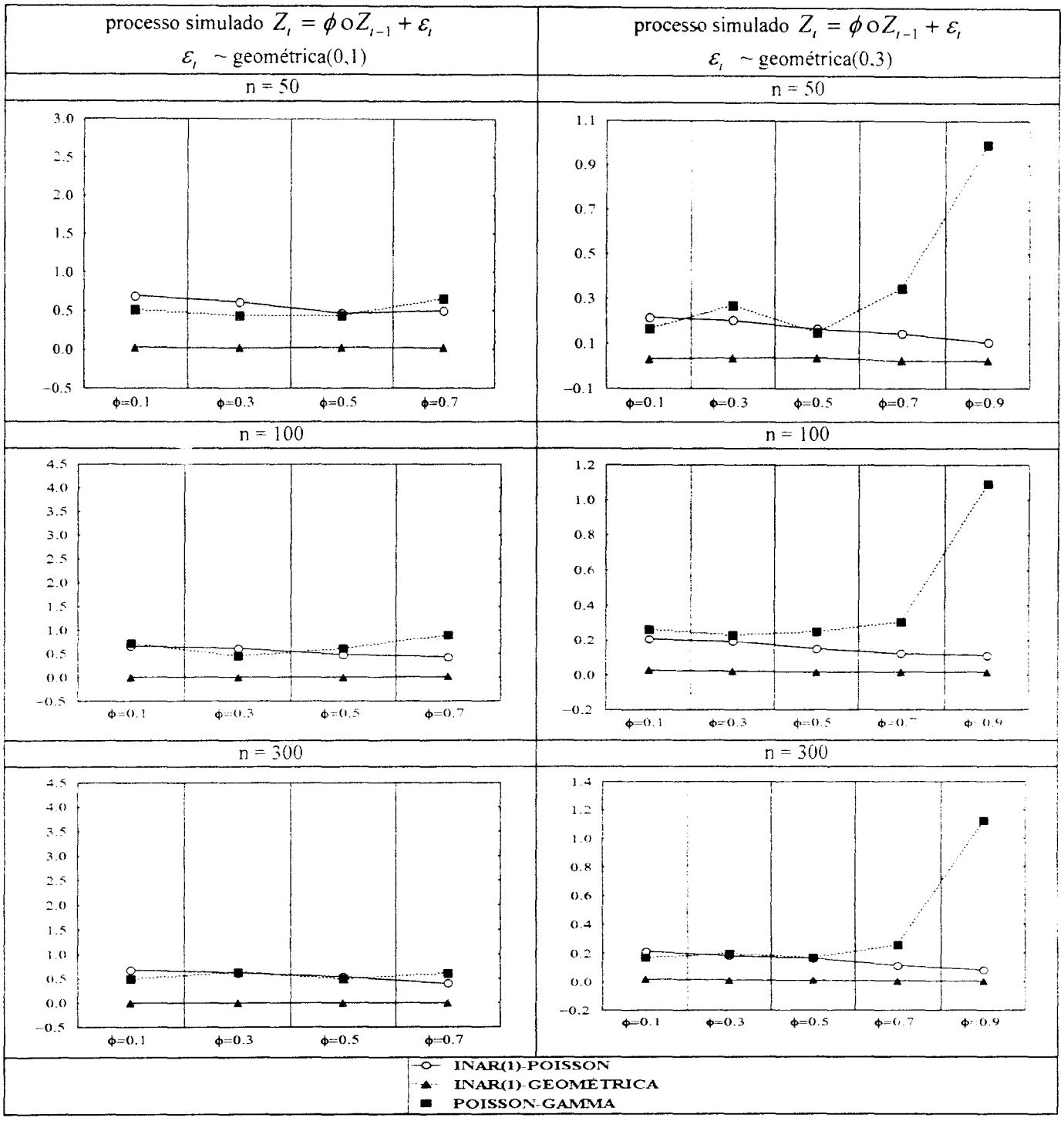


FIGURA 4.4.7b - Divergência de Kullback-Leibler entre a distribuição preditiva real e as distribuições preditivas dos modelos ajustados - simulação do processo INAR(1)-Geométrica.

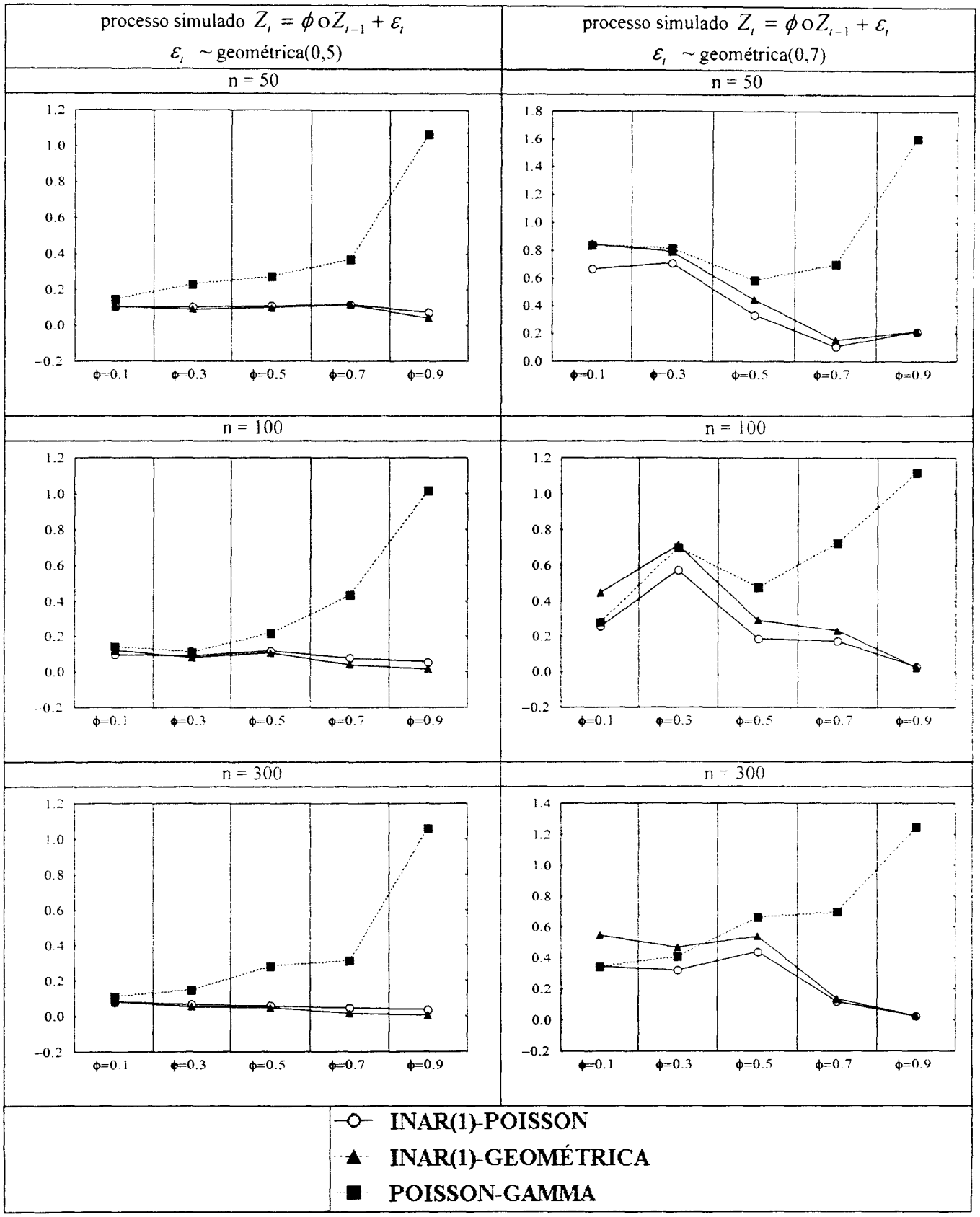
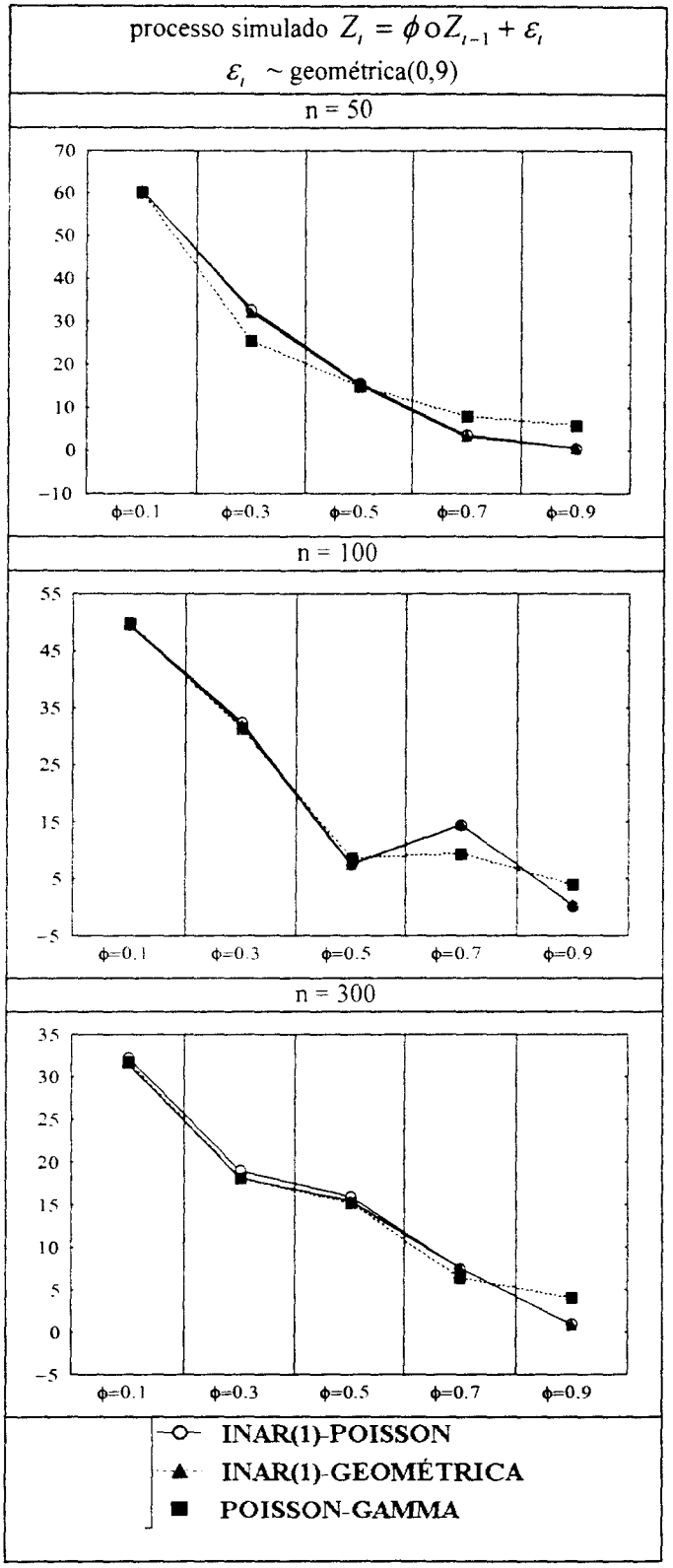


FIGURA 4.4.7c - Divergência de Kullback-Leibler entre a distribuição preditiva real e as distribuições preditivas dos modelos ajustados - simulação do processo INAR(1)-Geométrica.



Os resultados que encontramos para as divergências confirmam as distâncias que encontramos na média e variância das distribuições preditivas dos modelos ajustados em relação aos valores verdadeiros. Em séries de grande magnitude, θ = 0.1 e θ =0.3 , as divergências apresentadas pelos 3 modelos ajustados estiveram próximas quando a correlação serial do processo foi pequena. Na figura 4.4.7.a observamos que para θ =0.3, e a medida que aumentamos o valor de ϕ na simulação, a divergência calculada para o modelo Poisson-Gama começa a se distanciar dos demais , enquanto a do modelo INAR(1)-Poisson se aproxima do modelo real. Na Figura 4.4.7.b com θ = 0.5, os modelos INAR(1) apresentam comportamentos semelhantes, divergindo pouco da distribuição preditiva real, enquanto a divergência do modelo Poisson-Gama cresce com o aumento da correlação serial. A partir de valores de θ=0.7, notamos uma mudança no comportamento das divergências. Nos modelos INAR(1), embora as divergências destes modelos em relação a distribuição preditiva real estejam próximas, esta começa a se comportar de forma decrescente com os valores de ϕ. Lembramos aqui que a partir destes casos, a sobredispersão do processo simulado começa a diminuir e o modelo INAR(1)-Poisson começa, por isso, a ser uma boa aproximação dando a impressão na figura de ser melhor do que o modelo real. Para o modelo Poisson-Gama, a divergência para séries com correlação pequena está próxima dos modelos INAR(1), porém esta aumenta e se afasta da distribuição real com o aumento na correlação serial. A Figura 4.4.7.c ilustra as divergências para os processos simulados com magnitude extremamente pequena . As divergências apresentadas pelos 3 modelos são altas e decrescem a medida em que é utilizado um valor maior para ϕ e logo o processo passa a ter um número maior de observações diferentes de zero.

4.5 - RESULTADOS OBTIDOS A PARTIR DA SIMULAÇÃO DO PROCESSO POISSON-GAMA

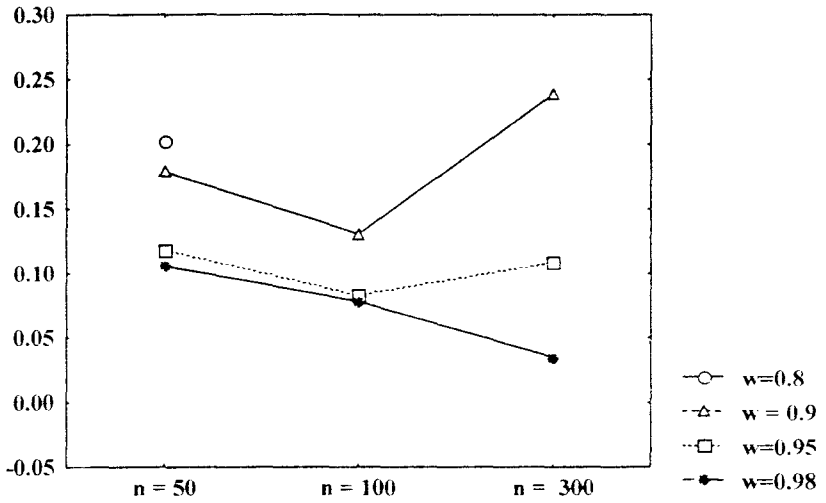
4.5.1 -Ajuste dos modelos.

As simulações do processo Poisson-Gama foram obtidas com valores inciais $a_0 = 10$ e $b_0 = 1$ e desta forma a magnitude das séries foi pequena. Como o modelo Poisson-Gama supõe nível localmente constante, séries não estacionárias podem ser simuladas , dificultando assim o ajuste dos modelos INAR(1). Em seguida, apresentamos os resultados encontrados para o ajuste do modelos INAR(1).

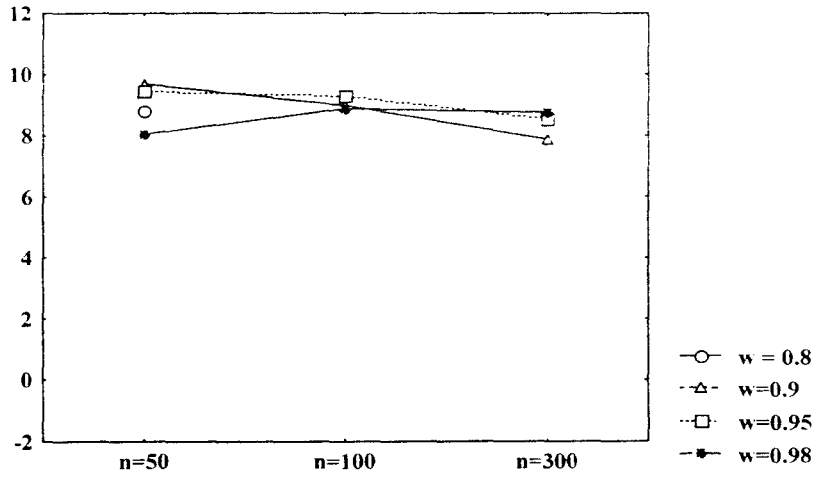
4.5.1 - Ajuste do modelo INAR(1)-Poisson ao processo Poisson-Gama simulado.

Figura 4.5.1 - Parâmetros estimados no ajuste do modelo INAR(1)-Poisson a processos Poisson-Gama simulados - média com base em 20 replicações.

Estimativas para α, parâmetro de correlação serial do modelo INAR(1)-Poisson



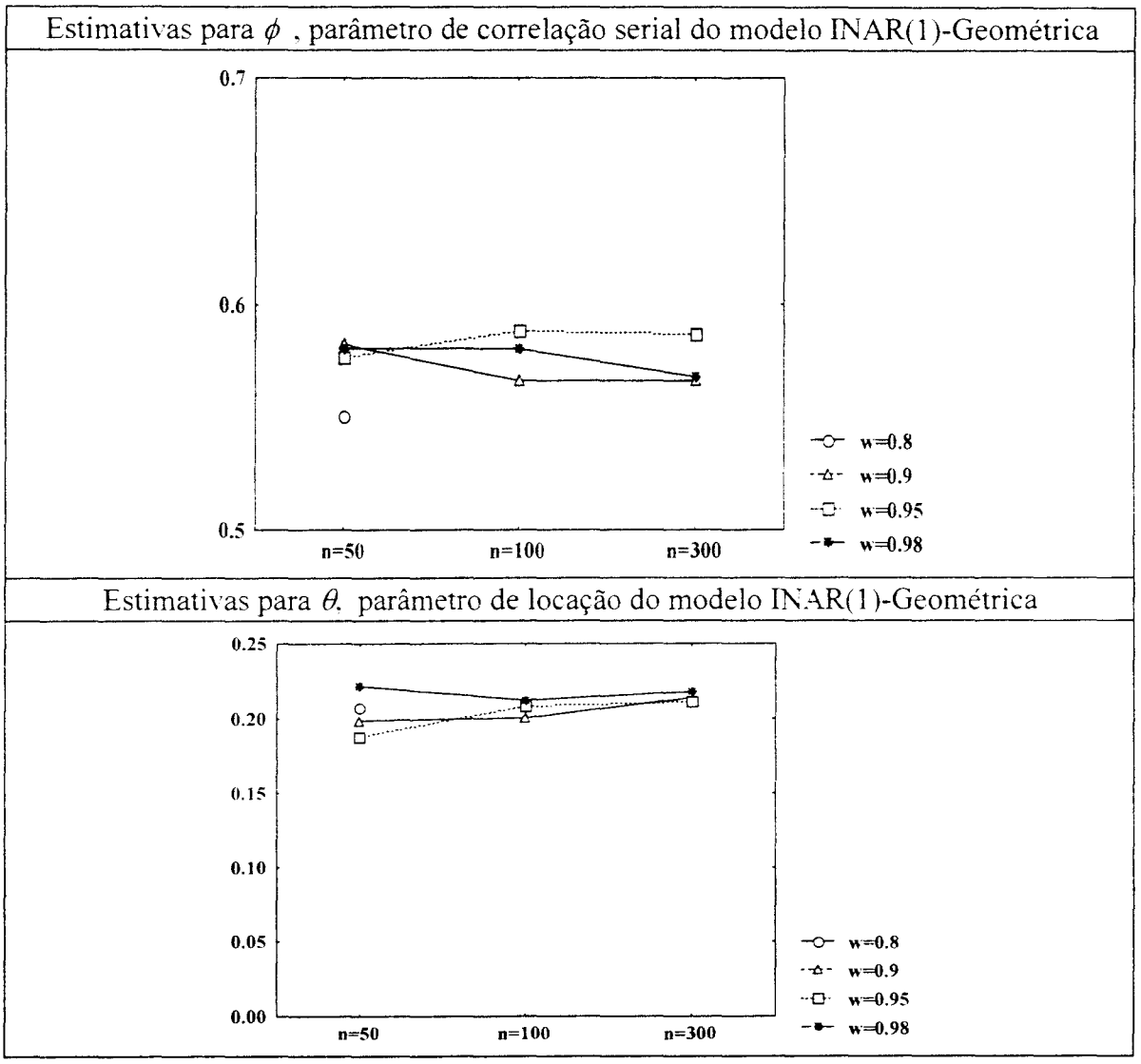
Estimativas para λ, parâmetro de locação do modelo INAR(1)-Poisson



Conforme discutimos na seção 4.3, ao fixarmos valores altos de ω, o processo Poisson-Gama fica mais próximo do comportamento estático e nestas situações o modelo INAR(1)-Poisson ajustado tem parâmetro de correlação serial próximo de zero, especialmente para séries de tamanho n=300. A medida que diminuimos os valores de ω, há um aumento na correlação serial. A estimativa para o parâmetro de locação esteve próxima do nível inicial (a_0/b_0=10) escolhido para a simulação, principalmente em séries de tamanho n=50 .

4.5.1.2 - Ajuste do modelo INAR(1)-Geométrica ao processo Poisson-Gama simulado.

Figura 4.5.2 - Parâmetros estimados no ajuste do modelo INAR(1)-Geométrica a processos Poisson-Gama simulados - média com base em 20 replicações.



Os valores médios ajustados tanto para a correlação como para o parâmetro de locação não foram influenciados pelas escolhas de ω e n . As estimativas estiveram, em média, próximas de 0.2 e 0.58 para ϕ e θ, respectivamente. Analogamente ao caso em que este modelo foi ajustado ao processo INAR(1)-Poisson, ao substiuirmos estes dois valores e calcularmos a média geral do processo , verificaremos que $\frac{(1-\hat{\theta})}{\hat{\theta}(1-\hat{\phi})} \approx 10$, ficando esta próxima do nível utilizado para iniciar a simulação. Comparando este ajuste com o feito pelo modelo INAR(1)-Poisson, verificamos que o parâmetro de correlação serial do modelo INAR(1)-Geométrica não é desprezível.

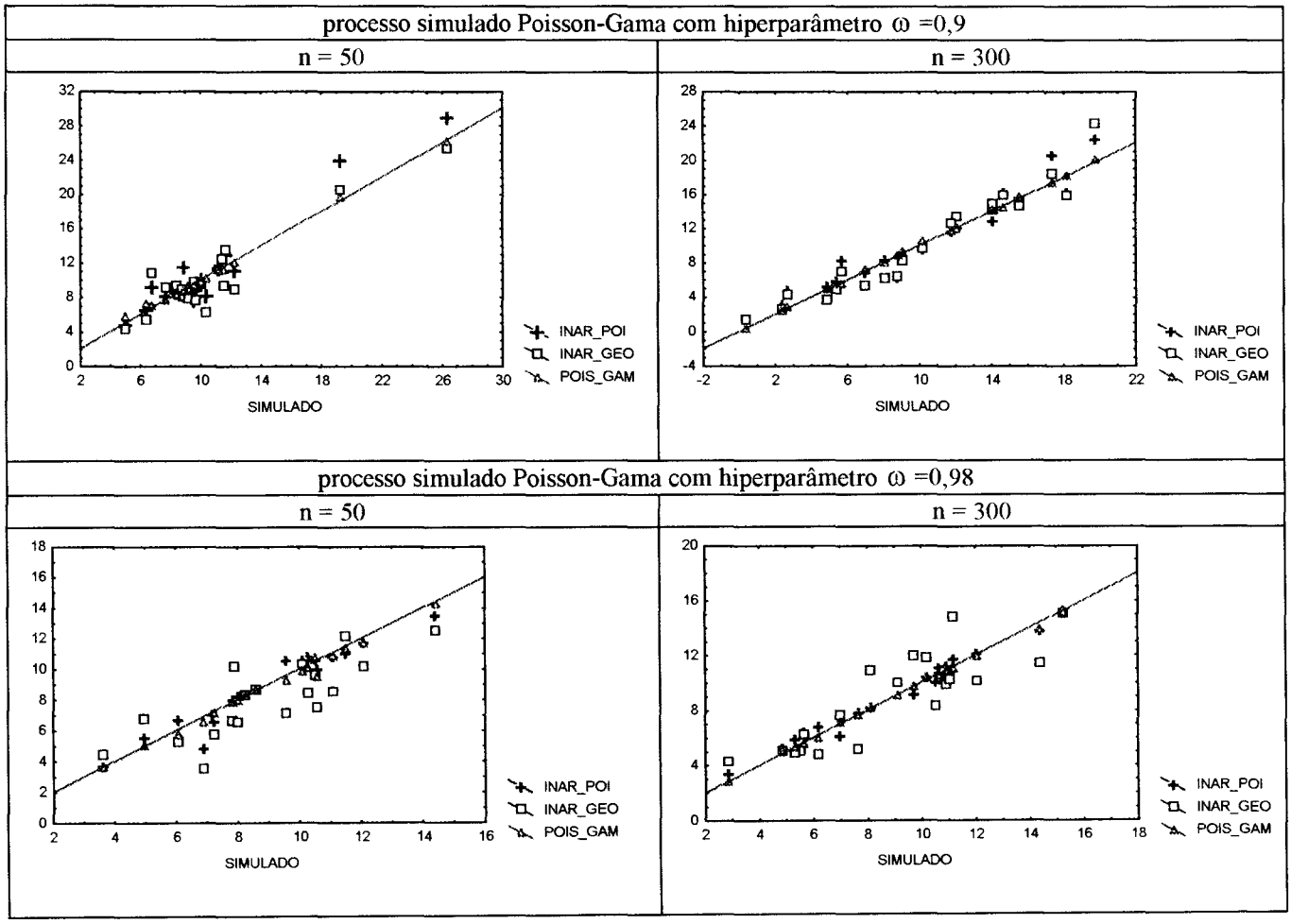
4.5.2 - Estatísticas principais da distribuição preditiva 1-passo-à-frente

Nesta sub-seção, o modelo real é o Poisson-Gama e passaremos a estudar o comportamento da distribuição preditiva 1-passo-à-frente com as distribuições obtidas nos ajustes dos modelos INAR(1) assim como a distribuição preditiva verdadeira . Na seção 4.2 em que esudamos o ajuste dos modelos a séries reais, verificamos que na série de furtos em Hyde Park, havia uma grande diferença entre a distribuição preditiva do modelo INAR(1)-Geométrica em relação as obtidas através dos modelos Poisson-Gama e INAR(1)-Poisson. Em processos INAR(1)-Poisson simulados, a tendência do modelo INAR(1)-Geométrica foi superestimar a variância, apresentando grandes desvios em relação a média verdadeira em várias simulações. Portanto utilizaremos dois diagramas de dispersão (Figuras 4.5.3 e 4.5.4) para verificar se esta situação ocorre também nas simulações do processo Poisson-Gama.

4.5.2.1 - Média da distribuição preditiva 1-passo-à-frente.

Apresentamos na Figura 4.5.3, o diagrama de dispersão entre as médias das distribuições preditivas dos modelos ajustados e os valores verdadeiros.

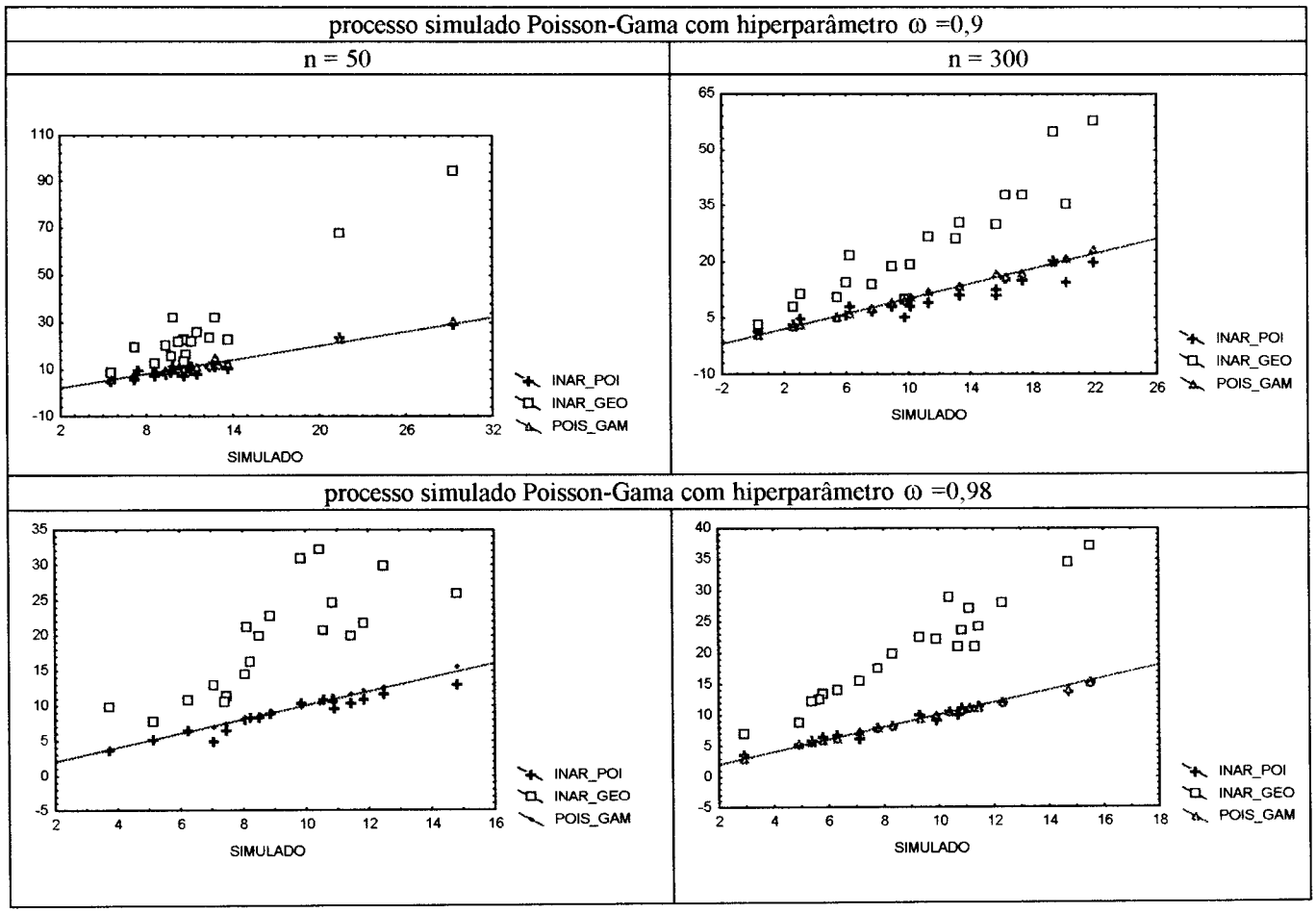
Figura 4.5.3 - Desvios das médias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo Poisson-Gama com hiperparâmetro ω =0,9



Podemos notar na figura 4.5.3 a tendência geral da estimativa da média da distribuição preditiva obtida com o modelo real (Poisson-Gama), é ficar muito próxima da média real em qualquer simulação e ter um desempenho muito superior aos demais. Entretanto, para séries de tamanho grande e valores altos ω, o modelo INAR(1)-Poisson também dá uma boa aproximação da média real. Fica claro também que o modelo INAR(1)-Geométrica ajustado é aquele cuja a média da distribuição preditiva apresenta os maiores desvios em relação ao seu valor verdadeiro, quaisquer que sejam os fatores utilizados na simulação. Entretanto, para ω=0,9, a aproximação deste modelo é um pouco melhor.

4.5.2.2 - Variância da distribuição preditiva 1-passo-à-frente

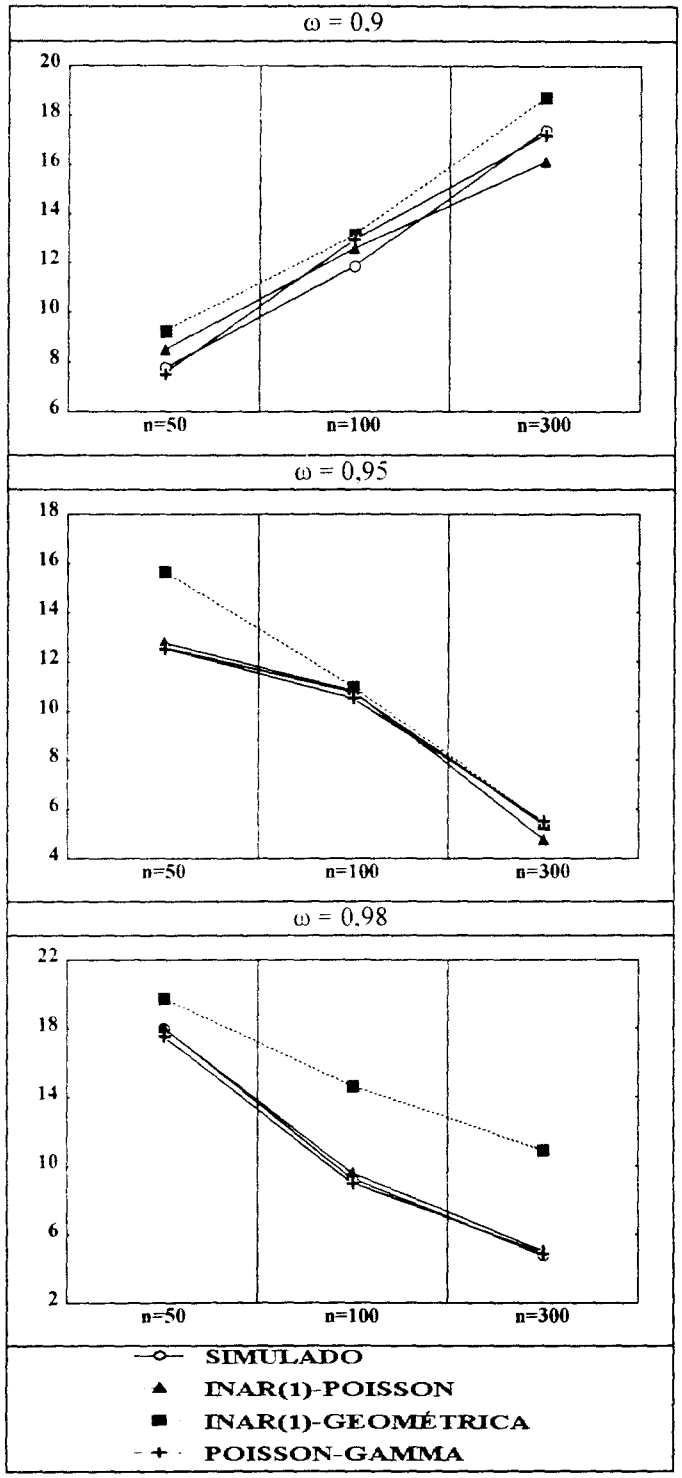
Figura 4.5.4 - Desvios das variâncias das distribuições preditivas um passo a frente dos modelos ajustados em relação a média da distribuição preditiva do modelo Simulado, para simulações do processo Poisson-Gama.



Tal como ocorreu quando o modelo INAR(1)-Geométrica foi ajustado às simulações do processo INAR(1)-Poisson, verificamos que este superestima a variância da distribuição preditiva 1-passo-à-frente em várias circunstâncias. A situação é esperada nestes processos pois o modelo Poisson-Gama supõe da mesma forma que as observações são geradas através da distribuição Poisson e ainda mais se utilizarmos valores altos para o hiperparâmetro na simulação, tal como foi feito acima, garantimos que não ocorram grandes oscilações no nível do processo. Por isso, que o melhor dos resultados apresentados pelo modelo INAR(1)-Geométrica na Figura 4.5.4 foi em simulações de pequeno tamanho, ou seja, um período curto demais para o nível do processo apresentar grandes oscilações, e hiperparâmetro ω=0,9 que garante torna o processo menos dinâmico.

4.5.2.3 - Erro Quadrático Médio da Previsão 1-passo-à-frente (EQMP)

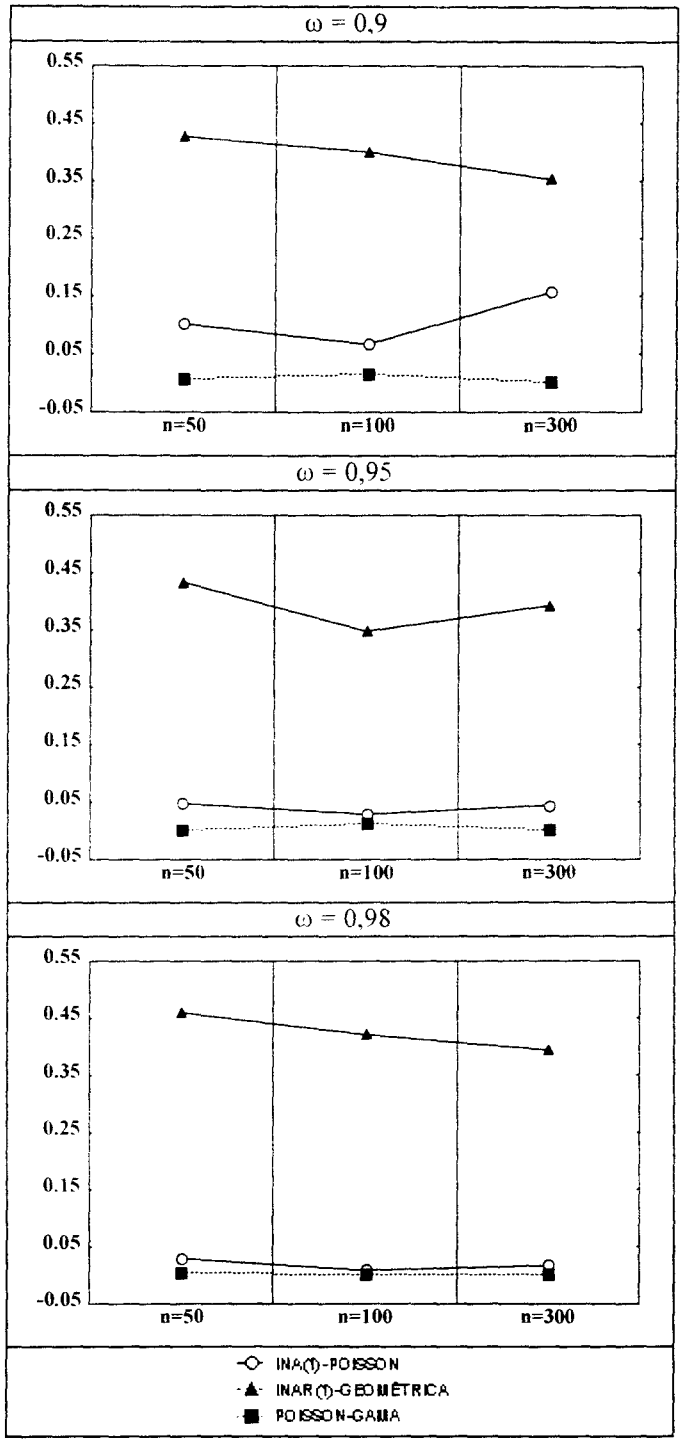
Figura 4.5.5 - Comportamento do erro quadrático - média em 20 replicações - obtido ao utilizar como previsão pontual 1-passo-à-frente a média das distribuições preditivas 1-passo-à-frente dos modelos ajustados ao processo Poisson-Gama simulado.



A Figura 4.5.5 indica que o erro de previsão nestes processos simulados é maior no ajuste do modelo INAR(1)-Geométrica, principalmente para $\omega = 0{,}98$. O resultado confirma o desempenho ruim deste modelo que superestimou a variância da distribuição preditiva 1-passo-à-frente e apresentou os maiores desvios em relação ao valor verdadeiro da média. Por outro lado, os modelos INAR(1)-Poisson fornecem EQMP levemente maiores do que o modelo real, mas quando o hiperparâmetro utilizado é alto e as simulações são de tamanho grande, estes se equivalem aos erros apresentados pelo modelo real. A próxima sub-seção nos traz os resultados para a medida de afinidade de Kullback-Leibler entre as distribuições preditivas dos modelo ajustados e a distribuição preditiva real. Durante a análise dos resultados para as estatísticas principais da distribuição preditiva 1-passo-à-frente na simulação deste processo, já acumulamos algumas pistas que indicam uma grande divergência do modelo INAR(1)-Geométrica em relação aos demais, pois este superestimou a variância e foi notadamente inferior no EMQP.

4.5.3 - Afinidade entre as distribuições preditivas 1-passo-à-frente

Figura 4.5.6 - Divergência de Kullback-Leibler - média em 20 replicações entre as distribuições preditivas 1-passo-à-frente dos modelos ajustados e a distribuição preditiva real - simulação do processo Poisson-Gama



A Figura 4.5.6 confirma a pequena afinidade da distribuição preditiva do modelo INAR(1)-Geométrica com a distribuição preditiva real já ilustrada quando comparamos a média, variância e erro quadrático de previsão. A afinidade medida para a distribuição preditiva do modelo INAR(1)-Poisson esteve bem próxima do resultado encontrado no ajuste do modelo real para o hiperparâmetro fixado em 0,98.

Nestes processos Poisson-Gama simulados, podemos resumir os resultados da seguinte forma : O modelo INAR(1)-Geométrica é deficiente quando tenta aproximar este processo, principalmente a variância; A medida em que o hiperparâmetro utilizado na simulação se aproxima de 1, o modelo INAR(1)-Poisson é uma boa aproximação para estes processos nas características que foram analisadas.

4.6- Seleção de modelos

Nesta seção, os modelos INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama são comparados, nas séries simuladas, em relação aos seus resultados em dois critérios tradicionalmente utilizados para a seleção do melhor dentre possíveis modelos. Os critérios AIC (Akaike, 1973) e BIC (Schwarz, 1978) já foram utilizados na seção 4.2 quando utilizamos séries reais. Naquela ocasião para a série de gols, os modelos INAR(1) tiveram um desempenho ligeiramente superior e para a série de furtos que apresentou maiores mudanças estruturais , o modelo Poisson-Gama obteve um desempenho consideravelmente melhor. Abaixo, apresentamos os dois critérios.

$$AIC = -2\log(L) + 2p$$
$$BIC = -2\log(L) + p\log(n)$$

$\log(L)$: log-verossimilhanca do modelo ajustado

p : número de parâmetros independentes utilizados no modelo ajustado

n : número de observações

Os resultados são apresentados em forma de tabelas contendo as freqüências de escolhas corretas de cada modelo. Ao simularmos, por exemplo, o processo INAR(1)-Poisson, estaremos interessados em selecionar um dos 3 modelos ajustados através do menor AIC ou BIC. As situações nas quais um modelo incorreto for selecionado em muitas replicações do PGD, teremos uma indicação da robustez deste modelo no sentido de que ele pode aproximar o modelo INAR(1)-Poisson.

4.6.1 - Seleção de modelos -simulação do processo INAR(1)-Poisson

O primeiro PGD a ser estudado é o INAR(1)-Poisson. A tabela 4.6.1 contém a freqüência de escolhas de cada modelo através dos critérios AIC e BIC.

Tabela 4.6.1 - Freqüência de escolhas , através dos critérios AIC e BIC, de cada modelo em 20 replicações do processo INAR(1)-Poisson.

α	λ	n	INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
			AIC	BIC	AIC	BIC	AIC	BIC
0.1	1	50	17	14	3	3	0	3
		100	18	15	0	0	2	5
		300	20	18	0	0	0	2
	2	50	19	14	0	0	1	6
		100	20	19	0	0	0	1
		300	20	18	0	0	0	2
	3	50	17	11	0	0	3	9
		100	20	19	0	0	0	1
		300	20	16	0	0	0	4
0.3	1	50	18	16	0	1	1	3
		100	20	19	0	0	0	1
		300	20	20	0	0	0	0
	2	50	19	18	0	0	1	2
		100	20	20	0	0	0	0
		300	20	20	0	0	0	0
	3	50	18	14	1	1	1	5
		100	20	19	0	0	0	1
		300	20	20	0	0	0	0
0.5	1	50	17	17	3	3	0	0
		100	18	18	2	2	0	0
		300	20	20	0	0	0	0
	2	50	20	20	0	0	0	0
		100	20	20	0	0	0	0
		300	20	20	0	0	0	0
	3	50	18	17	2	2	0	1
		100	19	19	1	1	0	0
		300	20	20	0	0	0	0
0.7	1	50	16	16	4	4	0	0
		100	17	17	3	3	0	0
		300	20	20	0	0	0	0
	2	50	19	19	1	1	0	0
		100	20	20	0	0	0	0
		300	20	20	0	0	0	0
	3	50	20	20	0	0	0	0
		100	20	20	0	0	0	0
		300	20	20	0	0	0	0
0.9	1	50	15	15	5	5	0	0
		100	18	18	2	2	0	0
		300	20	20	0	0	0	0
	2	50	19	19	1	1	0	0
		100	20	20	0	0	0	0
		300	20	20	0	0	0	0
	3	50	10	10	10	10	0	0
		100	19	19	1	1	0	0
		300	20	20	0	0	0	0

Podemos observar na tabela 4.6.1 que segundo os critérios AIC e BIC, o modelo INAR(1)-Poisson é corretamente identificado na grande maioria das simulações e a grande maioria de escolhas incorretas ocorre em simulações de tamanho pequeno. Ao utilizar o critério AIC, simulações com valores altos de α, resultam em algumas identificações do modelo INAR(1)-Geométrica como o correto, principalmente para séries de tamanho pequeno. Para valores pequenos de α, em especial, 0,1, o modelo Poisson-Gama é identificado algumas vezes como sendo o correto, através deste critério. Os resultados nas identificações dos modelos INAR não diferem quando utilizamos o critério BIC, porém para simulações com pequena correlação serial, há um aumento no número de escolhas do modelo Poisson-Gama.

Na tabela 4.6.2, apresentamos as freqüências de escolhas de cada modelo, segundo o valor fixo do parâmetro de dependência temporal do modelo INAR(1)-Poisson, representado por α, e o tamanho da série utilizados nas simulações.

Tabela 4.6.2 - Freqüência de escolhas dos modelos, através dos critérios AIC e BIC, segundo o parâmetro α, em 60 replicações do processo INAR(1)-Poisson simulado .

α	n	INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
		AIC	BIC	AIC	BIC	AIC	BIC
0,1	50	53	39	3	3	4	18
	100	58	53	0	0	2	7
	300	60	52	0	0	0	8
0,3	50	55	48	2	2	3	10
	100	60	58	0	0	0	2
	300	60	60	0	0	0	0
0,5	50	55	54	5	5	0	1
	100	57	57	3	3	0	0
	300	60	60	0	0	0	0
0,7	50	55	55	5	5	0	0
	100	57	57	3	3	0	0
	300	60	60	0	0	0	0
0,9	50	44	44	16	16	0	0
	100	57	57	3	3	0	0
	300	60	60	0	0	0	0

A tabela 4.6.2 , de acordo com os resultados obtidos na seção 4.3, nos mostra que os modelos INAR(1)-Poisson e INAR(1)-Geométrica apresentam maior semelhança para processos INAR(1)-Poisson simulados com alta correlação serial de forma que a escolha do modelo incorreto, dentre estes dois, aumenta com o valor de α. O modelo Poisson-Gama é identificado incorretamente por estes critérios, em séries com pequena correlação serial sendo que para valores de $\alpha > 0.5$ ocorre apenas 1 escolha deste modelo. Na tabela 4.6.3, observamos o comportamento dos modelos de acordo com o parâmetro de locação das inovações (λ).

Tabela 4.6.3 - Freqüência de escolhas dos modelos, através dos critério AIC e BIC, segundo o parâmetro de locação λ, em 100 replicações do processo INAR(1)-Poisson .

λ	n	INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
		AIC	BIC	AIC	BIC	AIC	BIC
1	50	83	78	16	16	1	6
	100	91	87	7	7	2	6
	300	100	98	0	0	0	2
2	50	96	90	2	2	2	8
	100	100	99	0	0	0	1
	300	100	98	0	0	0	2
3	50	83	72	13	13	4	15
	100	98	96	2	2	0	2
	300	100	96	0	0	0	4

Na tabela 4.6.3, verificamos que embora haja um aumento nas escolhas do modelo Poisson-Gama, através do critério BIC, de acordo com o aumento na locação das inovações, o fator preponderante nas escolhas incorretas é o tamanho da série.

Através dos resultados demonstrados nesta seção, verificamos que o AIC mostrou ser um critério melhor do que o BIC pois este seleciona o modelo correto com maior frequência, sendo que a escolha incorreta do modelo INAR(1)-Geométrica ocorre muitas vezes para séries de tamanho pequeno e com alta correlação serial.

4.6.2 - -Seleção de modelos - simulação do processo INAR(1)-Geométrica

Iremos proceder nesta sub-seção a escolha do melhor modelo, segundo os critérios AIC e BIC, para as simulações do processo INAR(1)-Geométrica. Os resultados obtidos através da utilização dos critérios AIC e BIC são apresentados na tabela 4.6.4.

Tabela 4.6.4 - Freqüência de escolhas, através dos critérios AIC e BIC, de cada modelo nas 20 replicações do processo INAR(1)-Geométrica - séries com pequena correlação serial (ϕ=0,1 e ϕ=0,3).

ϕ	θ	n	INAR(1)-POISSON		INAR(1) - GEOMÉTRICA		POISSON-GAMA	
			AIC	BIC	AIC	BIC	AIC	BIC
0.1	0.1	50	0	0	20	20	0	0
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.3	50	0	0	20	20	0	0
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.5	50	3	3	17	17	0	0
		100	1	0	19	19	0	0
		300	0	0	20	20	0	0
	0.7	50	12	11	8	7	0	2
		100	3	2	16	16	1	2
		300	2	0	18	18	0	0
	0.9	50	13	10	7	7	0	3
		100	13	11	7	7	0	2
		300	7	7	13	13	0	0
0.3	0.1	50	0	0	20	20	0	0
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.3	50	0	0	19	19	1	1
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.5	50	4	4	16	16	0	0
		100	3	3	17	17	0	0
		300	0	0	20	20	0	0
	0.7	50	5	5	15	13	0	2
		100	5	5	15	15	0	0
		300	0	0	20	20	0	0
	0.9	50	13	11	7	7	0	2
		100	14	11	6	6	0	3
		300	9	9	11	11	0	0

Na tabela 4.6.4, podemos constatar que em séries com pequena correlação serial e cuja média das inovações é alta *(θ =0,1 e 0,3)*, o acerto na escolha do modelo real é muito alto. A medida em que há uma diminuição na média das inovações, ou seja, uma aumento no valor de θ , o modelo INAR(1)-Poisson passa a ser identificado várias vezes como o correto, principalmente em séries de tamanhos pequenos. Nestes valores escolhidos para a simulação, o modelo Poisson-Gama não constitui uma boa alternativa, segundo os critérios AIC e BIC, principalmente este primeiro, pois há poucas identificações deste modelo como sendo o correto: estas ocorrendo em séries com valores altos de θ.

Tabela 4.6.5 - Freqüência de escolhas, através dos critérios AIC e BIC, de cada modelo nas 20 replicações do processo INAR(1)-Geométrica - séries com moderada correlação serial (ϕ=0.5 e ϕ=0.7).

ϕ	θ	n	INAR-POISSON		INAR-GEOMÉTRICA		POISSON-GAMA	
			AIC	BIC	AIC	BIC	AIC	BIC
0.5	0.1	50	0	0	20	20	0	0
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.3	50	1	0	18	18	1	1
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.5	50	5	5	14	14	1	0
		100	2	2	18	18	0	0
		300	0	0	20	20	0	0
	0.7	50	10	10	10	10	0	0
		100	3	3	17	17	0	0
		300	0	0	20	20	0	0
	0.9	50	15	14	5	5	0	0
		100	8	8	12	12	0	0
		300	7	7	13	13	0	0
0.7	0.1	50	0	0	20	20	0	0
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.3	50	2	2	18	18	0	0
		100	0	0	20	20	0	0
		300	0	0	20	20	0	0
	0.5	50	7	7	13	13	0	0
		100	1	1	19	19	0	0
		300	0	0	20	20	0	0
	0.7	50	4	4	16	16	0	0
		100	6	6	14	14	0	0
		300	2	2	18	18	0	0
	0.9	50	12	12	8	8	0	0
		100	9	9	11	11	0	0
		300	6	6	14	14	0	0

Tabela 4.6.6 - Freqüência de escolhas, através dos critérios AIC e BIC, de cada modelo nas 20 replicações do processo INAR(1)-Geométrica - séries com alta correlação serial (ϕ=0,9).

ϕ	θ	n	INAR-POISSON		INAR-GEOMÉTRICA		POISSON-GAMA	
			AIC	BIC	AIC	BIC	AIC	BIC
0.9	0.1	50	-	-	-	-	-	-
		100	-	-	-	-	-	-
		300	-	-	-	-	-	-
	0.3	50	1	1	19	19	0	0
		100	1	1	19	19	0	0
		300	0	0	20	20	0	0
	0.5	50	2	2	18	18	0	0
		100	4	4	16	16	0	0
		300	0	0	20	20	0	0
	0.7	50	5	5	15	15	0	0
		100	6	6	14	14	0	0
		300	0	0	20	20	0	0
	0.9	50	14	14	6	6	0	0
		100	9	9	11	11	0	0
		300	7	7	13	13	0	0

Quando aumentamos o valor da correlação serial utilizado na simulação do processo INAR(1)-Geométrica, podemos perceber nas tabelas 4.6.5 e 4.6.6, que : o modelo Poisson-Gama não é mais identificado como sendo o correto através destes critérios; o número de seleções do modelo correto aumenta em séries com inovações de pequena magnitude; o modelo INAR(1)-Poisson continua a ser selecionado na maioria das vezes em séries de tamanho pequeno com alto valor de θ.

Tal como fizemos na simulação do processo INAR(1)-Poisson, verificaremos nas tabelas 4.6.7 e 4.6.8, marginalmente os resultados, segundo os parâmetros de correlação e locação do processo INAR(1)-Geométrica.

Tabela 4.6.7 - Freqüência de escolhas dos modelos, através dos critérios AIC e BIC, segundo a correlação serial ϕ, em 100 replicações do processo INAR(1)-Geométrica .

ϕ	n	INAR(1)-POISSON		INAR(1)-Geométrica		Poisson-Gama	
		AIC	BIC	AIC	BIC	AIC	BIC
0,1	50	28	24	72	71	0	5
	100	17	14	82	82	1	4
	300	9	8	91	91	0	1
0,3	50	22	20	77	75	1	5
	100	22	19	78	78	0	3
	300	9	9	91	91	0	0
0,5	50	31	30	67	67	2	3
	100	13	13	87	87	0	0
	300	7	7	93	93	0	0
0,7	50	25	25	75	75	0	0
	100	16	16	84	84	0	0
	300	8	8	92	92	0	0
0,9	50	27	27	73	63	0	0
	100	20	20	50	70	0	0
	300	7	7	63	83	0	0

Tabela 4.6.8 - Freqüência de escolhas através dos critérios AIC e BIC, segundo o parâmetro de locação , em 100 replicações do processo INAR(1)-Geométrica .

θ	n	INAR(1)-POISSON		INAR(1)-Geométrica		Poisson-Gama	
		AIC	BIC	AIC	BIC	AIC	BIC
0,1	50	5	5	65	75	0	0
	100	0	0	70	70	0	0
	300	0	0	70	70	0	0
0,3	50	4	4	94	94	2	2
	100	1	1	99	99	0	0
	300	0	0	100	100	0	0
0,5	50	21	21	78	78	1	1
	100	11	11	89	89	0	0
	300	0	0	100	100	0	0
0,7	50	36	35	64	61	0	4
	100	23	22	76	76	1	2
	300	4	3	96	96	0	1
0,9	50	67	61	33	33	0	6
	100	53	48	47	47	0	5
	300	36	36	64	64	0	0

Nas tabelas 4.6.7 e 4.6.8 , constatamos que a seleção do modelo correto ocorre em grande proporção em grandes amostras. O aumento no parâmetro de locação do processo INAR(1)-Geométrica faz com que o modelo INAR(1)-Poisson seja identificado várias vezes como sendo o correto. Se recorrermos aos resultados encontrados na seção 4.4, constataremos que as séries com valores altos de θ e pequenos de ϕ são aquelas que apresentam uma magnitude muito pequena e pouca variação. Nestas séries, devemos levar em consideração o excesso de zeros que certamente prejudica o modelo Poisson-Gama pois este só inicializa seu procedimento iterativo para estimar o hiperparâmetro após a primeira observação diferente de zero..

Não aconteceram mudanças significativas nos resultados obtidos através do critério BIC em relação ao AIC, apenas houve um pequeno aumento na seleção do modelo Poisson-Gama em séries simuladas com pequena correlação serial.

4.6.3 - Seleção de modelos -simulação do processo Poisson-Gama

Apresentamos, nas tabelas 4.6.9 e 4.6.10, os resultados obtidos na seleção de modelos, através dos critérios AIC e BIC, para as simulações do processo Poisson-Gama.

Tabela 4.6.9 - Frequência de escolhas através dos critérios AIC e BIC, de cada modelo ajustado em 20 replicações do processo Poisson-Gama.

ω	n	INAR(1)-Poisson		INAR(1)-Geométrica		Poisson-Gama	
		AIC	BIC	AIC	BIC	AIC	BIC
0..8	50	7	5	0	0	13	15
	100	-	-	-	-	-	-
	300	-	-	-	-	-	-
0.9	50	14	12	0	0	6	8
	100	12	6	0	0	8	14
	300	1	1	0	0	19	19
0,95	50	16	14	0	0	4	6
	100	17	14	0	0	3	6
	300	6	3	0	0	14	17
0,98	50	17	10	0	0	3	10
	100	19	14	0	0	1	6
	300	12	9	0	0	8	11

Para as simulações do processo Poisson-Gama, o desempenho dos modelos INAR(1)-Geométrica, nos critérios AIC e BIC, foi muito ruim, não sendo este identificado em nehuma das replicações. Entretanto se verificarmos nas Figuras 4.5.3 e 4.5.4 os resultados para média e variância da distribuição preditiva nas simulações deste processo, veremos destacadamente a distância deste modelo em relação aos demais. Por sua vez, o modelo INAR(1)-Poisson é considerado como sendo o correto em muitas das situações acima, principalmente para os valores mais altos do hiperparâmetro utilizado na simulação e também em séries pequenas. A seleção de modelos através do critério BIC demonstrou ser a mais correta nas simulações do processo Poisson-Gama pois o modelo correto foi identificado um número maior de vezes.

4.6.4 - Considerações finais sobre a seleção de modelos.

Os resultados encontrados na seção 4.6.4 são resumidos nas tabelas 4.6.10a,b e c que mostram a freqüência percentual de escolhas corretas dos modelos nas simulações dos diferentes processos.

Tabela 4.6.10 a - Frequência relativa percentual de escolhas corretas -Simulação do processo INAR(1)-Poisson

		INAR(1)-POISSON	INAR(1)-GEOMÉTRICA	POISSON-GAMA
AIC	50	87,3 %	10,3 %	9,3 %
	100	96,3 %	3 %	0,7 %
	300	100 %	0 %	0 %
BIC	50	80 %	10,3 %	9,7 %
	100	94 %	3 %	3 %
	300	97,3 %	0 %	2,7 %

Tabela 4.6.10b - Frequência relativa percentual de escolhas corretas -Simulação do processo INAR(1)-Geométrica

		INAR(1)-POISSON	INAR(1)-GEOMÉTRICA	POISSON-GAMA
AIC	50	25 %	74,37 %	0,63 %
	100	15, 63 %	83,47 %	0,2 %
	300	6,87 %	92,83 %	0 %
BIC	50	24,17 %	73,75 %	2,08 %
	100	14,79 %	80 %	1,05 %
	300	6,7 %	88,96 %	0,21 %

Tabela 4.6.10 c - Frequência relativa percentual de escolhas corretas - Simulação do processo Poisson-Gama

		INAR(1)-POISSON	INAR(1)-GEOMÉTRICA	POISSON-GAMA
AIC	50	67,5 %	0 %	32,5 %
	100	80 %	0 %	20 %
	300	31,7 %	0 %	68,3 %
BIC	50	51,25 %	0 %	48,75 %
	100	56,7 %	0 %	43,3 %
	300	21,7 %	0 %	78,3 %

O critério AIC foi o mais adequado para selecionar os modelos corretos nas simulações dos processos INAR(1), enquanto o critério BIC funcionou melhor nas simulações do processo Poisson-Gama. De uma forma geral, o modelo que demonstrou ser mais robusto foi o INAR(1)-Poisson a medida em que este aproximou bem os outros dois modelos quando estes foram os corretos. Em amostras de tamanho grande (n = 300) , os três processos simulados foram aproximados de forma correta na grande maioria das simulações.

CAPÍTULO 5

CONCLUSÕES

Ao apresentarmos os modelos INAR(1) e Estruturais, nos deparamos com dois enfoques simples, parcimoniosos e de fácil uso computacional. Através das séries reais podemos notar a dificuldade dos modelos INAR(1) em acompanhar possíveis evoluções no nível de uma série, fornecendo modelos de características estáticas tal como o que foi ajustado para a série de gols da Inglaterra, muito embora nesta série este modelo tivesse obtido resultados um pouco melhores do que o modelo estrutural nos critérios de seleção. Na série de furtos em Hyde Park, que possuia observações de maior magnitude e aparente mudança no nível, o modelo estrutural teve um desempenho muito superior aos modelos INAR(1). Ao falarmos dos resultados obtidos através de simulações, quando o PGD foi o processso INAR(1)-Poisson, verificamos que o modelo INAR(1)-Geométrica deu uma boa aproximação para este processo quando a correlação de primeira ordem entre os valores da série foi alta. Nesta situação, a média, variância, EQMP e a afinidade calculadas para a distribuição preditiva estiveram bem próximas nos dois modelos. Além disto, na seleção de modelos na seção 4.6, o modelo INAR(1)-Geométrica demonstrou ser uma alternativa para o modelo INAR(1)-Poisson em séries de tamanho pequeno. O modelo Poisson-Gama aproxima bem a média do processo INAR(1)-Poisson e a variância, quando a correlação serial é pequena, porém ao utilizarmos critérios de seleção, este modelo não supera o ajuste dos modelos INAR(1) neste PGD. A simulação do processo INAR(1)-Geométrica trouxe dificuldades esperadas no ajuste dos modelos incorretos, pelo fato de sua variância exceder a média. A variância da distribuição preditiva nestes casos foi subestimada pelo modelo INAR(1)-Poisson e superestimada pelo modelo Poisson-Gama, Utilizando os critérios AIC e BIC, este PGD não conseguiu ser aproximado pelos outros modelos, principalmente em pontos onde sua sobredispersão é grande. A medida que o parâmetro de locação da distribuição geométrica se aproxima de 0,9 e a correlação serial é alta, este processo começa a ser bem aproximado pelo modelo INAR(1)-Poisson que foi selecionado muitas

vezes como o modelo correto nesta região. Nas poucas simulações do processo Poisson-Gama, todas feitas com valores altos para o hiperparâmetro, o modelo INAR(1)-Poisson demonstrou ser uma boa aproximação, principalmente para valores do hiperparâmetro próximos de 1. Com estes valores, o PGD simulou processos mais estáveis e com a distribuição poisson, logo, o modelo INAR(1)-Poisson descreveu estas séries como sendo uma seqüência de inovações da distribuição de Poisson, tal como o ajuste encontrado na série de gols, e obteve melhor desempenho nos critérios AIC e BIC. Por outro lado, ficaram evidentes as divergências entre os resultado obtidos no ajuste do modelo INAR(1)-Geométrica a processos Poisson-Gama simulados. Além de sua distribuição preditiva superestimar a variância correta, o EMQP e a divergência de Kullback-Leibler mostraram claramente as desvantagens em utilizar este modelo neste PGD. Adicionalmente, o INAR(1)-Geométrica não foi selecionado uma única vez como sendo o modelo correto na simulação do Poisson-Gama, através dos critérios AIC e BIC.

APÊNDICE A

Resultados para a simulação do processo INAR(1)-Poisson,

Tabela A1 - Estimativas médias e desvios-padrões obtidos para os parâmetros no ajuste dos modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama em 20 replicações do processo INAR(1)-Poisson.

PROCESSO SIMULADO $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$ ε_t poisson(λ)			INAR(1)-POISSON				INAR(1)-GEOMÉTRICA				POISSON-GAMMA	
		n	$\bar{\hat{\alpha}}$	$S(\bar{\hat{\alpha}})$	$\bar{\hat{\lambda}}$	$S(\bar{\hat{\lambda}})$	$\bar{\hat{\phi}}$	$S(\bar{\hat{\phi}})$	$\bar{\hat{\theta}}$	$S(\bar{\hat{\theta}})$	$\bar{\hat{\omega}}$	$S(\bar{\hat{\omega}})$
α=0,1	λ = 1	50	0,137	0.126	1,01	0.212	0,257	0.125	0,542	0.050	0,940	0.082
		100	0,108	0.096	0,99	0.169	0,282	0.088	0,560	0.035	0,978	0.047
		300	0,095	0.063	1,00	0.093	0,252	0.052	0,548	0.022	0,992	0.008
	λ = 2	50	0,101	0.127	1,92	0.357	0,368	0.100	0,433	0.045	0,973	0.057
		100	0,102	0.074	2,01	0.224	0,377	0.076	0,421	0.032	0,986	0.016
		300	0,112	0.071	1,97	0.207	0,392	0.036	0,427	0.019	0,989	0.015
	λ = 3	50	0,096	0.105	3,09	0.381	0,466	0.059	0,358	0.024	0,949	0.069
		100	0,092	0.077	3,12	0.363	0,460	0.055	0,353	0.028	0,979	0.014
		300	0,114	0.064	2,95	0.225	0,449	0.039	0,354	0.016	0,984	0.029
α=0,3	λ = 1	50	0,283	0.130	0,99	0.189	0,403	0.095	0,553	0.037	0,907	0.072
		100	0,317	0.117	0,98	0.162	0,442	0.082	0,557	0.033	0,939	0.068
		300	0,286	0.051	1,00	0.067	0,416	0.034	0,550	0.013	0,973	0.021
	λ = 2	50	0,339	0.126	1,89	0.411	0,524	0.066	0,429	0.041	0,893	0.089
		100	0,265	0.089	2,10	0.310	0,493	0.063	0,412	0.031	0,958	0.047
		300	0,281	0.041	2,07	0.161	0,503	0.025	0,413	0.015	0,966	0.034
	λ = 3	50	0,263	0.165	3,22	0.665	0,587	0.068	0,361	0.044	0,918	0.097
		100	0,300	0.102	2,97	0.496	0,572	0.051	0,359	0.030	0,941	0.058
		300	0,299	0.065	3,06	0.307	0,572	0.032	0,350	0.018	0,957	0.046
α=0,5	λ = 1	50	0,484	0.121	1,02	0.305	0,584	0.068	0,561	0.061	0,872	0.128
		100	0,514	0.056	0,95	0.107	0,605	0.040	0,569	0.027	0,905	0.091
		300	0,506	0.034	0,99	0.088	0,593	0.025	0,553	0.018	0,843	0.070
	λ = 2	50	0,501	0.137	2,01	0.656	0,648	0.067	0,429	0.059	0,857	0.115
		100	0,463	0.094	2,20	0.324	0,636	0.051	0,405	0.030	0,891	0.090
		300	0,481	0.032	2,12	0.188	0,647	0.020	0,411	0.017	0,873	0.062
	λ = 3	50	0,456	0.157	3,24	1.027	0,681	0.052	0,356	0.042	0,885	0.115
		100	0,492	0.092	3,08	0.612	0,690	0.037	0,352	0.031	0,846	0.092
		300	0,493	0.030	3,04	0.220	0,685	0.019	0,347	0.016	0,858	0.060
α=0,7	λ = 1	50	0,706	0.066	1,06	0.266	0,761	0.051	0,556	0.060	0,831	0.134
		100	0,704	0.042	1,04	0.166	0,754	0.033	0,544	0.029	0,820	0.090
		300	0,690	0.024	1,07	0.116	0,746	0.017	0,535	0.021	0,765	0.035
	λ = 2	50	0,699	0.081	2,12	0.565	0,784	0.041	0,412	0.037	0,821	0.118
		100	0,721	0.034	1,86	0.255	0,797	0.024	0,434	0.030	0,788	0.073
		300	0,696	0.029	2,02	0.148	0,785	0.020	0,416	0.015	0,786	0.054
	λ = 3	50	0,700	0.056	3,12	0.466	0,811	0.025	0,349	0.025	0,780	0.089
		100	0,694	0.059	3,13	0.644	0,810	0.022	0,347	0.029	0,801	0.080
		300	0,701	0.021	3,03	0.205	0,811	0.012	0,345	0.014	0,785	0.066
α=0,9	λ = 1	50	0,890	0.032	1,19	0.262	0,911	0.022	0,555	0.054	0,853	0.107
		100	0,901	0.017	1,06	0.191	0,918	0.013	0,558	0.040	0,816	0.051
		300	0,900	0.011	1,03	0.095	0,918	0.007	0,549	0.019	0,798	0.035
	λ = 2	50	0,897	0.018	2,54	0.407	0,930	0.008	0,402	0.029	0,844	0.102
		100	0,902	0.017	2,09	0.304	0,929	0.010	0,423	0.031	0,787	0.044
		300	0,899	0.010	2,08	0.179	0,929	0.006	0,414	0.017	0,812	0.031
	λ = 3	50	0,850	0.020	5,89	0.544	0,938	0.015	0,352	0.043	0,827	0.087
		100	0,891	0.015	3,52	0.449	0,933	0.006	0,336	0.019	0,798	0.027
		300	0,902	0.007	3,02	0.201	0,938	0.004	0,350	0.010	0,803	0.033

$\overline{\hat{\alpha}}$: média das estimativas encontradas para o parâmetro α do processo INAR(1)-Poisson em 20 simulações,

$S(\overline{\hat{\alpha}})$: desvio pardrão das estimativas encontradas para o parâmetro α do processo INAR(1)-Poisson em 20 simulações,

$\overline{\hat{\lambda}}$: média das estimativas encontradas para o parâmetro λ do processo INAR(1)-Poisson em 20 simulações,

$S(\overline{\hat{\lambda}})$: : desvio pardrão das estimativas encontradas para o parâmetro λ do processo INAR(1)-Poisson em 20 simulações,

$\overline{\hat{\phi}}$: média das estimativas encontradas para o parâmetro ϕ do processo INAR(1)-Geométrica em 20 simulações,

$S(\overline{\hat{\phi}})$: desvio padrão das estimativas encontradas para o parâmetro ϕ do processo INAR(1)-Geométrica em 20 simulações,

$\overline{\hat{\omega}}$: média das estimativas encontradas para o hiperparâmetro do processo Poisson-Gama em 20 simulações

$S(\overline{\hat{\omega}})$: desvio padrão das estimativas encontradas para o hiperparâmetro do processo Poisson-Gama em 20 simulações

Tabela A2 - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Poisson.

PROCESSO SIMULADO $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$ ε_t poisson(λ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
α	λ	n	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$
0,1	1	50	1,075	0.101	1,070	0,224	0,955	0.262	1,095	0.313
		100	1,094	0,072	1,095	0.179	1,012	0.240	1,043	0.232
		300	1,134	0.124	1,114	0.163	1,074	0.302	1,094	0.104
	2	50	2,180	0.113	2,093	0.285	1,864	0.398	2,095	0.347
		100	2,293	0.184	2,292	0.256	2,359	0.660	2,237	0.166
		300	2,263	0.179	2,269	0.278	2,249	0.667	2,247	0.169
	3	50	3,311	0.194	3,469	0.361	3,105	0.852	3,370	0.490
		100	3,281	0.129	3,399	0.257	2,996	0.552	3,407	0.294
		300	3,336	0.175	3,313	0.268	3,144	0.747	3,296	0.238
0,3	1	50	1,410	0.427	1,404	0.515	1,338	0.599	1,388	0.531
		100	1,559	0.430	1,562	0.474	1,551	0.595	1,506	0.309
		300	1,525	0.391	1,529	0.413	1,490	0.535	1,452	0.236
	2	50	2,888	0.464	2,864	0.588	2,753	0.773	2,775	0.660
		100	2,760	0.402	2,816	0.404	2,618	0.608	2,872	0.294
		300	3,087	0.730	3,091	0.723	3,095	1.123	2,954	0.436
	3	50	4,336	.574	4,483	0.689	4,318	1.121	4,556	0.682
		100	4,357	0.564	4,323	0.640	4,212	1.015	4,178	0.521
		300	4,435	0.701	4,488	0.771	4,434	1.262	4,161	0.409
0,5	1	50	1,870	0.737	1,797	0.821	1,735	0.890	1,712	0.647
		100	2,214	0.586	2,190	0.673	2,177	0.743	2,216	0.548
		300	1,823	0.692	1,814	0.713	1,754	0.804	1,818	0.669
	2	50	3,940	1.106	3,942	1.239	3,802	1.409	3,866	1.001
		100	4,393	0.960	4,410	0.955	4,391	1.184	4,377	0.935
		300	4,370	1.042	4,416	1.025	4,411	1.293	4,111	0.717
	3	50	5,893	1.121	5,893	1.028	5,647	1.359	5,711	0.908
		100	5,599	1.219	5,575	1.321	5,336	1.671	5,932	1.343
		300	5,924	0.906	5,919	0.945	5,784	1.220	5,931	1.016
0,7	1	50	2,990	1.137	3,058	1.264	2,953	1.301	3,109	0.986
		100	3,274	1.085	3,316	1.079	3,262	1.132	3,152	0.846
		300	2,855	1.096	2,888	1.113	2,830	1.181	3,127	1.040
	2	50	6,965	1.466	7,103	1.589	6,931	1.644	6,568	1.318
		100	6,619	1.495	6,623	1.607	6,552	1.735	6,267	1.160
		300	6,199	2.004	6,190	1.986	6,075	2.183	6,051	1.500
	3	50	10,419	2.085	10,535	2.199	10,412	2.442	10,145	1.963
		100	10,173	2.174	10,269	2.296	10,132	2.512	9,680	2.156
		300	9,930	2.030	9,953	2.034	9,865	2.305	10,099	1.721
0,9	1	50	9,364	1.887	9,458	1.922	9,253	1.932	8,990	1.760
		100	9,955	2.307	10,037	2.309	9,933	2.342	9,712	1.519
		300	9,190	3.315	9,222	3.313	9,169	3.371	9,059	2.678
	2	50	20,854	3.083	21,326	3.174	20,948	3.261	20,568	3.111
		100	19,550	4.519	19,668	4.708	19,489	4.774	19,354	4.178
		300	20,306	5.453	20,385	5.514	20,293	5.598	19,827	4.124
	3	50	29,505	5.814	30,846	5.628	29,493	6.150	29,010	5.246
		100	29,505	4.846	29,751	4.814	29,418	5.002	29,514	4.002
		300	29,505	6.030	29,624	6.024	29,478	6.247	29,302	5.319

$\bar{\hat{Z}}_{n+1}$: Previsão 1-passo-à-frente utilizando a média da distribuição preditiva (média em 20 replicações),

$S_{\hat{Z}_{n+1}}$: Desvio padrão das previsões 1-passo-à-frente em 20 replicações,

Tabela A3 - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Poisson.

PROCESSO SIMULADO $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$ ε_t poisson(λ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
α	λ	n	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$
0,1	1	50	1,044	0,090	1,035	0,210	1,337	0,401	1,136	0,332
		100	1,049	0,073	1,047	0,173	1,244	0,300	1,033	0,214
		300	1,095	0,101	1,076	0,138	1,345	0,212	1,081	0,104
	2	50	2,130	0,108	2,029	0,292	2,743	0,659	2,155	0,366
		100	2,228	0,159	2,220	0,233	3,053	0,534	2,250	0,191
		300	2,214	0,151	2,197	0,216	2,973	0,365	2,251	0,178
	3	50	3,252	0,174	3,359	0,292	4,601	0,712	3,569	0,628
		100	3,236	0,112	3,339	0,263	4,712	0,699	3,470	0,311
		300	3,286	0,155	3,247	0,244	4,742	0,558	3,340	0,272
0,3	1	50	1,256	0,271	1,230	0,320	1,491	0,313	1,465	0,554
		100	1,367	0,282	1,339	0,323	1,577	0,365	1,558	0,345
		300	1,338	0,249	1,339	0,237	1,570	0,214	1,451	0,228
	2	50	2,593	0,306	2,477	0,457	3,243	0,772	3,068	0,775
		100	2,517	0,268	2,587	0,313	3,458	0,501	2,980	0,347
		300	2,744	0,496	2,778	0,503	3,692	0,650	3,037	0,517
	3	50	3,910	0,394	3,973	0,479	5,218	1,212	5,008	1,055
		100	3,928	0,384	3,848	0,493	5,180	0,918	4,428	0,722
		300	3,991	0,480	4,024	0,504	5,521	0,561	4,339	0,445
0,5	1	50	1,423	0,362	1,422	0,553	1,674	0,600	1,896	0,747
		100	1,589	0,284	1,532	0,350	1,711	0,363	2,382	0,725
		300	1,406	0,341	1,393	0,371	1,699	0,326	2,078	0,737
	2	50	2,969	0,566	2,905	0,833	3,818	1,212	4,517	1,425
		100	3,191	0,477	3,347	0,577	4,237	0,721	4,914	1,338
		300	3,179	0,517	3,296	0,514	4,158	0,498	4,661	0,848
	3	50	4,443	0,565	4,511	0,851	5,924	1,372	6,510	1,272
		100	4,297	0,607	4,317	0,927	5,922	1,228	7,031	1,965
		300	4,459	0,451	4,495	0,557	6,095	0,702	6,879	1,263
0,7	1	50	1,595	0,340	1,644	0,525	1,937	0,637	3,726	1,399
		100	1,680	0,324	1,717	0,442	2,064	0,403	3,737	0,949
		300	1,556	0,329	1,642	0,411	2,069	0,388	3,954	1,274
	2	50	3,495	0,445	3,583	0,851	4,647	0,976	8,090	1,732
		100	3,386	0,450	3,178	0,578	4,030	0,751	7,897	1,508
		300	3,262	0,606	3,283	0,647	4,330	0,706	7,695	1,948
	3	50	5,229	0,630	5,320	1,001	6,977	1,317	13,227	3,478
		100	5,159	0,666	5,277	1,074	7,096	1,429	12,395	3,412
		300	5,079	0,609	5,107	0,689	6,901	0,817	12,999	2,614
0,9	1	50	1,878	0,299	2,149	0,596	2,421	0,914	11,039	2,751
		100	1,898	0,234	1,939	0,333	2,241	0,444	12,485	2,844
		300	1,819	0,332	1,852	0,388	2,226	0,443	11,732	4,251
	2	50	3,894	0,312	4,505	0,894	5,586	1,283	27,472	6,188
		100	3,756	0,453	3,828	0,783	4,727	1,095	25,946	6,704
		300	4,057	1,361	4,075	0,976	5,782	4,189	28,264	15,706
	3	50	5,660	0,590	10,378	2,454	7,958	2,379	40,928	11,526
		100	5,656	0,490	6,406	0,976	8,795	2,853	43,126	14,058
		300	5,653	0,610	5,600	0,479	7,519	1,872	41,771	18,233

$\bar{S}^2_{Z_{n+1}}$: Variância da distribuição preditiva 1-passo-a-frente (média em 20 replicações)

$S_{S^2_{Z_{n+1}}}$: Desvio padrão da variância da distribuição preditiva em 20 replicações

Tabela A4 - Erro quadrático da previsão 1-passo-à-frente utilizando a média da distribuição preditiva 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Poisson.

PROCESSO SIMULADO $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$ ε_t poisson(λ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
α	λ	n	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$
0,1	1	50	1,005	2,105	1,002	2,092	1,163	2,487	0,927	1,756
		100	1,736	3,209	1,866	3,457	1,757	3,216	1,869	3,573
		300	1,365	1,773	1,359	1,742	1,520	1,945	1,298	1,741
	2	50	2,213	2,047	1,850	2,045	2,303	2,430	1,923	2,421
		100	2,212	2,968	2,243	2,859	2,321	2,799	2,390	3,485
		300	1,817	2,596	1,755	2,572	2,138	3,324	1,810	2,537
	3	50	4,167	5,392	3,769	4,786	4,532	6,446	3,809	5,507
		100	2,930	3,447	3,291	3,831	3,707	4,298	3,311	3,953
		300	4,458	5,362	4,409	5,333	4,758	5,228	4,495	5,767
0,3	1	50	0,813	0,894	0,825	0,921	0,800	0,927	0,824	0,844
		100	0,972	1,016	0,878	0,874	1,020	1,072	0,758	1,204
		300	1,160	1,128	1,151	1,125	1,135	1,145	1,167	0,966
	2	50	3,602	4,081	3,669	4,080	4,068	4,685	3,592	4,260
		100	3,762	4,179	3,626	3,531	3,909	4,449	3,744	3,921
		300	2,230	4,148	2,159	3,994	2,520	4,339	2,197	3,933
	3	50	3,257	4,102	3,473	4,814	3,983	4,618	4,075	5,946
		100	2,820	3,025	2,890	3,181	3,445	4,138	2,466	3,374
		300	4,818	6,340	5,073	6,516	5,713	6,802	5,264	8,853
0,5	1	50	2,433	2,275	2,570	2,436	2,681	2,534	2,307	3,017
		100	2,319	2,598	2,416	2,362	2,585	2,559	2,094	2,663
		300	1,986	2,816	1,935	2,579	2,106	2,838	2,171	2,651
	2	50	3,956	5,177	3,785	4,918	4,103	5,147	4,137	4,596
		100	6,758	8,440	7,162	8,450	7,553	8,837	7,245	8,999
		300	4,578	5,595	4,604	5,715	5,065	6,467	4,190	4,339
	3	50	3,529	3,339	3,386	3,089	3,776	3,800	5,049	4,631
		100	7,324	9,256	7,473	9,781	9,185	11,980	4,386	7,014
		300	4,433	5,980	4,401	5,990	4,865	6,180	4,692	7,837
0,7	1	50	2,818	5,381	2,752	5,116	2,851	5,494	2,830	4,386
		100	2,627	3,193	2,669	3,638	2,722	3,510	2,916	3,914
		300	2,192	2,409	2,219	2,456	2,191	2,434	2,514	2,974
	2	50	4,701	6,786	4,995	6,693	4,821	6,559	3,979	4,400
		100	2,640	4,141	2,635	4,104	2,772	4,348	2,938	4,426
		300	4,814	8,112	4,809	7,929	5,308	8,045	2,971	5,003
	3	50	6,004	9,865	6,252	9,526	5,985	10,291	5,883	6,869
		100	18,539	20,983	18,688	21,448	19,045	21,730	16,428	15,737
		300	10,100	12,621	10,310	12,648	10,988	14,699	8,719	9,288
0,9	1	50	4,463	5,562	4,371	5,683	4,323	5,216	3,889	4,294
		100	5,254	7,974	5,241	7,979	5,242	8,097	5,377	11,514
		300	5,347	8,410	5,296	8,330	5,428	8,530	4,206	5,168
	2	50	5,978	10,032	6,595	11,874	6,147	10,756	6,313	10,066
		100	6,142	8,397	6,355	8,709	6,162	8,212	12,152	14,910
		300	6,505	11,865	6,600	11,725	6,845	12,278	10,391	18,530
	3	50	7,911	9,551	7,735	11,089	8,001	8,963	17,968	22,721
		100	8,650	9,945	8,887	9,734	8,833	10,619	14,641	17,636
		300	16,698	23,465	16,703	23,479	17,343	23,946	16,336	29,944

$EMQP_{n+1}$: Média em 20 replicações do erro quadrático de previsão

$S_{EMQP_{n+1}}$: Desvio padrão em 20 replicações do erro quadrático médio de previsão

Tabela A5 - Divergência de Kullback-Leibler entre as distribuições preditivas 1-passo-à-frente do modelo ajustado e do processo simulado - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Poisson.

PROCESSO SIMULADO $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$ ε_t poisson(λ)			INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
α	λ	n	$\overline{D}(P:Q)$	$S_{D(P:Q)}$	$\overline{D}(P:Q)$	$S_{D(P:Q)}$	$\overline{D}(P:Q)$	$S_{D(P:Q)}$
0,1	1	50	0,017	0,025	0,061	0,049	0,053	0,096
		100	0,010	0,015	0,031	0,036	0,029	0,048
		300	0,004	0,006	0,033	0,035	0,009	0,015
	2	50	0,019	0,017	0,095	0,083	0,027	0,035
		100	0,007	0,008	0,077	0,096	0,013	0,012
		300	0,005	0,006	0,085	0,076	0,003	0,003
	3	50	0,017	0,019	0,143	0,155	0,032	0,042
		100	0,010	0,013	0,110	0,105	0,015	0,023
		300	0,003	0,005	0,119	0,144	0,012	0,015
0,3	1	50	0,039	0,101	0,051	0,037	0,107	0,133
		100	0,014	0,016	0,031	0,044	0,073	0,108
		300	0,004	0,005	0,023	0,030	0,054	0,101
	2	50	0,014	0,015	0,057	0,041	0,062	0,136
		100	0,014	0,014	0,068	0,058	0,035	0,032
		300	0,004	0,005	0,071	0,079	0,063	0,060
	3	50	0,027	0,031	0,110	0,137	0,072	0,069
		100	0,013	0,021	0,080	0,095	0,056	0,057
		300	0,005	0,007	0,102	0,129	0,070	0,065
0,5	1	50	0,026	0,036	0,058	0,058	0,195	0,184
		100	0,011	0,012	0,028	0,038	0,175	0,170
		300	0,003	0,003	0,043	0,037	0,141	0,136
	2	50	0,030	0,034	0,086	0,068	0,147	0,108
		100	0,014	0,021	0,050	0,032	0,152	0,133
		300	0,004	0,011	0,053	0,041	0,132	0,118
	3	50	0,019	0,027	0,082	0,073	0,178	0,181
		100	0,010	0,010	0,113	0,085	0,244	0,253
		300	0,003	0,003	0,076	0,056	0,137	0,141
0,7	1	50	0,024	0,032	0,062	0,050	0,345	0,222
		100	0,008	0,008	0,039	0,025	0,313	0,179
		300	0,004	0,004	0,052	0,025	0,299	0,132
	2	50	0,023	0,017	0,065	0,037	0,397	0,228
		100	0,011	0,010	0,063	0,025	0,346	0,153
		300	0,004	0,003	0,074	0,042	0,384	0,187
	3	50	0,017	0,018	0,085	0,048	0,425	0,267
		100	0,012	0,011	0,088	0,047	0,427	0,292
		300	0,003	0,003	0,085	0,048	0,388	0,188
0,9	1	50	0,021	0,021	0,056	0,027	1,398	0,443
		100	0,011	0,012	0,046	0,017	1,383	0,254
		300	0,003	0,003	0,045	0,011	1,333	0,303
	2	50	0,043	0,027	0,077	0,025	1,588	0,426
		100	0,014	0,013	0,078	0,025	1,748	0,403
		300	0,004	0,005	0,072	0,025	1,762	0,477
	3	50	0,239	0,075	0,111	0,047	1,647	0,405
		100	0,016	0,013	0,094	0,032	1,746	0,379
		300	0,004	0,003	0,088	0,027	1,795	0,321

$\overline{D}(P:Q)$: Divergência de Kullback-Leibler entre a distribuição preditiva real P e a distribuição preditiva Q do modelo ajustado,

$S_{D(P:Q)}$: Desvio padrão da divergência de Kullback-Leibler em 20 replicações,

Tabela A6 - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Poisson.

PROCESSO SIMULADO			INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
$Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$								
ε_t poisson(λ)								
α	λ	n	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}
0,1	1	50	142,450	10,790	146,380	10,200	145,540	10,650
		100	272,120	16,580	283,230	17,500	275,480	17,060
		300	820,210	23,950	849,200	25,150	827,410	22,320
	2	50	175,380	10,790	185,100	9,840	177,820	9,940
		100	355,420	15,590	374,490	14,490	359,790	15,660
		300	1056,140	22,940	1108,980	27,100	1064,450	22,970
	3	50	200,760	8,340	212,660	6,960	202,350	8,060
		100	399,430	14,850	425,760	15,740	403,570	15,080
		300	1198,740	28,220	1269,010	25,250	1206,310	28,520
0,3	1	50	147,330	10,040	151,790	9,870	152,790	10,500
		100	294,210	12,800	302,370	13,320	306,780	13,810
		300	876,600	21,130	902,410	21,440	902,830	24,570
	2	50	186,580	9,790	195,340	10,800	191,840	9,530
		100	376,020	15,270	391,730	15,620	385,260	13,380
		300	1127,130	24,930	1172,840	26,290	1153,170	24,680
	3	50	209,320	11,680	218,520	12,360	214,470	12,030
		100	413,710	15,910	434,860	15,990	424,120	14,310
		300	1253,840	25,140	1310,740	29,420	1284,550	24,690
0,5	1	50	156,080	14,980	160,070	13,760	167,920	11,470
		100	306,130	12,250	313,300	14,010	330,570	14,120
		300	937,710	28,540	955,420	27,080	1011,710	29,790
	2	50	194,710	12,770	201,590	14,990	206,730	10,180
		100	396,330	14,120	409,600	15,700	418,850	12,800
		300	1178,84	28,470	1218,480	26,840	1245,050	27,640
	3	50	216,230	10,360	224,000	11,550	227,030	8,120
		100	433,360	13,820	448,380	17,550	456,510	11,260
		300	1299,180	27,580	1347,300	29,730	1375,320	30,090
0,7	1	50	168,440	14,020	171,140	14,280	193,070	13,000
		100	335,780	14,840	342,050	15,080	384,340	9,870
		300	1015,110	33,070	1033,720	34,830	1163,460	32,010
	2	50	205,100	8,940	211,150	9,670	227,200	6,340
		100	397,920	14,650	408,060	15,880	451,330	12,610
		300	1214,090	21,740	1243,080	22,680	1361,260	13,320
	3	50	224,810	7,950	231,320	8,310	250,680	7,680
		100	448,620	17,670	462,760	17,980	495,230	12,130
		300	1339,430	24,330	1380,720	27,460	1491,850	24,700
0,9	1	50	175,180	13,160	178,380	14,680	231,440	7,370
		100	344,150	19,960	350,160	22,560	464,110	13,740
		300	1041,750	27,610	1056,450	27,300	1403,370	17,350
	2	50	215,140	8,780	220,400	9,400	271,990	8,130
		100	415,190	15,320	423,430	15,900	539,370	11,290
		300	1250,390	28,630	1278,300	30,710	1608,140	23,930
	3	50	236,050	9,500	236,430	13,760	291,470	6,170
		100	466,260	11,540	478,360	13,070	579,230	7,730
		300	1368,180	22,070	1399,880	20,550	1733,880	21,620

$\overline{AIC}$: Média do AIC em 20 replicações

S_{AIC} : Desvio padrão do AIC em 20 replicações

Tabela A7 - Resultados obtidos pelos 3 modelos ajustados no critério BIC - média e desvio padrão em 20 replicações do processo INAR(1)-Poisson.

PROCESSO SIMULADO $Z_t = \alpha \circ Z_{t-1} + \varepsilon_t$ ε_t poisson(λ)			INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
α	λ	n	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}
0,1	1	50	146,310	10,790	150,240	10,200	147,480	10,650
		100	277,350	16,580	288,460	17,500	278,100	17,060
		300	827,630	23,950	856,610	25,150	831,110	22,320
	2	50	179,240	10,790	188,960	9,840	179,750	9,940
		100	360,650	15,590	379,720	14,490	362,410	15,660
		300	1063,560	22,940	1116,400	27,100	1068,160	22,970
	3	50	204,620	8,340	216,520	6,960	204,280	8,060
		100	404,660	14,850	430,990	15,740	406,190	15,080
		300	1206,150	28,220	1276,430	25,250	1210,020	28,520
0,3	1	50	151,200	10,040	155,660	9,870	154,720	10,500
		100	299,440	12,800	307,600	13,320	309,390	13,810
		300	884,010	21,130	909,830	21,440	906,540	24,570
	2	50	190,450	9,790	199,200	10,800	193,780	9,530
		100	381,250	15,270	396,960	15,620	387,880	13,380
		300	1134,550	24,930	1180,260	26,290	1156,880	24,680
	3	50	213,180	11,680	222,390	12,360	216,410	12,030
		100	418,940	15,910	440,090	15,990	426,740	14,310
		300	1261,250	25,140	1318,150	29,420	1288,260	24,690
0,5	1	50	159,950	14,980	163,930	13,760	169,850	11,470
		100	311,360	12,250	318,530	14,010	333,180	14,120
		300	945,120	28,540	962,840	27,080	1015,420	29,790
	2	50	198,580	12,770	205,450	14,990	208,670	10,180
		100	401,560	14,120	414,830	15,700	421,460	12,800
		300	1186,26	28,470	1225,890	26,840	1248,750	27,640
	3	50	220,100	10,360	227,870	11,550	228,960	8,120
		100	438,590	13,820	453,610	17,550	459,130	11,260
		300	1306,590	27,580	1354,720	29,730	1379,030	30,090
0,7	1	50	172,300	14,020	175,000	14,280	195,000	13,000
		100	341,010	14,840	347,280	15,080	386,950	9,870
		300	1022,520	33,070	1041,130	34,830	1167,170	32,010
	2	50	208,970	8,940	215,010	9,670	229,130	6,340
		100	403,150	14,650	413,290	15,880	453,950	12,610
		300	1221,510	21,740	1250,490	22,680	1364,960	13,320
	3	50	228,670	7,950	235,180	8,310	252,610	7,680
		100	453,850	17,670	467,990	17,980	497,840	12,130
		300	1346,840	24,330	1388,130	27,460	1495,560	24,700
0,9	1	50	179,040	13,160	182,240	14,680	233,370	7,370
		100	349,380	19,960	355,390	22,560	466,720	13,740
		300	1049,160	27,610	1063,860	27,300	1407,070	17,350
	2	50	219,000	8,780	224,270	9,400	273,920	8,130
		100	420,430	15,320	428,660	15,900	541,990	11,290
		300	1257,810	28,630	1285,720	30,710	1611,850	23,930
	3	50	239,910	9,500	240,290	13,760	293,400	6,170
		100	471,490	11,540	483,590	13,070	581,850	7,730
		300	1375,600	22,070	1407,300	20,550	1737,590	21,620

$\overline{BIC}$: Média do BIC em 20 replicações

S_{BIC} : Desvio padrão do BIC em 20 replicações

APÊNDICE B

Resultados obtidos para a simulação do processo INAR(1)-Geométrica.

Tabela B1.a - Estimativas dos parâmetros obtidos nos ajustes dos modelos INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geoméirica (ϕ=0,1 e ϕ=0,3).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			INAR(1)-POISSON				INAR(1)-GEOMÉTRICA				POISSON-GAMA	
ϕ	θ	n	$\bar{\hat{\alpha}}$	$S(\bar{\hat{\alpha}})$	$\bar{\hat{\lambda}}$	$S(\bar{\hat{\lambda}})$	$\bar{\hat{\phi}}$	$S(\bar{\hat{\phi}})$	$\bar{\hat{\theta}}$	$S(\bar{\hat{\theta}})$	$\bar{\hat{\omega}}$	$S(\bar{\hat{\omega}})$
0,1	0,1	50	0,052	0,059	9,569	1,628	0,090	0,051	0,101	0,017	0,121	0,037
		100	0,052	0,047	9,517	0,890	0,106	0,035	0,101	0,007	0,118	0,013
		300	0,035	0,025	9,669	0,698	0,096	0,019	0,100	0,007	0,114	0,016
	0,3	50	0,071	0,077	2,365	0,336	0,116	0,083	0,311	0,032	0,494	0,098
		100	0,072	0,048	2,322	0,348	0,094	0,054	0,309	0,029	0,500	0,092
		300	0,071	0,024	2,424	0,144	0,101	0,035	0,300	0,013	0,479	0,036
	0,5	50	0,107	0,059	1,069	0,251	0,093	0,071	0,486	0,054	0,786	0,139
		100	0,055	0,073	1,030	0,182	0,089	0,062	0,505	0,048	0,883	0,129
		300	0,067	0,030	1,025	0,086	0,094	0,031	0,502	0,020	0,856	0,090
	0,7	50	0,136	0,117	0,410	0,112	0,056	0,093	0,694	0,057	0,942	0,077
		100	0,102	0,074	0,439	0,073	0,084	0,068	0,693	0,028	0,931	0,060
		300	0,066	0,046	0,444	0,046	0,093	0,052	0,699	0,024	0,981	0,027
	0,9	50	0,080	0,130	0,095	0,051	0,031	0,065	0,908	0,050	0,991	0,023
		100	0,124	0,146	0,113	0,053	0,079	0,112	0,897	0,040	0,975	0,032
		300	0,100	0,070	0,110	0,025	0,070	0,055	0,899	0,022	0,992	0,011
0,3	0,1	50	0,103	0,078	11,659	1,617	0,276	0,036	0,098	0,015	0,147	0,024
		100	0,153	0,058	10,957	1,123	0,294	0,037	0,099	0,008	0,148	0,021
		300	0,174	0,034	10,560	0,526	0,300	0,015	0,101	0,004	0,152	0,011
	0,3	50	0,182	0,085	2,889	0,533	0,294	0,089	0,291	0,035	0,483	0,115
		100	0,198	0,071	2,593	0,447	0,308	0,045	0,313	0,033	0,548	0,142
		300	0,208	0,041	2,649	0,243	0,300	0,031	0,300	0,021	0,477	0,054
	0,5	50	0,211	0,118	1,189	0,253	0,262	0,086	0,478	0,046	0,722	0,123
		100	0,217	0,081	1,120	0,163	0,307	0,060	0,504	0,033	0,802	0,116
		300	0,246	0,049	1,128	0,110	0,307	0,043	0,492	0,024	0,719	0,078
	0,7	50	0,246	0,123	0,434	0,138	0,269	0,123	0,710	0,061	0,919	0,090
		100	0,244	0,111	0,434	0,114	0,272	0,098	0,709	0,044	0,923	0,057
		300	0,267	0,059	0,436	0,057	0,293	0,059	0,705	0,025	0,928	0,056
	0,9	50	0,246	0,184	0,105	0,041	0,239	0,183	0,906	0,034	0,964	0,052
		100	0,270	0,166	0,113	0,041	0,271	0,160	0,900	0,033	0,968	0,040
		300	0,272	0,086	0,110	0,023	0,276	0,086	0,902	0,019	0,981	0,018

Tabela B1.b - Estimativas dos parâmetros obtidos nos ajustes dos modelos INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geoméirica (ϕ=0,5 e ϕ=0,7).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			INAR(1)-POISSON				INAR(1)-GEOMÉTRICA				POISSON-GAMA	
ϕ	θ	n	$\bar{\hat{\alpha}}$	$S(\bar{\hat{\alpha}})$	$\bar{\hat{\lambda}}$	$S(\bar{\hat{\lambda}})$	$\bar{\hat{\phi}}$	$S(\bar{\hat{\phi}})$	$\bar{\hat{\theta}}$	$S(\bar{\hat{\theta}})$	$\bar{\hat{\omega}}$	$S(\bar{\hat{\omega}})$
0,5	0,1	50	0,287	0,051	13,258	2,445	0,497	0,045	0,100	0,020	0,220	0,063
		100	0,326	0,062	11,986	1,619	0,503	0,028	0,102	0,008	0,217	0,038
		300	0,310	0,030	12,384	0,930	0,494	0,015	0,099	0,006	0,246	0,125
	0,3	50	0,343	0,123	3,018	0,811	0,486	0,066	0,306	0,047	0,549	0,133
		100	0,343	0,054	2,982	0,392	0,485	0,029	0,302	0,028	0,503	0,073
		300	0,360	0,030	2,980	0,284	0,502	0,023	0,303	0,019	0,499	0,038
	0,5	50	0,400	0,148	1,275	0,391	0,514	0,100	0,500	0,054	0,756	0,164
		100	0,407	0,087	1,145	0,149	0,494	0,058	0,507	0,026	0,698	0,096
		300	0,423	0,037	1,132	0,107	0,498	0,037	0,505	0,024	0,686	0,046
	0,7	50	0,432	0,096	0,487	0,133	0,448	0,082	0,683	0,053	0,892	0,111
		100	0,462	0,087	0,446	0,104	0,478	0,083	0,702	0,047	0,826	0,081
		300	0,455	0,041	0,454	0,045	0,480	0,040	0,698	0,020	0,827	0,050
	0,9	50	0,454	0,155	0,113	0,041	0,443	0,153	0,898	0,032	0,936	0,065
		100	0,472	0,103	0,126	0,041	0,472	0,099	0,889	0,030	0,949	0,055
		300	0,439	0,080	0,110	0,026	0,440	0,079	0,902	0,021	0,946	0,030
0,7	0,1	50	0,387	0,074	19,240	3,412	0,694	0,027	0,096	0,011	0,293	0,057
		100	0,451	0,069	16,334	2,344	0,700	0,022	0,102	0,011	0,353	0,128
		300	0,460	0,035	15,938	1,509	0,697	0,010	0,101	0,007	0,411	0,187
	0,3	50	0,545	0,077	3,321	0,625	0,694	0,034	0,312	0,030	0,564	0,080
		100	0,548	0,062	3,467	0,759	0,696	0,032	0,305	0,037	0,567	0,080
		300	0,543	0,024	3,558	0,334	0,700	0,016	0,301	0,018	0,560	0,039
	0,5	50	0,608	0,093	1,360	0,396	0,702	0,058	0,499	0,060	0,752	0,119
		100	0,581	0,043	1,428	0,299	0,666	0,047	0,474	0,039	0,656	0,078
		300	0,617	0,033	1,257	0,121	0,692	0,024	0,498	0,020	0,675	0,047
	0,7	50	0,624	0,110	0,472	0,143	0,646	0,087	0,697	0,050	0,758	0,063
		100	0,661	0,067	0,462	0,087	0,681	0,059	0,699	0,034	0,765	0,081
		300	0,681	0,028	0,456	0,064	0,700	0,024	0,700	0,026	0,741	0,038
	0,9	50	0,651	0,117	0,106	0,051	0,661	0,121	0,907	0,046	0,901	0,083
		100	0,657	0,081	0,107	0,045	0,648	0,080	0,903	0,035	0,874	0,052
		300	0,687	0,047	0,111	0,026	0,686	0,047	0,901	0,020	0,867	0,035

Tabela B1,c - Estimativas dos parâmetros obtidos nos ajustes dos modelos INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geoméirica (ϕ=0.9).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			INAR(1)-Poisson				INAR(1)-Geométrica				Poisson-Gama	
ϕ	θ	n	$\bar{\hat{\alpha}}$	$S(\bar{\hat{\alpha}})$	$\bar{\hat{\lambda}}$	$S(\bar{\hat{\lambda}})$	$\bar{\hat{\phi}}$	$S(\bar{\hat{\phi}})$	$\bar{\hat{\theta}}$	$S(\bar{\hat{\theta}})$	$\bar{\hat{\omega}}$	$S(\bar{\hat{\omega}})$
0,9	0,1	50	---	---	---	---	---	---	---	---	---	---
		100	---	---	---	---	---	---	---	---	---	---
		300	---	---	---	---	---	---	---	---	---	---
	0,3	50	0,798	0,049	4,719	1,443	0,897	0,017	0,301	0,031	0,731	0,095
		100	0,801	0,036	4,880	1,155	0,899	0,010	0,291	0,032	0,709	0,082
		300	0,812	0,012	4,315	0,429	0,898	0,005	0,300	0,014	0,700	0,041
	0,5	50	0,855	0,042	1,488	0,664	0,894	0,021	0,496	0,071	0,784	0,092
		100	0,857	0,034	1,456	0,443	0,897	0,016	0,495	0,051	0,741	0,056
		300	0,865	0,013	1,358	0,175	0,899	0,008	0,499	0,023	0,753	0,026
	0,7	50	0,886	0,036	0,500	0,160	0,897	0,028	0,694	0,059	0,857	0,103
		100	0,885	0,028	0,469	0,094	0,892	0,026	0,697	0,039	0,767	0,064
		300	0,889	0,012	0,470	0,058	0,898	0,011	0,698	0,023	0,773	0,032
	0,9	50	0,870	0,102	0,116	0,058	0,908	0,049	0,912	0,055	0,843	0,089
		100	0,894	0,030	0,102	0,042	0,894	0,029	0,909	0,034	0,826	0,045
		300	0,900	0,020	0,112	0,023	0,900	0,020	0,900	0,018	0,807	0,029

$\overline{\hat{\alpha}}$: média das estimativas encontradas para o parâmetro α do processo INAR(1)-Poisson em 20 simulações.

$S(\overline{\hat{\alpha}})$: desvio pardrão das estimativas encontradas para o parâmetro α do processo INAR(1)-Poisson em 20 simulações.

$\overline{\hat{\lambda}}$: média das estimativas encontradas para o parâmetro λ do processo INAR(1)-Poisson em 20 simulações.

$S(\overline{\hat{\lambda}})$: : desvio pardrão das estimativas encontradas para o parâmetro λ do processo INAR(1)-Poisson em 20 simulações,

$\overline{\hat{\phi}}$: média das estimativas encontradas para o parâmetro ϕ do processo INAR(1)-Geométrica em 20 simulações.

$S(\overline{\hat{\phi}})$: desvio padrão das estimativas encontradas para o parâmetro ϕ do processo INAR(1)-Geométrica em 20 simulações.

$\overline{\hat{\omega}}$: média das estimativas encontradas par
a o hiperparâmetro do processo Poisson-Gama em 20 simulações

$S(\overline{\hat{\omega}})$: desvio padrão das estimativas encontradas para o hiperparâmetro do processo Poisson-Gama em 20 simulações

Tabela B2.a - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidas em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,1 e ϕ=0,3)

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
ϕ	θ	n	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$
0,1	0,1	50	10,184	1,058	10,358	2,111	10,344	2,056	10,630	10,297
		100	10,484	1,580	10,354	1,172	10,457	1,583	9,312	7,865
		300	10,075	0,909	10,103	0,873	10,108	1,223	8,908	6,798
	0,3	50	2,506	0,238	2,431	0,355	2,437	0,391	2,334	1,329
		100	2,590	0,321	2,531	0,439	2,628	0,558	2,238	1,582
		300	2,581	0,267	2,631	0,248	2,630	0,319	2,495	1,358
	0,5	50	1,047	0,099	1,168	0,276	1,174	0,271	1,102	0,581
		100	1,055	0,132	1,082	0,195	1,099	0,257	1,126	0,522
		300	1,032	0,084	1,080	0,108	1,071	0,124	1,064	0,428
	0,7	50	0,389	0,074	0,460	0,156	0,471	0,125	0,457	0,233
		100	0,407	0,061	0,470	0,091	0,471	0,091	0,433	0,189
		300	0,405	0,059	0,466	0,058	0,462	0,067	0,441	0,077
	0,9	50	0,076	0,054	0,099	0,058	0,109	0,070	0,101	0,063
		100	0,070	0,042	0,119	0,064	0,126	0,068	0,140	0,141
		300	0,043	0,037	0,129	0,081	0,133	0,080	0,125	0,041
0,3	0,1	50	13,298	3,817	13,393	2,738	13,409	4,209	13,185	7,949
		100	12,251	2,204	12,554	1,645	12,347	2,530	10,666	6,361
		300	12,221	2,471	12,552	1,786	12,191	2,711	12,270	8,920
	0,3	50	3,321	0,959	3,580	0,889	3,589	1,258	3,408	2,675
		100	3,015	0,709	3,057	0,626	2,982	0,854	2,889	1,844
		300	3,044	0,400	3,145	0,390	3,066	0,470	2,519	1,381
	0,5	50	1,388	0,471	1,529	0,432	1,505	0,485	1,562	1,117
		100	1,319	0,486	1,394	0,386	1,361	0,505	1,328	0,578
		300	1,350	0,301	1,447	0,321	1,435	0,375	1,271	0,752
	0,7	50	0,555	0,202	0,630	0,229	0,632	0,241	0,598	0,222
		100	0,596	0,257	0,677	0,339	0,671	0,355	0,621	0,303
		300	0,589	0,221	0,636	0,243	0,645	0,266	0,581	0,296
	0,9	50	0,053	0,038	0,189	0,189	0,189	0,187	0,116	0,067
		100	0,059	0,040	0,151	0,130	0,151	0,131	0,148	0,089
		300	0,034	0,028	0,146	0,103	0,146	0,105	0,144	0,073

Tabela B2.b - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidas em 20 replicações do processo INAR(1)-Geométrica (ϕ=*0,5 e* ϕ=*0,7*),

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
ϕ	θ	n	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$
0,5	0,1	50	18,956	5,553	19,150	5,140	19,228	6,793	19,606	7,763
		100	18,636	6,397	18,376	4,765	18,658	6,602	18,170	11,620
		300	17,425	6,401	17,724	4,389	17,437	6,522	19,185	8,300
	0,3	50	4,451	1,451	4,484	1,508	4,399	1,711	4,489	2,188
		100	4,332	1,268	4,369	1,003	4,305	1,308	3,590	1,661
		300	4,067	1,244	4,246	0,960	4,077	1,265	3,721	1,766
	0,5	50	2,208	0,830	2,450	0,855	2,406	1,005	2,427	0,960
		100	2,054	1,313	2,134	1,313	2,127	1,471	2,071	1,763
		300	1,946	0,844	1,956	0,768	1,942	0,900	2,173	1,368
	0,7	50	0,843	0,624	0,892	0,672	0,895	0,695	0,958	0,799
		100	0,829	0,506	0,860	0,556	0,863	0,577	0,831	0,433
		300	1,004	0,670	1,066	0,737	1,080	0,774	0,953	0,842
	0,9	50	0,055	0,044	0,222	0,234	0,222	0,229	0,190	0,139
		100	0,114	0,142	0,275	0,237	0,274	0,236	0,268	0,161
		300	0,141	0,286	0,278	0,351	0,277	0,350	0,211	0,170
0,7	0,1	50	31,016	11,265	31,619	8,004	31,438	11,822	31,369	14,279
		100	34,576	14,551	32,906	10,698	34,487	15,170	38,731	19,049
		300	32,878	11,132	31,456	7,682	32,697	11,302	31,097	12,552
	0,3	50	7,325	2,697	7,294	2,192	7,293	2,823	7,122	3,405
		100	7,739	2,921	7,772	2,536	7,734	3,061	7,794	3,030
		300	7,221	1,966	7,331	1,501	7,216	1,955	7,614	1,757
	0,5	50	3,666	1,917	3,811	2,000	3,831	2,190	3,576	1,972
		100	3,396	1,784	3,488	1,622	3,447	1,806	3,778	1,998
		300	3,159	1,370	3,153	1,258	3,150	1,397	3,280	1,795
	0,7	50	1,168	0,688	1,151	0,654	1,149	0,666	1,365	0,891
		100	1,632	0,850	1,665	0,890	1,668	0,911	1,334	0,804
		300	1,408	0,907	1,438	0,928	1,439	0,955	1,483	1,281
	0,9	50	0,231	0,296	0,276	0,329	0,272	0,328	0,393	0,280
		100	0,169	0,279	0,305	0,371	0,304	0,365	0,239	0,228
		300	0,256	0,314	0,429	0,594	0,429	0,593	0,423	0,457

Tabela B2.c - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidas em 20 replicações do processo INAR(1)-Geométrica (ϕ=*0,9*)

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
ϕ	θ	n	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$
0,9	0,1	50	--	---	--	-----	--	-----	--	-----
		100	--	---	--	-----	--	-----	--	-----
		300	--	---	--	-----	--	-----	--	-----
	0,3	50	22,555	8,555	22,755	8,350	22,630	8,984	22,576	7,467
		100	25,770	8,673	25,746	8,016	25,911	8,848	26,435	8,574
		300	23,032	5,391	22,992	4,959	22,994	5,410	22,793	5,398
	0,5	50	10,106	3,340	10,165	3,408	10,137	3,482	10,325	3,477
		100	9,544	2,904	9,601	2,833	9,576	2,945	9,942	2,838
		300	10,583	2,892	10,572	2,815	10,592	2,928	10,447	2,986
	0,7	50	4,457	2,002	4,551	2,105	4,548	2,123	4,471	1,820
		100	3,978	2,128	3,970	2,109	3,967	2,130	3,780	1,993
		300	4,649	2,229	4,652	2,227	4,655	2,248	4,450	2,234
	0,9	50	0,938	0,847	0,981	0,887	0,965	0,904	1,106	0,818
		100	0,947	0,928	0,958	0,973	0,958	0,973	1,113	0,923
		300	1,246	0,923	1,292	0,991	1,292	0,991	1,234	0,907

Tabela B3.a - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica ($\phi = 0,1$ e $\phi=0,3$).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
ϕ	θ	n	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$
0,1	0,1	50	87,583	3,487	10,250	1,958	93,040	26,643	97,260	102,760
		100	88,275	1,305	10,280	1,098	88,030	12,360	75,210	60,350
		300	88,174	1,296	10,080	0,858	89,760	12,813	75,080	54,050
	0,3	50	7,659	0,209	2,420	0,353	7,600	2,100	4,970	3,290
		100	7,739	0,287	2,510	0,427	7,800	1,910	4,480	3,230
		300	7,751	0,231	2,610	0,232	8,110	0,835	5,210	2,830
	0,5	50	1,925	0,110	1,160	0,273	2,380	0,789	1,420	0,790
		100	1,915	0,134	1,070	0,189	2,110	0,613	1,330	0,770
		300	1,898	0,076	1,070	0,102	2,060	0,255	1,260	0,550
	0,7	50	0,524	0,072	0,450	0,146	0,690	0,247	0,490	0,260
		100	0,527	0,056	0,470	0,085	0,670	0,135	0,470	0,220
		300	0,521	0,047	0,460	0,055	0,650	0,093	0,450	0,080
	0,9	50	0,093	0,046	0,100	0,058	0,120	0,084	0,100	0,060
		100	0,087	0,035	0,120	0,063	0,140	0,077	0,150	0,150
		300	0,069	0,038	0,130	0,070	0,140	0,074	0,130	0,040
0,3	0,1	50	90,166	3,068	13,100	2,412	100,260	29,657	92,360	64,680
		100	90,177	1,533	12,280	1,493	93,050	15,657	73,760	47,070
		300	90,131	1,662	12,170	1,433	89,170	7,059	80,630	58,600
	0,3	50	8,265	0,662	3,430	0,764	9,560	2,770	7,610	6,670
		100	8,013	0,495	2,960	0,579	7,860	2,209	5,560	3,990
		300	8,075	0,268	3,040	0,342	8,380	1,209	5,410	3,130
	0,5	50	2,166	0,322	1,430	0,333	2,660	0,722	2,240	1,720
		100	2,107	0,339	1,330	0,305	2,240	0,470	1,670	0,780
		300	2,143	0,205	1,360	0,235	2,390	0,395	1,780	1,080
	0,7	50	0,643	0,131	0,570	0,184	0,750	0,258	0,660	0,270
		100	0,654	0,178	0,600	0,231	0,760	0,269	0,680	0,360
		300	0,652	0,152	0,580	0,184	0,760	0,214	0,630	0,330
	0,9	50	0,085	0,049	0,150	0,102	0,160	0,106	0,120	0,070
		100	0,082	0,042	0,140	0,090	0,150	0,096	0,150	0,090
		300	0,067	0,045	0,140	0,079	0,150	0,084	0,150	0,070

Tabela B3.b - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica ($\phi = 0,5$ e $\phi=0,7$).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
$\varepsilon_t \sim$ geométrica(θ)										
ϕ	θ	n	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$
0.5	0.1	50	92,659	3,632	17,360	4,153	102,790	37,741	101,250	57,230
		100	92,891	2,399	16,210	3,178	90,450	14,323	85,480	60,480
		300	92,840	2,812	16,030	3,104	94,760	12,855	85,860	44,320
	0.3	50	8,664	0,754	3,920	1,154	9,080	3,032	8,810	5,040
		100	8,624	0,605	3,890	0,705	8,930	1,952	7,040	3,170
		300	8,504	0,599	3,790	0,637	8,590	1,214	7,500	3,770
	0.5	50	2,512	0,429	1,880	0,445	2,740	0,844	3,470	1,780
		100	2,643	1,469	1,690	0,701	2,500	0,750	3,120	3,100
		300	2,385	0,415	1,610	0,452	2,440	0,581	3,240	2,200
	0.7	50	0,764	0,326	0,710	0,398	0,940	0,500	1,120	1,070
		100	0,755	0,267	0,670	0,308	0,850	0,351	1,020	0,550
		300	0,898	0,435	0,790	0,391	0,960	0,399	1,160	1,090
	0.9	50	0,095	0,077	0,160	0,105	0,170	0,110	0,200	0,150
		100	0,124	0,091	0,200	0,118	0,220	0,122	0,290	0,180
		300	0,174	0,329	0,200	0,181	0,210	0,180	0,220	0,180
0.7	0.1	50	95,358	3,472	26,590	5,349	107,800	24,669	115,890	77,590
		100	96,527	4,072	25,230	6,169	96,650	22,738	121,280	77,240
		300	96,317	2,861	24,380	4,772	95,740	14,864	92,070	55,330
	0.3	50	9,420	1,310	5,070	1,150	8,810	1,742	12,820	6,420
		100	9,355	0,921	5,350	1,274	9,490	2,716	14,330	6,680
		300	9,198	0,603	5,290	0,806	9,270	1,349	13,670	3,380
	0.5	50	2,851	0,757	2,240	0,733	2,920	0,947	4,960	3,230
		100	2,714	0,599	2,250	0,625	3,160	0,766	5,790	3,210
		300	2,616	0,408	1,990	0,535	2,690	0,561	4,940	2,840
	0.7	50	0,803	0,208	0,720	0,277	0,890	0,308	1,800	1,190
		100	0,973	0,326	0,850	0,296	1,010	0,314	1,760	1,080
		300	0,926	0,389	0,770	0,313	0,920	0,310	1,970	1,630
	0.9	50	0,128	0,095	0,160	0,129	0,170	0,146	0,440	0,330
		100	0,178	0,254	0,170	0,132	0,190	0,137	0,270	0,270
		300	0,206	0,257	0,200	0,166	0,220	0,166	0,490	0,530

Tabela B3.c - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica ($\phi = 0,9$)

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
$\varepsilon_t \sim$ geométrica(θ)										
ϕ	θ	n	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\bar{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$
0.9	0.1	50	----	----	----	----	----	----	----	----
		100	----	----	----	----	----	----	----	----
		300	----	----	----	----	----	----	----	----
	0.3	50	10,454	2,767	8,240	2,703	10,070	2,782	31,140	10,620
		100	11,235	4,011	9,030	2,416	11,130	3,063	37,520	13,040
		300	9,855	0,555	7,820	1,133	9,930	1,171	32,780	8,370
	0.5	50	3,104	0,734	2,730	1,107	3,210	1,207	13,480	5,200
		100	2,889	0,369	2,610	0,805	3,030	0,819	13,310	3,390
		300	2,968	0,313	2,600	0,476	3,000	0,454	13,860	3,830
	0.7	50	1,229	0,852	0,940	0,327	1,070	0,371	5,330	2,400
		100	0,979	0,239	0,860	0,308	1,010	0,326	4,900	2,570
		300	1,074	0,332	0,930	0,270	1,060	0,271	5,800	3,000
	0.9	50	0,197	0,094	0,190	0,114	0,190	0,139	1,300	0,950
		100	0,212	0,120	0,190	0,125	0,200	0,130	1,350	1,120
		300	0,276	0,187	0,220	0,089	0,230	0,089	1,510	1,080

Tabela B4.a - Erro quadrático da previsão 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=*0,1 e* ϕ=*0,3*),

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
ϕ	θ	n	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$
0,1	0,1	50	141,595	342,640	156,810	379,150	159,980	378,300	238,280	469,510
		100	103,086	182,720	101,940	201,920	98,780	184,580	174,290	294,180
		300	45,496	37,130	46,840	38,410	47,530	40,440	54,810	88,900
	0,3	50	6,359	9,790	6,030	9,450	6,170	9,730	9,080	14,980
		100	6,347	5,280	5,580	4,810	5,960	4,880	11,450	11,820
		300	10,581	19,530	9,980	17,930	9,920	17,770	9,720	16,100
	0,5	50	1,086	1,810	1,190	1,740	1,200	1,750	1,030	1,510
		100	0,828	0,570	0,850	0,610	0,910	0,730	1,140	1,220
		300	1,701	2,480	1,750	2,430	1,730	2,430	1,940	2,910
	0,7	50	0,347	0,560	0,400	0,620	0,380	0,510	0,450	0,860
		100	0,778	2,630	0,800	2,570	0,780	2,490	0,740	2,520
		300	0,857	2,900	0,880	2,850	0,890	2,890	0,850	2,840
	0,9	50	0,008	0,010	0,010	0,010	0,020	0,020	0,010	0,020
		100	0,246	0,870	0,230	0,770	0,230	0,760	0,250	0,790
		300	0,090	0,270	0,100	0,240	0,100	0,230	0,090	0,230
0,3	0,1	50	80,882	73,540	60,560	54,140	76,620	70,420	112,750	131,980
		100	75,357	137,750	71,810	130,240	77,320	140,680	69,280	103,740
		300	124,010	202,110	122,360	204,760	126,870	209,010	113,330	192,230
	0,3	50	4,575	6,300	4,450	5,060	4,130	5,190	6,410	7,410
		100	9,599	18,520	9,460	17,280	9,560	18,450	9,340	20,380
		300	9,024	10,730	9,290	11,390	9,480	12,160	10,690	18,390
	0,5	50	0,740	1,020	0,760	0,980	0,810	1,120	0,910	1,240
		100	1,082	3,030	1,080	2,410	1,090	2,630	1,060	2,450
		300	7,755	20,890	7,590	20,370	7,640	20,540	6,760	17,840
	0,7	50	0,758	1,420	0,730	1,310	0,730	1,280	0,700	1,470
		100	0,507	0,500	0,630	0,750	0,650	0,780	0,570	0,710
		300	1,834	4,780	1,860	4,820	1,870	4,830	1,630	3,930
	0,9	50	0,172	0,350	0,220	0,320	0,210	0,320	0,170	0,320
		100	0,133	0,310	0,150	0,280	0,150	0,280	0,130	0,250
		300	0,288	0,900	0,250	0,810	0,250	0,810	0,250	0,740

Tabela B4.b - Erro quadrático da previsão 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,5 e ϕ=0,7).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica(θ)			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
ϕ	θ	n	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$
0,5	0,1	50	57,183	65,450	58,660	68,220	68,220	80,520	50,330	59,670
		100	295,146	634,270	293,250	609,190	297,710	625,540	281,200	741,110
		300	175,228	352,440	152,360	293,640	178,200	354,510	97,210	263,500
	0,3	50	15,542	30,480	16,460	33,690	15,480	30,110	12,930	28,350
		100	9,870	16,900	9,050	14,790	9,570	15,030	10,670	24,470
		300	5,047	5,870	5,170	6,480	4,980	6,460	2,420	3,410
	0,5	50	4,203	3,860	3,970	3,940	4,320	3,990	3,030	3,460
		100	5,263	12,330	4,990	11,430	5,330	12,050	5,730	11,500
		300	1,906	1,870	1,870	1,730	1,930	1,860	2,280	2,830
	0,7	50	0,954	1,900	1,030	2,320	1,070	2,410	1,200	3,550
		100	1,055	1,260	1,120	1,260	1,140	1,290	1,140	1,350
		300	1,093	1,200	1,140	1,160	1,150	1,150	1,320	2,200
	0,9	50	0,168	0,340	0,090	0,200	0,090	0,200	0,120	0,210
		100	0,108	0,240	0,100	0,150	0,100	0,150	0,170	0,210
		300	0,256	0,450	0,310	0,560	0,310	0,560	0,150	0,250
0,7	0,1	50	154,522	250,610	146,340	240,910	152,660	225,410	164,320	327,840
		100	238,336	516,230	213,720	532,160	244,980	514,000	92,840	158,660
		300	111,907	124,550	89,350	88,150	113,390	124,660	84,270	107,660
	0,3	50	10,453	19,550	10,270	21,040	10,390	20,790	6,210	13,420
		100	11,549	13,820	11,280	12,550	12,200	14,000	8,290	10,080
		300	15,742	27,290	13,550	23,440	15,170	25,830	10,020	14,210
	0,5	50	4,924	9,580	4,300	7,100	4,880	8,720	5,210	8,020
		100	5,358	5,460	5,760	5,390	5,580	5,380	5,550	8,230
		300	3,057	3,140	3,060	3,360	3,030	3,190	2,060	2,330
	0,7	50	1,142	2,590	1,260	3,100	1,240	3,000	0,910	1,590
		100	1,197	2,340	1,260	2,410	1,280	2,430	0,980	1,680
		300	1,911	2,410	1,850	2,330	1,820	2,260	1,230	1,930
	0,9	50	0,207	0,320	0,230	0,350	0,230	0,340	0,260	0,290
		100	0,198	0,340	0,170	0,300	0,170	0,290	0,210	0,310
		300	0,433	1,330	0,310	0,620	0,310	0,620	0,450	0,870

Tabela B4.c - Erro quadrático da previsão 1-passo-à-frente - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,9).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMÉTRICA		POISSON-GAMA	
ϕ	θ	n	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$
0,9	0,1	50	----	----	----	----	----	----	----	----
		100	----	----	----	----	----	----	----	----
		300	----	----	----	----	----	----	----	----
	0,3	50	13,268	15,560	12,210	13,340	14,110	17,790	13,020	8,300
		100	18,884	21,240	16,560	19,480	18,770	22,290	12,700	15,000
		300	16,878	27,850	15,010	24,990	16,480	27,250	9,780	18,620
	0,5	50	5,709	9,230	6,030	9,480	5,960	9,260	5,850	7,240
		100	11,482	18,490	11,580	18,880	11,550	18,710	12,110	17,290
		300	9,274	18,150	9,310	18,150	9,180	18,050	11,050	23,410
	0,7	50	3,984	10,990	3,950	11,310	3,940	11,250	4,240	7,420
		100	2,040	2,300	1,980	2,210	1,970	2,230	2,290	3,910
		300	1,074	1,730	1,050	1,680	1,070	1,710	1,090	1,660
	0,9	50	0,511	1,040	0,530	1,080	0,540	1,080	0,620	1,170
		100	0,587	1,040	0,620	1,120	0,620	1,120	0,730	0,750
		300	0,163	0,320	0,190	0,370	0,190	0,370	0,310	0,420

Tabela B5.a - Divergência de kullback leibler entre a distribuição preditiva real e a distribuição preditiva do modelo ajustado, - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,1 e ϕ=0,3)

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\overline{D}(P:Q)$	$S_{D(P:Q)}$	$\overline{D}(P:Q)$	$S_{D(P:Q)}$	$\overline{D}(P:Q)$	$S_{D(P:Q)}$
0.1	0.1	50	0,695	0,088	0,028	0,056	0,515	0,592
		100	0,661	0,128	0,008	0,008	0,718	0,555
		300	0,682	0,077	0,003	0,004	0,501	0,574
	0.3	50	0,216	0,055	0,033	0,023	0,165	0,109
		100	0,207	0,052	0,026	0,011	0,260	0,220
		300	0,215	0,042	0,018	0,008	0,169	0,171
	0.5	50	0,103	0,045	0,103	0,058	0,147	0,122
		100	0,095	0,037	0,120	0,073	0,145	0,122
		300	0,081	0,014	0,081	0,023	0,110	0,081
	0.7	50	0,666	0,683	0,846	0,567	0,843	0,784
		100	0,254	0,282	0,446	0,300	0,281	0,296
		300	0,345	0,278	0,546	0,259	0,346	0,320
	0.9	50	60,400	49,828	60,600	49,797	60,332	49,841
		100	49,657	49,070	49,501	49,037	49,871	48,878
		300	32,299	38,625	31,546	38,869	31,793	38,298
0.3	0.1	50	0,612	0,151	0,017	0,019	0,437	0,419
		100	0,614	0,145	0,006	0,005	0,460	0,366
		300	0,621	0,130	0,001	0,001	0,640	0,778
	0.3	50	0,202	0,101	0,037	0,053	0,271	0,277
		100	0,193	0,072	0,022	0,015	0,234	0,228
		300	0,182	0,047	0,013	0,007	0,198	0,236
	0.5	50	0,103	0,052	0,089	0,056	0,233	0,353
		100	0,090	0,038	0,080	0,062	0,111	0,078
		300	0,065	0,022	0,052	0,019	0,147	0,122
	0.7	50	0,709	1,036	0,792	0,866	0,818	1,090
		100	0,573	0,611	0,711	0,512	0,702	0,750
		300	0,319	0,383	0,468	0,381	0,410	0,364
	0.9	50	32,775	37,776	32,131	37,976	25,427	36,470
		100	32,562	40,835	31,966	41,159	31,441	39,912
		300	19,016	22,240	18,158	22,603	18,214	21,040

Tabela B5.b - Divergência de Kullback Leibler entre a distribuição preditiva real e a distribuição preditiva do modelo ajustado, - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,5 e ϕ=0,7)

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\bar{D}(P:Q)$	$S_{D(P:Q)}$	$\bar{D}(P:Q)$	$S_{D(P:Q)}$	$\bar{D}(P:Q)$	$S_{D(P:Q)}$
0,5	0,1	50	0,464	0,172	0,026	0,024	0,439	0,288
		100	0,481	0,174	0,007	0,011	0,618	0,793
		300	0,539	0,181	0,003	0,003	0,487	0,483
	0,3	50	0,164	0,089	0,038	0,054	0,151	0,142
		100	0,150	0,086	0,016	0,013	0,253	0,260
		300	0,165	0,094	0,008	0,005	0,167	0,134
	0,5	50	0,108	0,130	0,101	0,073	0,277	0,250
		100	0,118	0,194	0,107	0,236	0,219	0,333
		300	0,057	0,034	0,048	0,046	0,284	0,258
	0,7	50	0,338	0,664	0,448	0,578	0,587	0,509
		100	0,187	0,284	0,293	0,291	0,476	0,435
		300	0,440	0,766	0,538	0,770	0,663	0,787
	0,9	50	15,731	22,552	15,271	22,621	14,886	22,705
		100	7,548	6,522	7,403	6,160	8,632	6,252
		300	15,996	23,586	15,494	23,685	15,355	23,411
0,7	0,1	50	0,500	0,258	0,018	0,021	0,660	0,774
		100	0,437	0,259	0,012	0,012	0,897	1,016
		300	0,397	0,199	0,004	0,006	0,611	0,496
	0,3	50	0,143	0,108	0,023	0,030	0,344	0,204
		100	0,121	0,091	0,016	0,013	0,306	0,197
		300	0,114	0,062	0,006	0,008	0,258	0,161
	0,5	50	0,118	0,121	0,115	0,153	0,368	0,325
		100	0,076	0,055	0,038	0,041	0,436	0,302
		300	0,046	0,029	0,015	0,014	0,313	0,211
	0,7	50	0,107	0,180	0,152	0,213	0,699	0,584
		100	0,173	0,326	0,233	0,343	0,725	0,545
		300	0,118	0,335	0,135	0,335	0,701	0,949
	0,9	50	3,650	3,938	3,258	3,269	8,274	6,380
		100	14,513	26,199	14,444	25,640	9,451	10,791
		300	7,541	16,387	7,551	16,221	6,445	7,355

Tabela B5.c - Divergência de Kullback Leibler entre a distribuição preditiva real e a distribuição preditiva do modelo ajustado, - média e desvio padrão obtidos em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,9)

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\bar{D}(P:Q)$	$S_{D(P:Q)}$	$\bar{D}(P:Q)$	$S_{D(P:Q)}$	$\bar{D}(P:Q)$	$S_{D(P:Q)}$
0,9	0,1	50	----	----	----	----	----	----
		100	----	----	----	----	----	----
		300	----	----	----	----	----	----
	0,3	50	0,104	0,036	0,023	0,032	0,994	0,602
		100	0,111	0,084	0,016	0,026	1,095	0,590
		300	0,082	0,029	0,002	0,002	1,122	0,645
	0,5	50	0,071	0,049	0,039	0,043	1,067	0,428
		100	0,056	0,038	0,013	0,013	1,016	0,425
		300	0,036	0,017	0,004	0,004	1,061	0,289
	0,7	50	0,216	0,698	0,216	0,592	1,603	0,931
		100	0,028	0,020	0,024	0,041	1,119	0,616
		300	0,024	0,025	0,025	0,061	1,251	0,663
	0,9	50	0,583	1,035	0,579	1,003	5,941	5,677
		100	0,286	0,641	0,369	0,659	3,849	3,209
		300	0,896	2,469	0,884	2,257	4,016	4,916

Tabela B6.a - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,1 e ϕ=0,3).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}
0,1	0,1	50	572,390	86,060	332,920	16,890	375,170	21,950
		100	1127,510	65,560	663,140	15,250	747,820	17,310
		300	3426,740	278,800	1990,720	42,210	2260,590	58,990
	0,3	50	244,830	23,250	209,700	11,300	232,060	15,380
		100	487,880	40,210	416,240	22,940	460,810	26,030
		300	1502,500	60,900	1268,070	29,650	1412,120	35,770
	0,5	50	162,630	18,480	151,230	14,670	162,030	16,930
		100	307,880	36,990	288,850	24,630	305,400	31,920
		300	935,950	40,650	867,090	31,660	935,350	40,750
	0,7	50	93,530	14,380	94,200	12,930	95,910	13,670
		100	194,510	16,790	189,220	15,410	196,740	17,740
		300	568,140	38,860	554,670	31,360	573,770	38,440
	0,9	50	36,720	13,360	37,210	14,230	39,990	14,730
		100	79,320	26,640	79,980	25,180	84,010	26,110
		300	231,170	36,730	231,390	37,720	238,690	37,980
0,3	0,1	50	536,720	79,490	342,650	15,670	371,740	14,060
		100	1084,130	82,090	678,590	15,410	737,200	20,800
		300	3209,810	171,600	2026,350	20,320	2193,870	28,740
	0,3	50	260,270	26,570	227,180	12,990	242,670	17,770
		100	490,070	43,620	437,300	23,680	464,660	27,860
		300	1527,140	82,750	1332,740	43,000	1430,640	56,520
	0,5	50	174,620	17,670	163,350	13,120	173,810	15,740
		100	328,730	25,170	313,420	20,050	330,820	18,850
		300	1022,200	52,690	956,850	41,780	1023,300	44,130
	0,7	50	101,480	17,210	99,550	16,020	106,080	18,140
		100	200,410	23,180	197,130	22,930	208,360	22,690
		300	612,190	43,040	601,020	40,160	635,080	43,700
	0,9	50	42,250	11,480	42,170	11,240	47,530	12,810
		100	87,130	23,970	87,240	23,850	97,640	28,200
		300	252,900	40,230	252,340	39,140	274,360	46,520

Tabela B6.b - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=*0,5 e* ϕ=*0,7).*

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}
0.5	0.1	50	491,080	83,410	345,950	18,780	364,550	21,960
		100	977,270	78,410	684,580	15,320	720,030	20,140
		300	3039,070	208,220	2065,590	35,810	2232,080	187,340
	0.3	50	252,020	27,010	230,510	15,950	241,830	20,180
		100	515,960	38,720	461,280	19,840	487,510	21,970
		300	1528,890	64,430	1384,930	39,400	1469,490	41,570
	0.5	50	175,580	18,140	170,530	14,720	181,290	16,380
		100	345,410	17,470	332,440	12,970	358,110	17,300
		300	1036,240	43,710	998,230	35,890	1073,440	36,680
	0.7	50	116,980	15,410	116,250	13,750	125,570	14,520
		100	225,220	27,310	221,710	25,550	243,590	27,150
		300	677,290	40,390	663,700	37,090	734,620	40,990
	0.9	50	49,570	11,730	49,710	11,680	59,170	14,020
		100	103,540	20,640	103,480	20,700	122,630	22,190
		300	272,750	46,100	271,910	45,960	320,120	51,150
0.7	0.1	50	449,380	49,020	353,290	10,800	373,480	14,520
		100	897,150	82,080	694,280	20,260	740,100	46,450
		300	2755,520	107,520	2077,430	43,820	2299,540	181,250
	0.3	50	252,160	17,120	236,520	11,200	254,710	10,970
		100	506,420	37,800	475,620	23,990	511,700	23,330
		300	1540,020	65,140	1430,980	38,730	1544,200	38,970
	0.5	50	183,170	17,550	180,650	15,850	198,700	16,750
		100	384,720	26,770	367,840	23,220	407,190	23,540
		300	1103,080	36,580	1065,580	32,050	1187,250	27,990
	0.7	50	124,830	13,340	121,650	12,730	140,940	12,240
		100	244,980	22,870	242,030	22,300	286,590	25,950
		300	740,350	46,620	728,080	43,580	866,810	41,960
	0.9	50	53,130	15,010	70,310	78,970	70,870	19,530
		100	101,860	28,410	101,860	28,550	135,940	33,140
		300	314,320	53,610	313,380	53,550	430,220	70,180

Tabela B6,c - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=*0,9)*

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}
0.9	0.1	50	----	----	----	----	----	----
		100	----	----	----	----	----	----
		300	----	----	----	----	----	----
	0.3	50	255,530	13,060	245,790	10,630	288,120	10,250
		100	523,240	33,410	496,210	23,160	586,810	21,890
		300	1537,140	44,080	1470,510	30,380	1738,320	36,720
	0.5	50	192,860	21,200	188,640	19,410	238,580	13,840
		100	386,790	29,010	379,120	26,710	480,940	18,760
		300	1146,760	38,460	1119,000	34,620	1432,420	28,650
	0.7	50	136,680	15,010	135,540	14,800	191,970	11,930
		100	271,280	23,930	267,950	22,120	380,800	20,920
		300	805,860	41,910	795,700	39,810	1142,460	45,490
	0.9	50	62,640	17,570	104,530	133,090	115,660	28,960
		100	116,180	31,080	115,880	30,870	215,030	49,100
		300	367,920	48,780	367,080	48,660	688,590	78,540

Tabela B7.a - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,1 e ϕ=0,3).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}
0.1	0.1	50	576,250	86,060	336,780	16,890	377,100	21,950
		100	1132,740	65,560	668,370	15,250	750,430	17,310
		300	3434,150	278,800	1998,130	42,210	2264,290	58,990
	0.3	50	248,690	23,250	213,560	11,300	233,990	15,380
		100	493,110	40,210	421,470	22,940	463,430	26,030
		300	1509,910	60,900	1275,490	29,650	1415,830	35,770
	0.5	50	166,490	18,480	155,090	14,670	163,960	16,930
		100	313,110	36,990	294,080	24,630	308,020	31,920
		300	943,370	40,650	874,510	31,660	939,060	40,750
	0.7	50	97,390	14,380	98,060	12,930	97,840	13,670
		100	199,740	16,790	194,450	15,410	199,350	17,740
		300	575,550	38,860	562,080	31,360	577,480	38,440
	0.9	50	40,580	13,360	41,070	14,230	41,920	14,730
		100	84,550	26,640	85,210	25,180	86,630	26,110
		300	238,580	36,730	238,800	37,720	242,400	37,980
0.3	0.1	50	540,580	79,490	346,520	15,670	373,670	14,060
		100	1089,360	82,090	683,820	15,410	739,810	20,800
		300	3217,220	171,600	2033,760	20,320	2197,570	28,740
	0.3	50	264,130	26,570	231,040	12,990	244,600	17,770
		100	495,300	43,620	442,530	23,680	467,280	27,860
		300	1534,560	82,750	1340,160	43,000	1434,350	56,520
	0.5	50	178,490	17,670	167,220	13,120	175,740	15,740
		100	333,960	25,170	318,650	20,050	333,430	18,850
		300	1029,610	52,690	964,270	41,780	1027,010	44,130
	0.7	50	105,340	17,210	103,410	16,020	108,010	18,140
		100	205,640	23,180	202,360	22,930	210,970	22,690
		300	619,600	43,040	608,430	40,160	638,790	43,700
	0.9	50	46,110	11,480	46,040	11,240	49,460	12,810
		100	92,360	23,970	92,470	23,850	100,250	28,200
		300	260,320	40,230	259,750	39,140	278,070	46,520

Tabela B7.b - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,5 e ϕ=0,7).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}
0.5	0.1	50	494,94	83,41	349,81	18,78	366,48	21,96
		100	982,50	78,41	689,81	15,32	722,65	20,14
		300	3046,48	208,22	2073,00	35,81	2235,78	187,34
	0.3	50	255,88	27,01	234,37	15,95	243,76	20,18
		100	521,19	38,72	466,51	19,84	490,13	21,97
		300	1536,30	64,43	1392,35	39,40	1473,19	41,57
	0.5	50	179,45	18,14	174,39	14,72	183,23	16,38
		100	350,64	17,47	337,67	12,97	360,72	17,30
		300	1043,65	43,71	1005,64	35,89	1077,14	36,68
	0.7	50	120,85	15,41	120,11	13,75	127,50	14,52
		100	230,45	27,31	226,94	25,55	246,21	27,15
		300	684,70	40,39	671,12	37,09	738,33	40,99
	0.9	50	53,43	11,73	53,57	11,68	61,10	14,02
		100	108,77	20,64	108,71	20,70	125,24	22,19
		300	280,17	46,10	279,33	45,96	323,82	51,15
0.7	0.1	50	453,24	49,02	357,16	10,80	375,41	14,52
		100	902,38	82,08	699,51	20,26	742,72	46,45
		300	2762,94	107,52	2084,84	43,82	2303,24	181,25
	0.3	50	256,03	17,12	240,38	11,20	256,64	10,97
		100	511,65	37,80	480,85	23,99	514,31	23,33
		300	1547,44	65,14	1438,39	38,73	1547,90	38,97
	0.5	50	187,03	17,55	184,52	15,85	200,63	16,75
		100	389,95	26,77	373,07	23,22	409,81	23,54
		300	1110,49	36,58	1072,99	32,05	1190,96	27,99
	0.7	50	128,70	13,34	125,51	12,73	142,88	12,24
		100	250,21	22,87	247,26	22,30	289,20	25,95
		300	747,76	46,62	735,50	43,58	870,52	41,96
	0.9	50	57,00	15,01	74,18	78,97	72,81	19,53
		100	107,09	28,41	107,09	28,55	138,56	33,14
		300	321,73	53,61	320,79	53,55	433,93	70,18

Tabela B7.c - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo INAR(1)-Geométrica (ϕ=0,9).

$Z_t = \phi \circ Z_{t-1} + \varepsilon_t$ $\varepsilon_t \sim$ geométrica			INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ϕ	θ	n	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}
0.9	0.1	50	----	----	----	----	----	----
		100	----	----	----	----	----	----
		300	----	----	----	----	----	----
	0.3	50	259,39	13,06	249,65	10,63	290,05	10,25
		100	528,47	33,41	501,44	23,16	589,43	21,89
		300	1544,56	44,08	1477,92	30,38	1742,03	36,72
	0.5	50	196,73	21,20	192,50	19,41	240,51	13,84
		100	392,02	29,01	384,35	26,71	483,56	18,76
		300	1154,18	38,46	1126,41	34,62	1436,13	28,65
	0.7	50	140,55	15,01	139,40	14,80	193,90	11,93
		100	276,51	23,93	273,18	22,12	383,42	20,92
		300	813,27	41,91	803,11	39,81	1146,16	45,49
	0.9	50	66,50	17,57	108,39	133,09	117,59	28,96
		100	121,41	31,08	121,11	30,87	217,64	49,10
		300	375,34	48,78	374,50	48,66	692,30	78,54

APÊNDICE C

Resultados obtidos para a simulação do processo Poisson-Gama

Tabela C1 - Estimativas médias e desvios-padrões obtidos para os parâmetros no ajuste dos modelos : INAR(1)-Poisson, INAR(1)-Geométrica e Poisson-Gama em processos INAR(1)-Poisson simulados.

		INAR(1)-POISSON				INAR(1)-GEOMETRICA				POISSON-GAMA	
ω	n	$\bar{\hat{\alpha}}$	$S(\bar{\hat{\alpha}})$	$\bar{\hat{\lambda}}$	$S(\bar{\hat{\lambda}})$	$\bar{\hat{\phi}}$	$S(\bar{\hat{\phi}})$	$\bar{\hat{\theta}}$	$S(\bar{\hat{\theta}})$	$\bar{\hat{\omega}}$	$S(\bar{\hat{\omega}})$
0,8[1]	50	0,202	0,123	8,821	5,466	0,550	0,125	0,208	0,070	0,847	0,077
	100	--	--	--	--	--	--	--	--	--	--
	300	--	--	--	--	--	--	--	--	--	--
0,9	50	0,179	0,112	9,697	5,516	0,583	0,083	0,199	0,048	0,892	0,056
	100	0,131	0,087	8,981	3,332	0,566	0,092	0,201	0,059	0,924	0,054
	300	0,239	0,099	7,868	4,064	0,566	0,087	0,214	0,072	0,903	0,023
0,95	50	0,118	0,141	9,461	2,785	0,576	0,087	0,188	0,041	0,927	0,031
	100	0,083	0,070	9,288	3,010	0,589	0,052	0,209	0,044	0,970	0,031
	300	0,109	0,052	8,551	2,581	0,587	0,045	0,211	0,036	0,961	0,020
0,98	50	0,106	0,097	8,050	2,401	0,581	0,076	0,222	0,046	0,949	0,022
	100	0,078	0,109	8,874	3,083	0,581	0,075	0,212	0,042	0,991	0,025
	300	0,035	0,039	8,752	2,963	0,568	0,063	0,218	0,047	0,983	0,013

Tabela C2 - Média da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo Poisson-Gama

		SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ω	n	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$	$\bar{\hat{Z}}_{n+1}$	$S_{\hat{Z}_{n+1}}$
0,8	50	10,291	4,522	10,925	6,256	10,550	5,105	10,214	4,322
	100	--	--	--	--	--	--	--	--
	300	--	--	--	--	--	--	--	--
0,9	50	10,716	4,686	11,183	5,649	10,325	4,887	10,729	4,595
	100	10,266	3,615	10,279	3,831	10,205	4,338	10,011	3,381
	300	10,122	5,610	10,404	5,697	10,253	6,141	10,226	5,630
0,95	50	10,473	2,845	10,728	3,060	10,539	3,401	10,474	2,905
	100	9,721	2,994	10,047	3,135	9,611	3,208	9,806	3,209
	300	9,743	3,607	9,619	2,814	9,744	3,499	9,721	3,567
0,98	50	8,835	2,570	8,724	2,521	7,917	2,425	8,649	2,513
	100	9,514	2,875	9,576	2,862	10,262	3,063	9,412	2,881
	300	8,922	3,277	9,059	3,094	8,921	3,380	8,882	3,208

[1] A simulacoes de processos com hiperparametro igual a 0,8, foram absorvidas pelo estado 0 em series de tamanho n=100 e n=300.

Tabela C3 - Variância da distribuição preditiva 1-passo-à-frente - média e desvio padrão obtidos em 20 replicações do processo Poisson-Gama.

		SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ω	n	$\overline{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\overline{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\overline{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$	$\overline{S}^2_{Z_{n+1}}$	$S_{S^2_{Z_{n+1}}}$
0,8	50	12,864	5,653	10,384	6,060	26,684	16,466	12,219	5,391
	100	--	--	--	--	--	--	--	--
	300	--	--	--	--	--	--	--	--
0,9	50	11,913	5,209	10,863	5,694	27,424	19,917	12,179	5,570
	100	11,407	4,016	10,046	3,741	25,597	11,811	10,873	3,693
	300	11,246	6,233	9,704	5,208	24,870	14,833	11,396	6,373
0,95	50	11,067	3,006	10,362	2,846	27,983	11,795	11,375	3,281
	100	10,236	3,152	9,942	3,110	22,384	8,518	10,167	3,313
	300	10,256	3,797	9,475	2,769	21,642	8,643	10,110	3,703
0,98	50	9,115	2,652	8,606	2,508	19,406	7,378	9,190	2,776
	100	9,737	2,943	9,426	2,880	21,751	8,127	9,582	2,940
	300	9,104	3,344	9,035	3,082	20,566	8,112	9,027	3,203

Tabela C4 - Erro quadrático da previsão 1-passo-à-frente utilizando a média da distribuição preditiva 1-passo-à-frente - média e desvio padrão em 20 replicações do processo Poisson-Gama.

		SIMULADO		INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ω	n	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$	$EMQP_{n+1}$	$S_{EMQP_{n+1}}$
0,8	50	13,282	16,825	22,958	37,197	18,463	24,467	13,169	17,842
	100	--	--	--	--	--	--	--	--
	300	--	--	--	--	--	--	--	--
0,9	50	7,755	8,342	8,482	9,319	9,226	11,394	7,545	8,533
	100	11,831	15,280	12,553	16,023	13,116	18,777	12,910	20,450
	300	17,411	24,719	16,097	14,465	18,759	21,257	17,214	24,686
0,95	50	12,578	15,300	12,767	15,393	15,655	21,631	12,548	15,620
	100	10,785	15,030	10,842	17,855	11,022	15,115	10,538	14,815
	300	5,396	6,929	4,796	6,394	5,431	7,088	5,512	7,341
0,98	50	18,049	28,011	18,015	28,882	19,733	36,622	17,544	28,471
	100	9,273	10,993	9,600	11,062	14,599	17,602	9,019	10,507
	300	4,804	5,814	5,053	5,194	10,951	12,361	4,940	6,007

Tabela C5 - Divergência de Kullback-Leibler entre as distribuições preditivas 1-passo-à-frente do modelo ajustado e do processo simulado - média e desvio padrão obtidos em 20 replicações do processo Poisson-Gama.

		INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ω	n	$\overline{D}(P:Q)$	$S_{D(P:Q)}$	$\overline{D}(P:Q)$	$S_{D(P:Q)}$	$\overline{D}(P:Q)$	$S_{D(P:Q)}$
0,8	50	0,251	0,295	0,266	0,143	0,021	0,047
	100	--	--	--	--	--	--
	300	--	--	--	--	--	--
0,9	50	0,102	0,137	0,427	0,337	0,008	0,015
	100	0,067	0,127	0,400	0,304	0,014	0,032
	300	0,157	0,219	0,353	0,152	0,002	0,002
0,95	50	0,048	0,093	0,434	0,239	0,003	0,005
	100	0,029	0,035	0,348	0,350	0,013	0,029
	300	0,044	0,059	0,393	0,264	0,001	0,001
0,98	50	0,029	0,074	0,459	0,331	0,006	0,014
	100	0,009	0,015	0,421	0,273	0,003	0,009
	300	0,017	0,018	0,394	0,198	0,001	0,002

Tabela C6 - Resultados obtidos no AIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo Poisson-Gama.

		INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ω	n	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}	$\overline{AIC}$	S_{AIC}
0,8	50	275,89	23,458	288,13	22,359	270,64	18,406
	100	--	--	--	--	--	--
	300	--	--	--	--	--	--
0,9	50	267,30	24,028	289,52	26,532	268,30	23,872
	100	527,16	35,049	562,95	40,960	524,53	40,451
	300	1617,58	190,543	1652,26	187,742	1524,31	177,770
0,95	50	264,02	19,363	288,25	19,899	265,29	18,778
	100	512,23	37,118	561,53	41,024	515,20	35,938
	300	1532,01	96,204	1650,87	104,913	1519,99	88,362
0,98	50	258,04	17,000	281,83	21,819	259,24	15,890
	100	504,85	37,377	554,44	44,058	507,86	36,968
	300	1496,03	109,387	1628,42	130,481	1496,07	115,194

Tabela C7 - Resultados obtidos no BIC dos 3 modelos ajustados - média e desvio padrão em 20 replicações do processo Poisson-Gama.

		INAR(1)-POISSON		INAR(1)-GEOMETRICA		POISSON-GAMA	
ω	n	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}	$\overline{BIC}$	S_{BIC}
0,8	50	279,76	23,458	291,99	22,359	272,57	18,406
	100	--	--	--	--	--	--
	300	--	--	--	--	--	--
0,9	50	271,16	24,028	293,38	26,532	270,23	23,872
	100	532,39	35,049	568,18	40,960	527,14	40,451
	300	1624,99	190,543	1659,68	187,742	1528,02	177,770
0,95	50	267,88	19,363	292,12	19,899	267,22	18,778
	100	517,46	37,118	566,76	41,024	517,82	35,938
	300	1539,43	96,204	1658,29	104,913	1523,70	88,362
0,98	50	261,90	17,000	285,69	21,819	261,17	15,890
	100	510,08	37,377	559,67	44,058	510,48	36,968
	300	1503,45	109,387	1635,84	130,481	1499,77	115,194

Apêndice D

Programas em MATLAB

Programa D.1 – Simulação do processo INAR(1)-Poisson

```
%****Programa gerador da estrutura INAR(1) com a distribuição de Poisson
*****************************************************************
%****Estabelecendo parâmetros e valor inicial********
%*********************************************

L = input('Entre com a locação das inovações' )
a = input('Entre com o parâmetro de dependência temporal')
n = input('Entre com o tamanho da série')

%****Valor inicial do processo********

y(1) = L/(1-a);

%****Simulação de n observações do processo INAR(1)-Poisson*****

for t = 2: (n+1)

y(t) = binornd(y(t-1),a,1,1) + poissrnd(L,1,1)

end

save serie y
```

Programa D.2 – Simulação do processo INAR(1)-Geométrica

```
%****Programa gerador da estrutura INAR(1) com a distribuição Geométrica
*****************************************************************
%****Estabelecendo parâmetros e valor inicial********
%*********************************************

T = input('Entre com a locação das inovações' )
f = input('Entre com o parâmetro de dependência temporal')
n = input('Entre com o tamanho da série')

%****Valor inicial do processo********

y(1) = (1-T)/(T*(1-f));

%****Simulação de  n observações do processo INAR(1)-Geométrica*****

for t = 2: (n+1)

y(t) = binornd(y(t-1),a,1,1) + geornd(L,1,1)

end

save serie y
```

Programa D3 – Simulação do processo Poisson-Gama

```
%*****Estabelecendo valores iniciais e o hiperparâmetro*********
%*******************************************************

w = input('entre com o valor do hiperparâmetro')
posta(1) = input('entre com o valor inicial (a0) da distribuição gama');
postb(1) = input('entre com o valor inicial (b0) da distribuição gama');
n = input(' entre com o tamanho da série a ser simulada');


%******Simulação de n valores do processo Poisson-Gama
%*************************************************

for t = 2: (n+1)

        prioa(t) = w*posta(t-1)
        priob(t) = w*postb(t-1)

R(t) = gamarnd(prioa(t),1/priob(t),1,1);
. y(t) = poissrnd(R(t),1,1);


        posta(t) =  prioa(t) + y(t) ;
        postb(t) =  priob(t)+1;

end


save serie y
```

Programa D.4 – Estimação dos parâmetros do modelo INAR(1)-Poisson

```
%***** Programa para a estimação dos parâmetros do modelo INAR(1)-Poisson
%***** por máxima verossimilhança, condicionada ao valor inicial do processo
%*****************************************************************


function [s1] = mvinpoi(a)

%*****Carregando a série a ser analisada********
%*****************************************


load serie

n = length(y);

%******Escrevendo o parâmetro de locação em função
%****** da dependência temporal****************

L = (sum(y)-a*(sum(y)-y(n)))/(n-1)

%*****Calculando o valor da função de verossimilhança para o
%***** instante t*************************************

for t = 2:n

        c = min(y(t),y(t-1));
        h = 0;

                for i = 0 : c

                deno = deno + ((l.^(y(t)-i))*(prod(1:y(t-1)))*((a.^i))*((1-a).^(y(t-1)-i)))/((prod(1:(y(t)-
i)))*(prod(1:(y(t-1)-i)))*(prod(1:i)));

                nume = nume +  ((l.^(y(t)-i-1))*(prod(1:y(t-1)))*((a.^i))*((1-a).^(y(t-1)-i)))/((prod(1:(y(t)-
i-1)))*(prod(1:(y(t-1)-i)))*(prod(1:i)));

                end

                p(t) = (nume/deno);

                h = h + ((y(t)-a*(y(t-1)))-(l*(p(t))))/(a*(1-a));

                end

s1 = h;
```

Programa D.5 – Estimação dos parâmetros do modelo INAR(1)-Geométrica

```
%***** Programa para a estimação dos parâmetros do modelo INAR(1)-Geométrica
%***** por máxima verossimilhança, condicionada ao valor inicial do processo
%*****************************************************************


function [s2] = mvingeo(f)

load serie

n = length(y);

VERO=0;

T=(n-1)/((sum(y)-y(1))+(n-1)-p*(sum(y)-y(n)));
 p = ((n-1)/(sum(y)-y(n)))*(((sum(y))/(n-1))-((y(1))/(n-1))-((1-T)/T));

    for t=2:n

        c = min(y(t),y(t-1));



        TERMO(t)=0;

        P(t)=0;

       for i=0:c

          COMB= (prod(1:y(t-1)))/((prod(1:(y(t-1)-i)))*prod(1:i));

          BINGEO = (T)*((1-T).^(y(t)-i))*COMB*(p.^i)*((1-p).^(y(t-1)-i));

                TERMO(t) = TERMO(t) + (i-y(t))*BINGEO;

          P(t) = P(t)+BINGEO;


       end

       a1(t) = ((P(t))/T);
       b1(t) = ((1/(1-T))*(TERMO(t)));

       logvero = logvero  +  (a1(t)+b1(t))/(P(t));

        end

   s2 = logvero
```

Programa D.6 – Estimação do hiperparâmetro do modelo Poisson-Gama

```
%***** Programa para a estimação dos hiperparâmetro do modelo Poisson-Gama
%***** por máxima verossimilhança, condicionada ao valor inicial do processo
%*****************************************************************

***** Definição da função de máxima verossimilhança
*******************************************

function maxv = mvstrut(w)

Carregamento dos valores da série

        load serie
        n= length(y);

***** índice da primeira observação diferente de zero.
*******************************************
        ind=find(y);
        pas1 = ind(1)+1;

***** Inicialização do procedimento e a estimação por máxima verossimilhança
*****************************************************************

        for i=pas1:n
      prioa(i) = 0;
      priob(i) = 0;
      c = i-1;
        for k=1:c
     prioa(i) = prioa(i)+(w.^k)*(y(i-k));
     priob(i) = priob(i)+(w.^k);
       end

         end

  f1 = 0;
   for j=pas1:n

   f1 = f1+ (gammaln(prioa(j)+y(j))) + (prioa(j))*log(priob(j))-(gammaln(y(j)+1))- (gammaln(prioa(j)))-
(prioa(j)+y(j))*log(1+(priob(j)));
   end

   maxv =(-f1)


disp(w)
```

REFERÊNCIAS BIBLIOGRÁFICAS

Al-Osh & Alzaid, A.A. (1987). *First-order integer-valued autoregressive (INAR(1)) processes*. J. Time Ser. Anal., 8, 261-275.

Alzaid, A.A. & Al-Osh (1988). *First order integer-valued autoregressive (INAR(1)) process : distributional and regression properties*. Statistica Neerlandica, 42, 53-61.

Akaike, H. (1973). *Information theory and extension of the maximum likelihood principle*. Second International Symposium on Information Theory. 267-281. B.N. Petrov & F. Csáki (eds). Budapest : Akadémiai Kiadó.

Box, G.E.P. & Jenkins, G.M. (1970). *Time Series Analysis: Forecasting and Control.* Hoden-Day, Inc., San Francisco.

Box, G.E.P. & Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control* (revised edn). Hoden-Day, Inc., San Francisco.

Cox, D.R. (1981). *Statistical analysis of time series, some recent developments.* Scand. J. Statist. 8, 93-115.

Fernandes, C. (1990). *Non-Gaussian Structural Models, thesis submitted for the degree of Doctor of Philosophy at London School of Economics,* University of London.

Harvey, A.C. (1984). *A unified view of statistical forecasting procedures.* Journal of Forecasting 3 : 245-275.

Harvey, A.C. (1989) *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.

Harvey, A.C & Fernandes, C. (1989) *Time Series Models for count or qualitative observations*. Journal of Business and Economic Statistics, 7, 407-422

Hoaglin. D.C, Mosteller. F.S. & Tukey, J.W. (1983). *Análise Exploratória de Dados - Técnicas Robustas*. Edições Salamandra.

Kalman, R.E. (1960). *A New Aproach to Linear Filtering and Prediction Problems.* Journal of Basic Engineering, 82, 34-45.

Kendall, M. & Ord, J.K. (1990). *Time Series.* 3rd edição, A Charles Griffin Title.

Kullback, J. (1959), *Information Theory and Statistics, New York: John Wiley.*

Jacobs, P. A. & Lewis, P. A. W. (1978a) *Discrete Time Series Generated by Mixtures I: Correlational and Runs Properties*. Journal of The Royal Statistical Society. B40, 94-105.

Jacobs, P. A & Lewis, P. A. W. (1978b) *Discrete Time Series Generated by Mixtures II: Asymptotic Properties*. Journal of The Royal Statistical Society. 40, 222-228.

Jacobs, P.A & Lewis, P.A.W. (1983) *Stationary Discrete Autoregressive-Moving Average Time Series Generated by Mixtures*, Journal of Time Series Analysis Vol 4, n° 1 pp 19-36.

Lawrance, A.J. (1980). *Some autoregressive models for point processes, in: P.Bartfai and J.Tomko (eds), Point Processes and Queueing Problems, pp. 257-275.*

Jin-Guan, D. & Yuan, L. (1991) *The integer-valued autoregressive (INAR(p)) model.* Journal of Time Series Analysis. Vol 12, No 2, 129-42.

Jonhson, N. L. & Kotz, S. (1969). *Distributions in Statistics : Discrete Distributions*, John Wiley & Sons.

Klimko, L. A . & Nelson, P.I. (1978) *On conditional least squares estimation for stochastic processes* Ann. Statist. 6, 629-42.

McCleary, R.& Hay. R. A,Jr. (1980). *Applied Time Series Analysis for the Social Sciences,* Berverly Hills. CA: Sage Publications.

Mc Kenzie, E. (1986), *Autoregressive-moving average processes with negative binomial and geometric marginal distributions*, Advances in applied probability 18, 679-705.

Schwarz, G. (1978) "*Estimating the dimension of a model* ". *The Annals of Statistics*. 6: 461-464.

Shephard, N. G. (1994). *A local Scale Model : Na unobserved component alternative to integrated GARCH processes.* Journal of Econometrics 60, no. 1-2, 181-202.

Smith, R.L. & Miller, J.E. (1986) *A non-Gaussian state space model and application to prediction of records*. Journal of the Royal Statistical Society, Series b, 48, 79-88.

Sprott, D.A, (1983) *Estimating the parameters of a convolution by maximum likelihood.* J Amer. Statist, Ass, 78, 457-60.

Steutel, F.W. & Van Harn, K. (1979) *Discrete analogues of self-decomposability and stability*. Ann. Prob. 7, 893-9.

Zeger, S.L. (1988). *A Regression Model for Time Series of Counts.* Biometrika 75, 621-629.