

Dâmocles Aurélio Nascimento da Silva

**Abordagem Clássica e Bayesiana em Modelos Autoregressivos para Valores
Inteiros: Estudo de Caso em Número de dias com precipitação em
Garanhuns**

RECIFE-PE

SETEMBRO/2014



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Abordagem Clássica e Bayesiana em Modelos Autoregressivos para Valores
Inteiros: Estudo de Caso em Número de dias com precipitação em
Garanhuns

Tese apresentada ao Programa de Pós-
Graduação em Biometria e Estatística
Aplicada como exigência parcial à ob-
tenção do título de Doutor.

Área de Concentração: Biometria e Es-
tatística Aplicada

Orientador: Prof. Dr. Moacyr Cunha
Filho
Co-orientador: Prof. Dr. Eufrázio de
Souza Santos

RECIFE-PE
SETEMBRO/2014

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Abordagem Clássica e Bayesiana em Modelos Autoregressivos para Valores
Inteiros: Estudo de Caso em Número de dias com precipitação em
Garanhuns

Dâmocles Aurélio Nascimento da Silva

Tese julgada adequada para obtenção do título de Doutor em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 24/09/2014 pela Comissão Examinadora.

Orientador:

Prof. Dr. Moacyr Cunha Filho
Universidade Federal Rural de Pernambuco

Banca Examinadora:

Prof^a. Dr^a. Tatijana Stosic
Universidade Federal Rural de Pernambuco

Prof. Dr. Lucian Bogdan Bejan
Universidade Federal Rural de Pernambuco

Prof^a. Dr^a. Ana Patrícia Siqueira Tavares Falcão
Instituto Federal de Pernambuco

Prof. Dr. Irami Buarque do Amazonas
Universidade de Pernambuco

*Dedico,
a minha esposa Gabriela Isabel Limoeiro Al-
ves por todo companheirismo. As minhas
filhas Louise e Geovanna por todo carinho.
Aos meus pais por tudo que me ensinaram
durante toda a minha vida e por me educa-
rem sempre com muita atenção e carinho, a
meus irmãos Sara e Helfarne que sempre me
incentivaram e deram-me força para seguir
em frente.*

Agradecimentos

A Deus que sempre esteve presente em todas as fases da minha vida.

Ao meu orientador Professor Doutor Moacyr Cunha Filho, pela orientação, paciência e praticidade na execução deste trabalho;

Ao professor Eufrázio de Souza Santos, pelo aprendizado e parceria nos trabalhos acadêmicos;

A coordenadora do curso de Biometria e Estatística Aplicada professora doutora Tatijana Stosic, pela orientação e esforço pelo curso.

Ao Programa de Pós-Graduação em Biometria e Estatística Aplicada da Universidade Federal Rural de Pernambuco pela contribuição no meu crescimento intelectual e profissional.

Aos professores do Programa de Doutorado em Biometria e Estatística Aplicada/UFRPE pelos conhecimentos transmitidos.

Aos colegas do Programa de Doutorado em Biometria e Estatística Aplicada/UFRPE, pelo bom convívio. Em especial a Samuel, Rodrigo, Dennis, Milton, Diego, Lêda, Rosilda e Sílvio.

Aos funcionários Marco Santos e Zuleide, pelo carinho, respeito e amizade.

Aos colegas da UPE Campus Garanhuns pela compreensão, apoio e incentivo dado na fase final do doutorado. Em especial ao Professor José Elizângelo Lopes Luna.

A minha família que sempre me acompanhou durante todas as etapas de minha vida. Meus pais Damocles Silva e Maria do Carmo, meus irmãos Sara e Helfarne.

A minha esposa Gabriela Isabel pela dedicação, companheirismo, ajuda e força nos momentos mais difíceis.

A minha tia Sevi (*in memoriam*), que sempre acreditou na realização desse sonho.

A todos que de alguma forma contribuíram para o crescimento de cada momento para realização deste trabalho.

“Não é no silêncio que os homens se fazem, mas na palavra, no trabalho, na ação-reflexão.”

Paulo Freire

Resumo

Muitos aspectos do ciclo meteorológico poderiam ser descrito por dados de séries temporais. Os meteorologistas costumam usar dados de séries temporais para avaliar as condições climáticas e previsões. Esses modelos em geral são modelos contínuos. O interesse foi analisar dados meteorológicos discretos com o modelo INAR(1), através de abordagem clássica e bayesiana na estimação dos parâmetros. A proposta é analisar os dados da série utilizando modelos mistos com abordagem bayesiana. Sendo assim, neste trabalho é descrito uma sequência de procedimentos para estimar parâmetros de modelos autoregressivos de ordem $p = 1$, para valores inteiros INAR(1), por meio de inferência clássica via estimador de máxima verossimilhança e inferência bayesiana via simulação de Monte Carlo em Cadeias de Markov (MCMC). Duas alternativas são consideradas para a densidade a priori dos parâmetros do modelo. Para o primeiro caso, adota-se uma densidade a priori não-informativa. Para o segundo, adota-se uma densidade conjugada beta-gama. A análise a posteriori é efetuada por meio de algoritmos de simulação MCMC. Avalia-se também a previsão de novos valores da série número de dias com precipitação. O período de análise compreendeu 30/11/1993 à 29/02/2012 e obteve previsões do período de 31/03/2012 à 28/02/2013. Foram utilizados um modelo INAR(1) de estimação clássica dos parâmetros e quatro modelos INAR(1) de estimação bayesiana para os parâmetros. A escolha do modelo mais adequado foi utilizado o critério de informação de Akaike (AIC). A análise dos erros de previsão foi um instrumento utilizado para verificar qual modelo se adequou melhor aos dados. Conclui-se que o uso de simulação MCMC torna o processo de inferência bayesiana mais flexível, podendo ser estendido para problemas de dimensão maior. Os modelos mistos apresentaram melhores desempenho do que o modelo clássico, dentre os mistos o que teve resultados mais robustos foi o Modelo INAR(1) Poisson-Normal utilizando a priori Beta. Logo propomos o uso da operação thinning em modelos mistos.

Palavras-chave: Modelos INAR; Inferência Bayesiana; MCMC; Modelos mistos.

Abstract

Many aspects of the weather cycle could be described by time series data. Meteorologists often use time series data to assess climate conditions and forecasts. Such models are generally continuous models. The interest was to analyze discrete weather data with the INAR (1) model, using classical and Bayesian approach to parameter estimation. The proposal is to analyze the data series using mixed models with Bayesian approach. Thus, this work is described a sequence of procedures for estimating parameters of autoregressive models of order $p = 1$, for integer values INAR(1), by classical inference via maximum likelihood estimator and Bayesian inference via simulation Monte Carlo Markov Chain (MCMC). Two alternatives are considered for the a priori density of the model parameters. For the former case is adopted a density non-priori information. For the second, we adopt a density combined beta-gamma. A posteriori analysis is performed by algorithms of MCMC simulation. Also evaluates the prediction of new values of the series number of days with precipitation. The period of analysis comprised 30/11/ 1993 to 29/02/2012 and obtained estimates of the period of 31/03/2012 to 28/02/2013. One INAR (1) model of classical parameter estimation and four models INAR (1) Bayesian estimation for the parameters were used. The choice of the most appropriate model the Akaike information criterion (AIC) was used. The analysis of forecast errors was an instrument used to determine which model is best suited to the data. We conclude that the use of MCMC simulation makes the process more flexible Bayesian inference and can be extended to larger problems. Mixed models showed better performance than the classical model, mixed among the more robust results than had been the model INAR (1) Poisson-Normal using a priori Beta. Soon we propose the use of thinning operation in mixed models.

Keywords: INAR models; Bayesian inference; MCMC; Mixed models

Lista de Figuras

2.1	Linha evolutiva da modelagem INAR	p. 6
3.1	Disposição geográfica do Município de Garanhuns-PE	p. 27
4.1	Número de dias com precipitação em Garanhuns, 1993 a 2012	p. 37
4.2	Média mensal dos números de dias com precipitação	p. 39
4.3	Função de autocorrelação parcial, com 5% de limites de significância.	p. 39
4.4	Histograma da distribuição de probabilidade dos resultados do Modelo 1.	p. 41
4.5	Box-plot da série temporal	p. 42
4.6	Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 2.	p. 42
4.7	Convergência dos parâmetros do Modelo 2	p. 43
4.8	Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 2).	p. 43
4.9	Função de autocorrelação parcial para os parâmetros α e λ do Modelo 2.	p. 44
4.10	Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 3.	p. 44
4.11	Convergência dos parâmetros do Modelo 3	p. 45
4.12	Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 3).	p. 45
4.13	Função de autocorrelação parcial para os parâmetros α e λ do Modelo 3.	p. 46
4.14	Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 4.	p. 46
4.15	Convergência dos parâmetros do Modelo 4.	p. 47
4.16	Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 4).	p. 47
4.17	Função de autocorrelação parcial para os parâmetros α e λ do Modelo 4.	p. 48

4.18	Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 5.	p. 48
4.19	Convergência dos parâmetros do Modelo 5.	p. 49
4.20	Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 5).	p. 49
4.21	Função de autocorrelação parcial para os parâmetros α e λ do Modelo 5.	p. 50
4.22	Teste de aderência (Anderson-Darling) para as distribuições Beta e Normal	p. 52
4.23	Função marginal à posteriori dos parâmetros $\alpha = 0,6539$ e $\lambda = 12,68$, do Modelo 2, com $\alpha = 0,15$	p. 54
4.24	Função marginal à posteriori dos parâmetros $\alpha = 0,6558$ e $\lambda = 12,68$, do Modelo 2, com $\alpha = 0,50$	p. 54
4.25	Função marginal à posteriori dos parâmetros $\alpha = 0,6534$ e $\lambda = 12,64$, do Modelo 3, com $\alpha = 0,15$	p. 55
4.26	Função marginal à posteriori dos parâmetros $\alpha = 0,6546$ e $\lambda = 12,64$, do Modelo 3, com $\alpha = 0,50$	p. 55
4.27	Função marginal à posteriori dos parâmetros $\alpha = 0,6536$ e $\lambda = 12,67$, do Modelo 4, com $\alpha = 0,15$	p. 56
4.28	Função marginal à posteriori dos parâmetros $\alpha = 0,6547$ e $\lambda = 12,66$, do Modelo 4, com $\alpha = 0,50$	p. 56
4.29	Função marginal à posteriori dos parâmetros $\alpha = 0,6560$ e $\lambda = 12,64$, do Modelo 5, com $\alpha = 0,15$	p. 56
4.30	Função marginal à posteriori dos parâmetros $\alpha = 0,6544$ e $\lambda = 12,62$, do Modelo 5, com $\alpha = 0,50$	p. 57
4.31	Função marginal à posteriori dos parâmetros $\alpha = 0,6542$ e $\lambda = 12,68$, do Modelo 2, com 50000 iterações, sendo 4000 de <i>Burn-in</i>	p. 58
4.32	Função marginal à posteriori dos parâmetros $\alpha = 0,6543$ e $\lambda = 12,68$, do Modelo 2, com 500000 iterações, sendo 4000 de <i>Burn-in</i>	p. 58
4.33	Função marginal à posteriori dos parâmetros $\alpha = 0,6505$ e $\lambda = 12,64$, do Modelo 3, com 50000 iterações, sendo 4000 de <i>Burn-in</i>	p. 58
4.34	Função marginal à posteriori dos parâmetros $\alpha = 0,6505$ e $\lambda = 12,64$, do Modelo 3, com 500000 iterações, sendo 4000 de <i>Burn-in</i>	p. 59

- 4.35 Função marginal à posteriori dos parâmetros $\alpha = 0,6540$ e $\lambda = 12,66$,
do Modelo 4, com 50000 iterações, sendo 4000 de *Burn-in*. p. 59
- 4.36 Função marginal à posteriori dos parâmetros $\alpha = 0,6540$ e $\lambda = 12,66$,
do Modelo 4, com 500000 iterações, sendo 4000 de *Burn-in*. p. 60
- 4.37 Função marginal à posteriori dos parâmetros $\alpha = 0,6528$ e $\lambda = 12,67$,
do Modelo 5, com 50000 iterações, sendo 4000 de *Burn-in*. p. 60
- 4.38 Função marginal à posteriori dos parâmetros $\alpha = 0,6529$ e $\lambda = 12,66$,
do Modelo 5, com 500000 iterações, sendo 4000 de *Burn-in*. p. 60

Lista de Tabelas

4.1	Medidas descritivas da série de número de dias com precipitação	p. 38
4.2	Medidas descritivas da série de número de dias com precipitação por mês	p. 38
4.3	Valores de AIC obtidos nos modelos	p. 50
4.4	Valores de AICc obtidos nos modelos	p. 51
4.5	Valores de BIC obtidos nos modelos	p. 51
4.6	Valores previstos pelos modelos (12 passos à frente) e erros de previsão	p. 53
4.7	Amplitude dos parâmetros α e λ variando α	p. 57
4.8	Amplitude dos parâmetros α e λ variando quantidade de iterações . . .	p. 61

Sumário

1	Introdução	p. 1
2	Fundamentação Teórica	p. 6
2.1	Modelo autoregressivo de valores inteiros e suas propriedades	p. 6
2.1.1	Modelo autoregressivo de valores inteiros (INAR)	p. 7
2.1.2	Operador <i>thinning</i>	p. 7
2.1.3	Modelo INAR(1)	p. 8
2.1.4	Previsão INAR	p. 11
2.1.4.1	Intervalo de previsão	p. 12
2.2	Distribuição de Poisson	p. 12
2.2.1	Função de verossimilhança	p. 13
2.3	Critérios de seleção de modelos	p. 14
2.4	Estatística Bayesiana	p. 15
2.4.1	Inferência Bayesiana	p. 16
2.4.2	Análise Bayesiana em Séries Temporais	p. 16
2.4.3	Teorema de Bayes	p. 17
2.4.4	Distribuição a priori	p. 19
2.4.4.1	Priori não informativa	p. 19
2.4.4.2	Priori conjugada	p. 20
2.4.5	Monte Carlo com Cadeias de Markov	p. 21
2.4.5.1	Amostrador de Gibbs	p. 23
2.5	Modelos para série temporal número de dias com precipitação	p. 24

2.5.1	Modelo Clássico	p. 24
2.5.2	Modelo Poisson	p. 24
2.5.3	Modelo Poisson-Normal	p. 24
3	Material e Métodos	p. 26
3.1	Descrição dos dados	p. 26
3.2	Inferência bayesiana para processos INAR(1) de Poisson	p. 27
3.3	Inferência bayesiana para processos INAR(1) de Poisson-Normal	p. 30
3.4	Implementação computacional	p. 32
3.4.1	WinBUGS	p. 32
3.4.2	Diagnóstico de Convergência	p. 33
3.4.3	Método Informal	p. 34
3.5	Medidas de Erros de Previsão	p. 34
3.6	Gráficos de probabilidade	p. 35
3.7	Método de Anderson-Darling	p. 36
4	Resultados e Discussões	p. 37
4.1	Análise descritiva dos dados	p. 38
4.2	Modelos ajustados	p. 40
4.2.1	Clássica - Modelo INAR(1) Poisson (Modelo 1)	p. 41
4.2.2	Bayesiana - Modelo INAR(1) Poisson, priori Beta (Modelo 2)	p. 42
4.2.3	Bayesiana - Modelo INAR(1) Poisson, priori Normal (Modelo 3)	p. 44
4.2.4	Bayesiana - Modelo INAR(1) Poisson-Normal, priori Beta (Modelo 4)	p. 46
4.2.5	Bayesiana - Modelo INAR(1) Poisson-Normal, priori Normal (Modelo 5)	p. 48
4.2.6	Adequação dos modelos	p. 50
4.3	Distribuição de Probabilidade	p. 51

4.4	Previsão	p. 53
4.5	Análises Complementares	p. 54
4.5.1	Parte I	p. 54
4.5.2	Parte II	p. 57
5	Conclusão	p. 62
	Referências	p. 64
	Apêndice	p. 67

1 Introdução

O continente americano experimentou, nos últimos anos, uma sucessão de acontecimentos radicais: chuvas torrenciais na Venezuela, inundações nos pampas argentinos, secas na Amazônia, tempestades de granizo na Bolívia e uma temporada recorde de furacões no Caribe. Ao mesmo tempo, as chuvas diminuem no Chile, no sul do Peru e no sudoeste da Argentina. Com a elevação de temperaturas já registrada ($+1^{\circ}C$ na América Central e na América do Sul em um século, ante a média mundial de $+0,74^{\circ}C$), os glaciares andinos estão retrocedendo. A disponibilidade de água destinada ao consumo e à geração de eletricidade já está comprometida e o problema se agravará no futuro, tornando-se crônico caso medidas não sejam tomadas, afirma o relatório do IPCC (Painel Intergovernamental sobre Mudanças Climáticas) para a América Latina (SHIKLOMANOV et al., 2000).

O Brasil tem posição privilegiada no mundo, em relação à disponibilidade de recursos hídricos. A vazão média anual dos rios em território brasileiro é de cerca de 180 mil m^3/s . Esse valor corresponde a aproximadamente 12% da disponibilidade mundial de recursos hídricos, que é de 1,5 milhão de m^3/s (SHIKLOMANOV et al., 2000). Se forem levadas em conta as vazões oriundas em território estrangeiro e que ingressam no país (Amazônica: 86.321 mil m^3/s ; Uruguai: 878 m^3/s e Paraguai: 595 m^3/s), a vazão média total atinge valores da ordem de 267 mil m^3/s (18% da disponibilidade mundial).

Em relação às chuvas, observa-se a tendência já detectada, em estudos anteriores do IPCC, o aumento de até 30%/década da chuva na bacia do Prata e em algumas áreas isoladas do Nordeste. Essa região do Brasil possui apenas 3% de água doce. Em Pernambuco, existem apenas 1.320 litros de água por ano por habitante (TRENBERTH; DAI, 2007).

Se os governos dos países carentes de água não adotarem medidas urgentes para estabilizar a população e elevar a produtividade hídrica, a escassez de água em pouco tempo se transformará em falta de alimentos. Logo, só a chuva não é suficiente para repor os recursos hídricos necessários a vida, mas ter uma forma de identificar o quanto vai

chover num período, auxilia a população para se proteger e se organizar no enfrentamento à escassez de água, enquanto as políticas públicas, nesse assunto, não for prioridade.

Em 2013, várias cidades do Agreste de Pernambuco viveram um cenário difícil. As plantas secas e sem folhas, a terra rachada e os açudes vazios geraram ausência do verde na região. Segundo Agência Pernambucana de Águas e Clima (APAC), essa situação se deu devido ao déficit de chuvas em 2012. Nessa lógica apresentada pelos órgãos, 2014 deveria ter sido um ano difícil, pois o mesmo *déficit* ocorreu em 2013. Fato que até agosto do corrente ano, não aconteceu (APAC, 2013).

Assim o objetivo será realizar uma modelagem nos dados referente a precipitação de chuvas do município de Garanhuns, uma das cidades pertencente ao Agreste de Pernambuco, usando de uma metodologia que analise ao longo do tempo esse evento, que se adeque a dados discretos (inteiros) e seja um modelo de previsão satisfatório. Essa modelagem será comparada com o modelo utilizado por Pavlopoulos e Karlis (2007).

Muitos aspectos do ciclo meteorológico poderiam ser descrito por dados de séries temporais. Os meteorologistas costumam usar dados de séries temporais para avaliar as condições climáticas e previsões. Em muitos estudos, hidrólogos também usam séries temporais para análise dos dados referentes à quantidade de chuva precipitada em uma região para os últimos dias, anos ou um período de 10 anos (GUIMARÃES; SANTOS, 2011; LEE; LEE, 2000). Sendo assim, abordaremos o problema utilizando séries temporais para dados discretos, sob duas óticas: abordagem de estimação dos parâmetros clássica e abordagem de estimação dos parâmetros bayesiana.

Dentre os modelos de Séries Temporais, uma classe de modelos ainda pouco explorado pela metodologia bayesiana é a modelagem de séries temporais para variáveis discretas. Variáveis discretas de séries temporais são contempladas em áreas como: Indústria (número de acidentes de uma determinada fábrica em cada mês), Hospitais (número de pacientes tratados por hora num hospital ou clínica médica), Comércio (quantidade de produto comercializado por hora em um shopping). Estes dados podem também surgir da “discretização” de uma variável contínua de séries temporais. Já o inverso, muitas vezes leva a nos deparar com um elevado grau de dificuldade.

Na literatura, encontramos modelos conhecidos como autoregressivos de valores inteiros INAR(p), que são uma adaptação dos modelos AR a dados inteiros em que a operação multiplicação é substituída pela operação *thinning* definida por (STEUTEL; HARN, 1979).

A distribuição comumente utilizada para analisar dados na forma de contagem é a

distribuição de Poisson, que apesar de bastante útil apresenta uma restrição forte, a igualdade da média e da variância dos dados. Por outro lado, a aproximação bayesiana fornece uma estrutura coerente que facilita a análise de problemas de decisão sobre incertezas (BERGER, 1995). Inicia-se a análise bayesiana para seleção do modelo através da denominação das probabilidades usadas para cada modelo e, posteriormente, escolhe-se as distribuições a priori para os parâmetros que ainda são desconhecidos. Tipicamente, seleciona-se o modelo ou modelos com a maior probabilidade a posteriori.

A investigação de diferentes distribuições de probabilidade a priori é objeto deste estudo, já que estas distribuições são características intrínsecas dos parâmetros dos modelos. Segundo Berger (1995), a primeira vantagem do procedimento bayesiano é a simplicidade na interpretação das conclusões relacionadas às probabilidades a posteriori.

A segunda vantagem é a consistência, quando se aumenta o número de dados, pois garante a seleção do modelo mais apropriado ou de um modelo próximo. Os métodos clássicos, tipicamente, falham neste critério mínimo já que os modelos selecionados se tornam complexos quando há uma grande quantidade de dados. E, a terceira vantagem, apresentada pelo autor, é a possibilidade de se poder incluir a incerteza na seleção do modelo bayesiano. Além das vantagens, anteriormente, apresentadas, ainda existe a possibilidade da seleção do modelo bayesiano ser aplicada na comparação de diversos modelos e aplicada, de modo genérico, ou seja, sem que estes modelos pertençam à família padrão e nem estar agrupados segundo o tipo.

Conforme Berger e Insua (1996), as duas dificuldades na utilização do modelo Bayesiano são a escolha da distribuição a priori e a computação do modelo escolhido. Porém, a escolha da distribuição a priori é considerada o maior problema. Este pode ser o caso em que o conhecimento subjetivo sobre os parâmetros desconhecidos é avaliado e pode ser incorporado à subjetividade própria das densidades a priori para estes parâmetros. Isto é, claramente, desejável se puder ser realizado. Algumas ferramentas computacionais recentes têm permitido a aplicação de métodos bayesianos para modelos de alta complexidade e não padronizados. Na verdade, para modelos mais complicados, a análise bayesiana tenha, talvez, se tornado o mais simples, e frequentemente o único, método de análise.

Segundo Berger e Insua (1996) a base da aplicação da teoria bayesiana é conceitual e prática: fornece uma estrutura coerente e facilita a análise de problemas de decisão sobre incertezas. As críticas relativas aos métodos bayesianos estão centradas em três aspectos:

- Computacionais: A implementação do suporte bayesiano muitas vezes aponta para

as dificuldades dos problemas computacionais. Como uma consequência, a implementação, é algumas vezes, questionada numa tentativa de chamar a atenção para modelos muito simplificados ou para empreender uma análise não bayesiana. O recente desenvolvimento dos métodos de Monte Carlo via Cadeia de Markov (MCMC) e outras ferramentas computacionais bayesianas têm ultrapassado estas críticas, levando em consideração uma modelação mais realística (e tipicamente mais complexa).

- **Imprecisão:** É frequentemente apontado que o suporte bayesiano demanda precisão excessiva nos julgamentos do decisor, particularmente, quanto à especificação da distribuição a priori. O autor recorda que o suporte bayesiano é normativo. Quando aplicado, este suporte serve como um guia para agir sob a incerteza. Portanto, em estágios anteriores de uma análise, o modelo pode não ser capaz de produzir informações precisas (prioris, modelo, utilidade). Para esclarecer este fato, métodos bayesianos robustos têm sido desenvolvidos. Esses métodos consistem em empreender uma família de análises bayesianas e as suas conclusões em razões comuns. Se há também muita discrepância, as ferramentas bayesianas robustas sugerem como resolvê-las, guiando-as à elucidação.
- **Descritiva:** Algumas vezes, discute-se que decisores atuais não confirmam os postulados bayesianos. Berger e Insua (1996) em seu trabalho mostram que muitos estudos experimentais tem apontado que muitos decisores violam os postulados bayesianos em tarefas desassistidas o que sugere fragilidade do procedimento bayesiano como uma teoria descritiva. Alguns autores interpretam e tratam o assunto com status normativo, embora tal interpretação não carregue nenhuma força lógica. Em muitos casos, novas teorias têm surgido tentando aperfeiçoar, de modo geral, a análise bayesiana sob um ponto de vista descritivo.

A construção de um modelo é baseada em informações obtidas da realidade através das observações ou medidas. Em geral, é difícil afirmar, com certeza, se um modelo matemático idealizado é ou não adequado, antes que algum teste de observação seja realizado (MARTINS et al., 2001).

O objetivo deste trabalho foi modelar uma série temporal de dados discretos, com abordagem clássica e comparar com o modelo sob abordagem bayesiana, na estimação dos parâmetros. Na abordagem bayesiana, sugerimos um modelo misto para ajuste dos dados, comparando os modelos entre si, modificando as distribuições à priori e analisando a partir dos critérios de seleção de modelo e os tipos de erros. Os dados discretos utilizados

referem-se ao número de dias com precipitação no município de Garanhuns, no período de 11/1993 à 02/2013, totalizando 220 observações.

Portanto, o trabalho está estruturado da seguinte forma: na seção 2, consta a fundamentação teórica; na seção 3, aborda-se uma metodologia clássica e a Bayesiana para analisar os dados; na seção 4, discute-se os resultados das metodologias propostas; e finalmente, na seção 5, apresenta-se as conclusões.

2 Fundamentação Teórica

2.1 Modelo autoregressivo de valores inteiros e suas propriedades

A modelagem INAR iniciou-se com a apresentação do operador *thinning* binomial por Steutel e Van Harn (1979). Em 2009, foi publicada a comparação entre as previsões tradicional *versus* a Bayesiana para o modelo INAR. A Figura 2.1 mostra um esquema cronológico do desenvolvimento dessa modelagem.

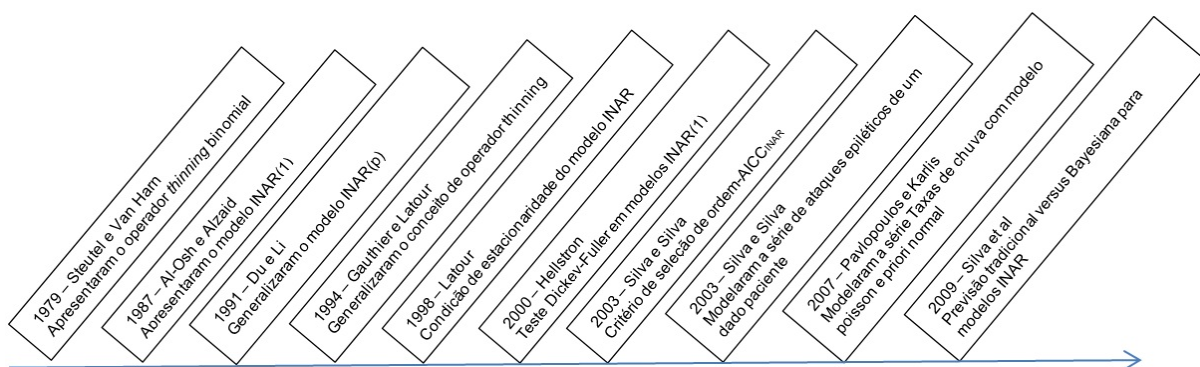


Figura 2.1: Linha evolutiva da modelagem INAR

Os modelos de Séries Temporais para dados discretos ainda são pouco explorados pela metodologia bayesiana. Mesmo que dados nesse formato apareçam nas mais diversas áreas, a recente ascensão pelo uso da metodologia bayesiana se dá após os avanços computacionais. Logo a abordagem bayesiana ganhou força nos últimos 40 anos, sendo assim um tempo ainda incipiente para modelar as diversas áreas do conhecimento.

Na literatura bayesiana, ainda encontramos muitas discussões acerca da dificuldade de escolher uma priori adequada para os modelos propostos. Essa abordagem causa curiosidade de investigação, pois é passível de se encontrar uma priori que se ajuste melhor a modelagem em comparação com resultados encontrados anteriormente. A condiciona-

lidade à qual o contexto bayesiano é posta, causa esse efeito de atualização do modelo, onde um modelo atual hoje pode cair facilmente por terra amanhã.

Os modelos INAR (Autoregressivo de Valores Inteiros), são uma adaptação dos modelos AR (autoregressivo) para dados inteiros em que a operação multiplicação é modificada pela operação *thinning*, proposta por Steutel e Van Harn(1979). Em 1987, Al-Osh e Alzaid, utilizaram a idéia do operador *thinnig* para modelo autoregressivo, sendo assim chamado de INAR. Antes da proposta apresentada utilizava-se o modelo contínuo para dados discretos (AL-OSH; ALZAID, 1987).

Uma modelagem foi proposta por Pavlopoulos e Karlis (2007) para valores inteiros de taxas de chuva (mm/h) em uma região remota no mar da China $(-2, 156)$ que cobre uma área de $240 \times 240 km^2$. Os dados foram coletados via radar de precipitação acoplado em um navio. Na seleção da ordem dos modelos INAR, usou-se os tradicionais critérios automáticos AIC e BIC proposto por Akaike. A melhor ordem para os modelos $INAR(p)$ foi a ordem $p = 1$. Silva, Pereira e Silva (2009) apresentaram duas metodologias para a previsão em modelos INAR. Os autores compararam a metodologia clássica com a metodologia Bayesiana com a finalidade de obter previsões coerentes para o modelo INAR. Obtiveram por meio de simulações previsões teóricas para as duas metodologias, em seguida foram apresentados os intervalos de confiança para cada metodologia. Nas simulações a metodologia Bayesiana obteve melhores resultados. Do ponto de vista prático foi modelado a série de trabalhadores envolvidos em acidentes por queimadura, para um conjunto de dados de 120 observações com previsão de até 6 dias, com seus respectivos intervalos de confiança.

2.1.1 Modelo autoregressivo de valores inteiros (INAR)

O modelo Autoregressivo de Valores Inteiros de ordem (1), apresentado por Al-Osh e Alzaid (1987), baseia-se no operador *thinning* proposto por Steutel e Van Harn (1979). A generalizaram o modelo INAR para ordem p , isto é $INAR(p)$, foi feita por autores em 1991, (DU; LI, 1991).

2.1.2 Operador *thinning*

Os processos INAR é uma adequação dos modelos AR a dados inteiros em que a operação multiplicação é substituída pela operação *thinning* definida por Steutel e Van Harn (1979). A operação *thinning* representa-se por “ \circ ”.

A operação *thinning* entre o parâmetro α e a variável aleatória Y define-se como,

$$\alpha \circ Y = \sum_{i=1}^Y U_i(\alpha), \quad (2.1)$$

onde $\{U_i(\alpha)\}$ é uma sucessão de i.i.d de v.a's de Bernoulli com probabilidade de sucesso α , isto é $U_i(\alpha) \sim B(1, \alpha)$. A sucessão de $\{U_i(\alpha)\}$ designa-se por processo de contagem. Assim, $1 \circ Y = Y$ e $0 \circ Y = 0$.

Deste modo, a operação *thinning* é uma operação aleatória discreta com uma distribuição de probabilidade associada. Portanto, $\alpha \circ Y|Y$ tem distribuição binomial de parâmetros Y e α . Por este motivo, a operação dada em 2.1 é também designada operação *thinning* binomial.

Algumas propriedades da operação *thinning* binomial foram estudadas por Silva e Silva (2003).

- i) $E[\alpha \circ Y] = \alpha E[Y]$
- ii) $E[W(\alpha \circ Y)] = \alpha E[WY]$
- iii) $E[(\alpha \circ Y)(\beta \circ Z)] = \alpha\beta E(YZ)$

onde Y, W e Z são variáveis aleatórias

2.1.3 Modelo INAR(1)

O processo autoregressivo de valores inteiros foi introduzido por McKenzie (1986) e posteriormente por AL-Osh & Alzaid(1987). Recorreram à operação *thinning* binomial definida por Steutel e van Harn (1979) para substituir a operação de multiplicação usual no processo AR(1) e propuseram o processo INAR(1) particularmente usado para séries de contagem. Assim um processo estocástico discreto de valor inteiro não negativo $\{Y_t\}$ será um processo de contagem INAR(1) se satisfizer a condição:

$$Y_t = \alpha \circ Y_{t-1} + \epsilon_t, t \geq 2, \quad (2.2)$$

onde $\alpha \in]0, 1]$, a operação \circ é a operação *thinning* binomial definida em 2.1 e $\{\epsilon_t\}$ é uma sucessão de variáveis aleatórias inteiras, não negativas, independentes e identicamente distribuídas, com média μ_t , variância finita σ_t^2 e independentes de y_{t-1} . Assim, este

processo indica que, de modo independente, em cada momento t cada elemento pode continuar no processo com probabilidade α ou abdicar com probabilidade $1 - \alpha$.

A distribuição marginal do modelo 2.1 pode ser expressa em termos da sequência $\{\epsilon\}$ como:

$$\begin{aligned}
Y_1 &= \epsilon_1 \\
Y_2 &= \alpha \circ \epsilon_1 + \epsilon_2 \\
Y_3 &= \alpha \circ (\alpha \circ \epsilon_1 + \epsilon_2) + \epsilon_3 \\
&= \alpha \circ (\alpha \circ \epsilon_1) \alpha \circ \epsilon_2 + \epsilon_3 \\
&= \alpha^2 \circ \epsilon_1 + \alpha \circ \epsilon_2 + \epsilon_3 \\
&\vdots \\
Y_n &\stackrel{d}{=} \sum_{j=0}^{\infty} \alpha^j \circ \epsilon_{n-j}
\end{aligned} \tag{2.3}$$

Para o processo INAR(1), Al-Osh e Alzaid (1987), demonstrou que pode se escrever o valor esperado como

$$E(Y_t) = E(\alpha \circ Y_{t-1} + \epsilon_t) = E(\alpha \circ Y_{t-1}) + E(\epsilon_t) = \alpha E(Y_{t-1}) + \mu_t, \tag{2.4}$$

onde, devido a estacionariedade do processo dada em 2.2, tem-se $E(Y_t) = E(Y_{t-1})$ e, portanto, a equação 2.4 pode ser reescrita da seguinte forma:

$$E(Y_t) = \alpha E(Y_{t-1}) + \mu_t \Rightarrow (1 - \alpha)E(Y_t) = \mu_t \Rightarrow \frac{E(Y_t)}{(1 - \alpha)}. \tag{2.5}$$

A variância é dada por:

$$\begin{aligned}
Var(Y_t) &= Var(\alpha \circ Y_{t-1} + \epsilon_t) = Var(\alpha \circ Y_{t-1}) + Var(\epsilon_t) = Var(\sum_{i=1}^{Y_{t-1}} U_i) + \sigma_t^2 \\
&= Var[E(\alpha \circ Y_{t-1} | Y_{t-1})] + E[(Var(\alpha \circ Y_{t-1} | Y_{t-1})) + \sigma_t^2] \\
&= Var[\alpha Y_{t-1}] + E[\alpha(1 - \alpha)Y_{t-1}] + \sigma_t^2 \\
&= \alpha^2 Var[Y_{t-1}] + \alpha E(Y_{t-1}) + \sigma_t^2
\end{aligned} \tag{2.6}$$

Analogamente, devida a estacionariedade do processo dada em 2.5, tem-se $Var(Y_t) = Var(Y_{t-1})$... e, portanto, a equação 2.6 pode ser reescrita do seguinte modo:

$$\begin{aligned}
\text{Var}(Y_t) &= \alpha^2 \text{Var}[Y_{t-1}] + \alpha E(Y_{t-1}) + \sigma_t^2 \\
&= \alpha^2 \text{Var}[Y_t] + \alpha E(Y_{t-1}) + \sigma_t^2 \\
&= (1 - \alpha^2) \text{Var}(Y_t) \\
&= \alpha \mu_t + \sigma_t^2 \\
&= (\alpha \mu_t + \sigma_t^2) / (1 - \alpha^2)
\end{aligned}$$

Logo, o processo INAR(1) dado por $Y_t = \alpha \circ Y_{t-1} + \epsilon_t$ é estacionário quando $|\alpha| < 1$ e não estacionário quando $|\alpha| = 1$.

As funções de autocovariância e autocorrelação são definidas pelas expressões seguintes:

$$\gamma(k) = \text{Cov}(Y_t, Y_{t+k}) = \alpha^k \gamma(0) \quad (2.7)$$

$$\rho(k) = \text{Corr}(Y_t, Y_{t+k}) = \alpha^k \quad (2.8)$$

Portanto, estas propriedades tem algumas semelhanças com as propriedades do processo AR(1) visto que ambos os modelos tem a mesma estrutura de autocorrelação. Em 1991, Du e Li propuseram uma generalização para o modelo INAR de ordem p :

$$Y_t = \alpha_1 \circ Y_{t-1} + \alpha_2 \circ Y_{t-2} + \dots + \alpha_p \circ Y_{t-p} + \epsilon_t, \quad (2.9)$$

onde $\{\epsilon_t\}$ é uma sucessão de variáveis aleatórias independentes, $\alpha_i \in [0, 1]$, $i = 1, \dots, p$, $\alpha_p \neq 0$ e as séries de contagem $\alpha_k \circ Y_{t-k}$, $k = 1, \dots, p$ são mutuamente independentes e independentes de $\{\epsilon_t\}$.

Conforme Du e Li (1991) as condições de estacionariedade de segunda ordem INAR(p) são as mesmas de um processo AR(p), ou seja, um processo INAR(p) é estacionário se as raízes do polinômio autoregressivo de ordem p estiverem fora do círculo unitário. Mais tarde, (LATOURE, 1998) mostrou que a condição de estacionariedade é equivalente a $\sum_{k=1}^p \alpha_k < 1$. Estes mesmos autores propuseram estimadores de Yule-Walker para os parâmetros desconhecidos e demonstraram que o método de mínimos quadrados condicionais fornece estimativas assintoticamente normais.

Trabalhos foram desenvolvidos com incidência nos modelos INAR de Poisson com um estudo comparativo entre as abordagens clássica e bayesiana (SILVA, 2005). Posteriormente trabalhos com incidência nos modelos INAR de Poisson, comparando os argumentos p , em abordagem bayesiana foram desenvolvidos.

2.1.4 Previsão INAR

Após o término do processo de identificação, estimação dos parâmetros e verificação, o último passo é realizar as previsões de observações futuras dado que temos observado uma série até o tempo n , ou seja, $y_n = (y_1, y_2, \dots)$. Partindo-se da origem em n e supondo que o objetivo é prever a série em período futuro h , \hat{Y}_{n+h} representa a previsão para um período $n + h$ feita em n . A previsão de múltiplos passos para o período $n + h$ é normalmente construída a partir de sucessivas previsões para os períodos $n + 1, n + 2, \dots, n + h - 1$ (JOHNSON; MONTGOMERY, 1976). Neste procedimento, o valor de Y_{n+h} , o qual não se conhece no tempo n , é substituído pela sua previsão \hat{Y}_{n+h} . Portanto, a previsão se torna o cálculo do valor esperado de uma futura observação condicionada aos valores passados e ao valor presente da variável. Ou seja, \hat{Y}_{n+h} é o valor previsto para um horizonte h períodos de tempo futuro e n é o período de origem da previsão, então,

$$\hat{Y}_{n+h}|y_n = E(\hat{Y}_{n+h}|Y_n) = \alpha^h \left[Y_n - \frac{\lambda}{1-\alpha} \right] + \frac{\lambda}{1-\alpha}, h = 1, 2, 3, \dots \quad (2.10)$$

A função de geração do momento de Y_{n+h} dado Y_n é:

$$\phi_{Y_{n+h}|Y_n}(s) = [\alpha^h e^s + (1 - \alpha^h)]^{X_n} \exp \left[\lambda \frac{1 - \alpha^h}{1 - \alpha} (e^s - 1) \right] \quad (2.11)$$

A expressão 2.11 mostra que a distribuição dos $Y_{n+h}|Y_n$ é uma combinação de uma distribuição binomial com parâmetros α^h e Y_n e uma distribuição de Poisson com parâmetro $\lambda(1 - \alpha^h)/(1 - \alpha)$ (FREELAND; MCCABE, 2004). Ou seja, a função de probabilidade de $Y_{n+h}|Y_n$ é dada por:

$$\begin{aligned} f(y_{n+h}|y_n) &= P(Y_{n+h} = y_{n+h}|Y_n = y_n) \\ &= \exp \left(-\lambda \frac{1 - \alpha^h}{1 - \alpha} \right) \sum_{i=0}^{M_h} \frac{1}{y_{n+h-i}!} \times \left(\lambda \frac{1 - \alpha^h}{1 - \alpha} \right)^{y_{n+h-i}} \binom{y_n}{i} (\alpha^h)^i (1 - \alpha^h)^{y_n - i}. \end{aligned} \quad (2.12)$$

Onde $M_h = \min(Y_{n+h}, Y_n)$ e $y_{n+h} = 0, 1, \dots$

A fim de obter previsões para Y_{n+h} Freeland e McCabe (2004) sugerem a utilização do valor que minimiza o erro esperado absoluto, dada a amostra, isto é, $E[|Y_{n+h} - \hat{Y}_{n+h}| | Y_n]$. Assim, eles concluíram que $\hat{Y}_{n+h} = \hat{m}_{n+h}$ é a mediana do h -passo à frente da distribuição condicional $f(y_{n+h}|y_n) = P(Y_{n+h} = y_{n+h}|Y_n = y_n)$

2.1.4.1 Intervalo de previsão

Um intervalo de confiança para o preditor \hat{Y}_{n+h} pode ser calculado através da função de probabilidade do h -passo à frente do erro de previsão, dada por:

$$e_{n+h}|y_n = Y_{n+h}|y_n - \hat{Y}_{n+h}|y_n \quad (2.13)$$

Substituindo $Y_{n+h}|y_n$ pelo seu preditor 2.10, obtemos

$$e_{n+h}|y_n = k - \alpha^h y_n - \lambda \frac{1 - \alpha^h}{1 - \alpha}, k = 0, 1, 2, \dots \quad (2.14)$$

Assim,

$$\begin{aligned} P(e_{n+h} = k - \alpha^h y_n - \lambda \frac{1 - \alpha^h}{1 - \alpha} | x_n) &= P(Y_{n+h} = k | Y_n = y_n) \\ &= \exp \left\{ -\lambda \frac{1 - \alpha^h}{1 - \alpha} \right\} \sum_{i=0}^{M_h} \frac{1}{(k-i)!} \left(\lambda \frac{1 - \alpha^h}{1 - \alpha} \right)^{k-i} \binom{y_n}{i} (\alpha^h)^i (1 - \alpha^h)^{x_n - i} \end{aligned} \quad (2.15)$$

A partir da expressão 2.15, obtemos um intervalo de confiança para $Y_n + h$.

$$(\hat{Y}_{n+h} + e_{t1}, \hat{Y}_{n+h} + e_{t2}), \quad (2.16)$$

onde \hat{Y}_{n+h} é dada por 2.10, e_{t1} é o maior valor de e_{n+h} , com $P(e_{n+h} \leq e_{t1}) \leq (1 - \gamma)/2$ e e_{t2} é o menor valor de e_{n+h} , com $P(e_{n+h} \leq e_{t2}) \geq (1 - \gamma)/2$ ao nível γ .

2.2 Distribuição de Poisson

A distribuição de Poisson foi descoberta por Siméon-Denis Poisson e publicada, conjuntamente com sua teoria de probabilidade, em 1837, no seu trabalho *Recherches sur la probabilité des jugements en matières criminelles et matière civile* (Investigação sobre a probabilidade de decisões em assuntos criminais e civis). A distribuição de Poisson é de grande importância tanto teórica como aplicada, principalmente, para análise de dados na forma de contagens.

Seja Y uma variável aleatória com distribuição de Poisson, $Y \sim \text{Poisson}(\lambda)$, $\lambda > 0$, com função de probabilidade de Y dada por:

$$P(Y = y) = \frac{e^{-\lambda} \lambda^y}{y!}, y = 0, 1, 2, \dots \quad (2.17)$$

A esperança de Y é dada por:

$$\begin{aligned} E(Y) &= \sum_{y=0}^{\infty} y P(Y = y) = \sum_{y=0}^{\infty} y \frac{e^{-\lambda} \lambda^y}{y!} \\ &= \lambda e^{-\lambda} \sum_{y=1}^{\infty} \frac{\lambda^{y-1}}{(y-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda, \end{aligned} \quad (2.18)$$

pois $\sum_{y=a}^{\infty} \frac{\lambda^{y-a}}{(y-a)!} = e^{\lambda}$, em que a é uma constante.

A variância de Y pode ser calculada da seguinte forma:

$$Var(Y) = E(Y^2) - [E(Y)]^2 = E(Y^2 - Y) + E(Y) - [E(Y)]^2. \quad (2.19)$$

Mas,

$$E(Y^2 - Y) = \sum_{y=0}^{\infty} (y^2 - y) P(Y = y) = \lambda^2 e^{-\lambda} \sum_{y=2}^{\infty} \frac{\lambda^{y-2}}{(y-2)!} = \lambda^2. \quad (2.20)$$

Assim, $Var(Y) = \lambda$ e, portanto, para a distribuição de Poisson tem-se que a média e a variância são iguais.

2.2.1 Função de verossimilhança

Seja Y_1, Y_2, \dots, Y_n uma amostra aleatória de uma distribuição de Poisson com média $\lambda > 0$. A função de verossimilhança para λ é dada por:

$$L(\lambda) = \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^n y_i}}{\prod_{i=1}^n y_i!} \quad (2.21)$$

cujo logaritmo é dado por:

$$l(\lambda) = -n\lambda + \sum_{i=1}^n y_i \log(\lambda) - \sum_{i=1}^n \log y_i!. \quad (2.22)$$

Resolvendo-se a equação $\frac{dl(\lambda)}{d\lambda} = 0$, obtém-se a estimativa de máxima verossimilhança,

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n y_i \quad (2.23)$$

que é a média amostral dos dados observados.

2.3 Critérios de seleção de modelos

A seleção da ordem do modelo é uma etapa de extrema importância na análise de séries temporais, pois é a partir dessa seleção que a escolha do modelo mais parcimonioso para a série é feita. A identificação da ordem do modelo consiste na determinação do valor de p , ou seja, do número de termos a serem considerados em um modelo auto-regressivo em cada mês. Os principais métodos utilizados nesta fase são as análises das funções de auto-correlação periódica e auto-correlação parcial periódica e os critérios: o critério de informação de Akaike (AIC), critério de informação de Akaike corrigido (AICc) e o critério de informação Bayesiano (BIC).

Um modelo com mais parâmetros pode ter um ajuste melhor, mas não necessariamente será preferível em termos de critério de informação, (EHLERS, 2007). Ou seja, para que haja a inclusão de parâmetros extras, vê-se a necessidade destes melhorarem sensivelmente o ajuste do modelo e, caso não ocorra, o critério de informação se mostrará maior. A regra básica consiste em selecionar o modelo cujo critério de informação calculado seja mínimo.

O critério de informação de Akaike, denotado por AIC, é um dos critérios de informação mais comumente utilizado para seleção de modelos matemáticos de previsão. A definição a ser utilizada é:

$$AIC = -2\hat{l}(\underline{\theta}) + 2m, \quad (2.24)$$

onde $\hat{l}(\underline{\theta})$ é a *log*-verossimilhança maximizada de $\underline{\theta}$ e m é o número de parâmetros. Para dados normalmente distribuídos e usando-se estimativas de máxima verossimilhança para os parâmetros pode-se mostrar que:

$$AIC = n\log\hat{\sigma}_\epsilon^2 + 2m. \quad (2.25)$$

Assumindo que os erros possuam uma função de distribuição de variância $\hat{\sigma}_\epsilon^2 = \frac{1}{n} \sum \hat{\epsilon}_t^2$.

Uma das modificações do AIC ocasionou o surgimento do AICc (AIC corrigido), que pode aumentar apreciavelmente a probabilidade de se escolher um modelo adequado, particularmente na seleção de modelos de regressão e de séries temporais,(BOZDOGAN, 1987). Sua expressão é dada por:

$$AICc = -2\hat{l}(\underline{\theta}) + 2m + 2\frac{m(m+1)}{n-m-1}, \quad (2.26)$$

Outro critério é o de informação Bayesiano, denotado por BIC e dado por:

$$BIC = -2\hat{l}(\underline{\theta}) + m\log(n). \quad (2.27)$$

Para a utilização de tais critérios na comparação de modelos, a estimação precisa ser feita no mesmo período amostral de maneira que os modelos sejam comparáveis.

2.4 Estatística Bayesiana

A análise Bayesiana a cerca dos dados é uma alternativa importante aos procedimentos clássicos de modelagem, estimação e previsão de dados, que tem tido uma crescente aplicação em problemas de várias áreas. Consiste num procedimento genérico de inferência a partir de dados, utilizando modelos probabilísticos para descrever variabilidade em quantidades observadas ou descrever graus de incertezas em quantidades sobre as quais se quer compreender (geralmente os parâmetros dos modelos ou funções destes parâmetros).

Uma abordagem de muito interesse na análise de séries temporais é a bayesiana. Antelman (1997) esclarece que a inferência bayesiana é constituída por três estruturas básicas: a distribuição a priori, que indica o estado atual de informação do pesquisador, ou seja representa o que é conhecido adicionalmente ao experimento antes da observação dos dados; a função de verossimilhança, que expressa todo o conhecimento do experimento contido nos dados, ou seja, codifica toda a informação relevante contida nos dados sobre o parâmetro; e a distribuição a posteriori, que representa o conhecimento sobre o experimento atualizado pelos dados, especificando o estado da informação sobre o parâmetro de interesse, após a observação dos dados. Silva e Suárez (2000) destacam que a abordagem clássica é principalmente empírica utilizando somente a informação amostral como base para estimar e testar hipóteses a respeito de parâmetros populacionais, enquanto que a abordagem bayesiana utiliza toda informação amostral além do julgamento pessoal para escolha de uma distribuição de probabilidade a priori adequada para o cálculo de estimativas. Box e Tiao (2011) discutiram as idéias de Jeffreys (1967), sobre a distribuição a priori para representar o estado de ausência de informação ou ignorância a respeito do comportamento probabilístico dos parâmetros.

2.4.1 Inferência Bayesiana

O rápido crescimento do uso do paradigma bayesiano em ciências aplicadas ao longo das últimas décadas, foi facilitado pelo surgimento e disponibilidade de vários algoritmos computacionais usados no cálculo de integrais que são necessárias em uma análise bayesiana.

Na perspectiva bayesiana, a inferência estatística sobre qualquer quantidade de interesse é descrita como a modificação que se processa nas incertezas à medida que são observados novos dados ou novos resultados. É o Teorema de Bayes que permite quantificar esta modificação.

A inferência Bayesiana paramétrica, assim como a frequentista, considera que um conjunto de observações y pode ser descrito por uma distribuição de probabilidade $f(y|\theta)$, em que θ é o vetor desconhecido de parâmetros. Pode-se dizer que um dos principais objetivos da estatística é fazer inferências sobre θ e muitas vezes o pesquisador possui alguma informação a priori a respeito dessa quantidade. A teoria Bayesiana diferencia-se da frequentista por permitir que essas informações prévias a respeito dos parâmetros de interesse para a análise sejam incorporadas ao modelo. Assim, a teoria Bayesiana permite a incorporação da incerteza relativa às quantidades desconhecidas na análise dos dados, enquanto a frequentista não.

A abordagem Bayesiana necessita, além da informação sobre os dados, que é incorporada na análise por meio da função de verossimilhança, de uma distribuição a priori para as quantidades desconhecidas no modelo (parâmetros e dados faltantes). Essas duas informações são usadas para obter a distribuição dos parâmetros, condicional aos dados observados, chamada distribuição a posteriori, da qual resultam todas as inferências estatísticas sobre os parâmetros do modelo (CARLIN; LOUIS, 1997).

2.4.2 Análise Bayesiana em Séries Temporais

Uma importante etapa na análise de séries temporais meteorológicas é a estimativa dos parâmetros do modelo, através da determinação da distribuição a priori da série. Entre os métodos para determinação dos parâmetros temos aquele comumente chamado de clássico e também um novo método que vêm despertando interesse da comunidade científica, o chamado método da inferência bayesiana (LIMA, 2013).

De acordo com Box e Tiao (2011) o tratamento estatístico das séries temporais meteo-

rológicas através da metodologia bayesiana durante muitas décadas oscilou dentre rejeição e a aceitação. Contudo a técnica se consolidou nas pesquisas científicas fundamentalmente por alguns fatores principais:

- i) Pode-se afirmar que o processo de aquisição de dados para que uma pesquisa seja realizada além de ser laborioso também é oneroso, e por isso espera-se resultados qualitativamente e quantitativamente adequados. No entanto, enquanto outras teorias de inferência produziam soluções viáveis em casos onde hipóteses especiais, tais como normalidade e independência de erros, em outros casos, tais como, onde não existiam estatísticas suficientes, as soluções eram, frequentemente, insatisfatórias e confusas. Embora seja verdadeiro que estas hipóteses especiais cubram um número de situações de interesse científico, seria inútil achar que o conjunto dos problemas estatísticos cuja solução tenha sido ou viria a ser necessária ao pesquisador coincida com o conjunto dos problemas e desta maneira tratado de forma conveniente. É razoável, então, que esses dados sejam analisados sob diferentes pontos de vista. Na seleção dos pontos de vista, os métodos bayesianos permitem que uma ênfase maior seja dada ao interesse científico e uma ênfase menor à conveniência matemática.
- ii) Os métodos para resolução de problemas matemáticos durante as últimas décadas vêm sofrendo um processo de substituição, no qual, as calculadoras caíram em desuso, dando espaço ao computador. Logo, as soluções viáveis baseadas nas hipóteses especiais tornaram-se populares por serem fáceis de computar.

Com o surgimento dos computadores, o tempo computacional gasto na busca de cada solução foi reduzido e um número maior de análises pode ser realizado. Segundo Box e Tiao (2011), na teoria da inferência bayesiana os parâmetros de um modelo são considerados variáveis aleatórias, para os quais temos alguma informação a priori, resumida na densidade de probabilidade a priori, a qual expressa a incerteza sobre estes parâmetros. Essa distribuição desempenha um papel fundamental na análise bayesiana. Diferentemente da análise clássica a inferência bayesiana possui uma estrutura compacta que possibilita a abatimento do esforço computacional, agilizando assim o trabalho estatístico.

2.4.3 Teorema de Bayes

Tomemos a definição de Box e Tiao (1973), com relação a natureza da inferência bayesiana. Suponha que $y' = (y_1, \dots, y_n)$ é um vetor n observações cuja distribuição

de probabilidade $p(y|\theta)$ depende dos k valores do parâmetro $\theta' = (\theta_1, \dots, \theta_k)$. Suponha também que a distribuição de probabilidade de θ é $p(\theta)$. Então $p(y|\theta)p(\theta) = p(y, \theta) = p(\theta|y)p(y)$.

Dado y observado, a distribuição condicional de θ é:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}. \quad (2.28)$$

Também, podemos escrever:

$$p(y) = E p(y|\theta) = c^{-1} = \begin{cases} \int p(y|\theta)p(\theta)d\theta, & \text{Se } \theta \text{ contínuo,} \\ \sum p(y|\theta)p(\theta), & \theta \text{ discreto.} \end{cases} \quad (2.29)$$

onde a soma ou a integral é tomada sobre a amplitude admissível de θ e onde $E[f(\theta)]$ é a esperança matemática de $f(\theta)$ com relação a distribuição $p(\theta)$. Assim podemos escrever 2.28 alternativamente como

$$p(\theta|y) = c p(y|\theta)p(\theta). \quad (2.30)$$

A condição em 2.28, ou seu equivalente em 2.30 é referido com o Teorema de Bayes. Nesta expressão, $p(\theta)$ diz o que é conhecido a respeito de θ sem o conhecimento dos dados e chamada de distribuição a priori de θ . Correspondentemente, $p(\theta|y)$ diz o que é conhecido dos dados e chamada de distribuição a posteriori de θ para y . A quantidade de c é simplesmente uma constante de “normalização” necessária para assegurar-se de que a distribuição a posteriori $p(\theta|y)$ integrada ou somada (em caso discreto) é 1.

Para o Teorema de Bayes e a função de verossimilhança tomemos as observações dadas y , $p(y|\theta)$ em 2.30. Elas podem ser consideradas como uma função não de y , mas de θ . Quando for assim considerada, será função de verossimilhança de θ dado y e pode ser escrito $l(\theta|y)$. Podemos assim escrever a fórmula de Bayes como:

$$p(\theta|y) = K.l(\theta|y)p(\theta) \quad (2.31)$$

Em outras palavras, o Teorema de Bayes diz que a distribuição de probabilidade de θ a posteriori para y , é proporcional ao produto entre a distribuição priori de θ e a verossimilhança de θ dado y . Isto é,

Distribuição a posteriori \propto distribuição a priori \times verossimilhança.

Quanto à natureza sequencial do Teorema de Bayes, suponhamos que temos uma amostra inicial y_1 e assim temos pela fórmula:

$$p(\theta|y_1) \propto p(\theta)L(\theta|y_1), \quad (2.32)$$

supondo uma segunda amostra y_2 , distribuída independentemente da primeira

$$p(\theta|y_2) \propto p(\theta)L(\theta|y_2). \quad (2.33)$$

Obviamente este processo pode ter n repetições independentes que a distribuição a posteriori será recalculada a cada nova observação, de modo que no m -ésimo estágio a verossimilhança associada à m -ésima observação seja combinada com a posteriori de θ depois de $m - 1$ observações:

$$p(\theta|y_1, \dots, y_m) \propto p(\theta|y_1, \dots, y_{m-1})L(\theta|Y_m), m = 2, \dots, n \quad (2.34)$$

onde

$$p(\theta|y_1) \propto p(\theta)L(\theta|y_1).$$

2.4.4 Distribuição a priori

A distribuição a priori constitui-se no único elemento novo na análise bayesiana em relação à análise clássica. A premissa desta distribuição é que ela reflita a incerteza sobre os possíveis valores dos parâmetros antes da coleta de dados. A seguir serão apresentadas algumas formas de especificação da distribuição a priori.

2.4.4.1 Priori não informativa

Existem situações em que o conhecimento sobre determinado fenômeno é vago ou inexistente. Nestes casos, a distribuição a priori é dita vaga, difusa ou não informativa, não precisa ou de variância alta, significando que a densidade a priori reflete a ignorância, sendo caracterizada pela ausência total ou quantidade mínima de informação, pois não privilegia nenhum valor do parâmetro sobre qualquer outro (JEFFREYS, 1998). Também utiliza-se esse tipo de priori quando se espera que os dados transmitam toda a informação necessária, isto é, a informação dos dados seja dominante.

Método de Jeffreys

A crítica da inconsistência da distribuição uniforme na representação formal da ig-

norância implica que esta deve ser invariante sob transformações injetivas. Entre os procedimentos que asseguram esta invariância está aquele proposto por Jeffreys, sendo baseado no uso da informação esperada de Fisher sobre $\theta \in \mathfrak{R}$.

$$I(\theta) = E \left[\left(\frac{\partial \log L(x|\theta)}{\partial \theta} \right)^2 \middle| \theta \right] = -E \left[\left(\frac{\partial^2 \log L(x|\theta)}{\partial \theta^2} \right) \middle| \theta \right]$$

Seja $\psi = g(\theta)$ e $\theta = g^{-1}(\psi)$, então

$$\begin{aligned} I(\psi) &= -E \left[\left(\frac{\partial^2 \log L(x|\psi)}{\partial \psi^2} \right) \right] \\ &= -E \left[\left(\frac{\partial^2 \log L(x|\theta=g^{-1}(\psi))}{\partial \theta^2} \right) \right] \left| \frac{\partial g^{-1}(\psi)}{\partial \psi} \right|^2 \\ &= I(\theta) \left| \frac{d\theta}{d\psi} \right|^2 \end{aligned}$$

De fato, seja $h(\theta)$ uma distribuição a priori para θ , então

$$h(\psi) = h[\theta(\psi)] \left| \frac{d\theta}{d\psi} \right| \quad (2.35)$$

deve ser a correspondente distribuição para a reparametrização injetiva ψ , que não é necessariamente uniforme quando $h(\theta)$ o é (considere transformações não lineares). Extraíndo a raiz quadrada nesta última igualdade e, em seguida comparando com 2.35, tem-se

$$h(\theta) \propto |I(\theta)|^{\frac{1}{2}} \quad (2.36)$$

e assim pode-se enunciar a priori de Jeffreys para o caso uniparamétrico.

2.4.4.2 Priori conjugada

Existem situações práticas em que há informação a priori mais ou menos substancial sobre os parâmetros do modelo, quer por parte do pesquisador ou por parte de outros indivíduos a quem o pesquisador pode recorrer. A questão de como obter e quantificar essa informação de modo a transformá-la em uma distribuição a priori que possa ser utilizada para prosseguir com a metodologia bayesiana é um tema que vem sendo amplamente tratado na literatura científica. A escolha da densidade a priori de um parâmetro $\theta \in \Theta$ deve ser baseada em questões relacionadas diretamente a ele e isso depende do tipo de parâmetro, dos conhecimentos do pesquisador, entre outros.

Uma densidade a priori é dita conjugada com função de verossimilhança se a distribuição a posteriori resultante é da mesma família de distribuições da priori.

Uma vantagem do uso de priori conjugadas é que, em geral, facilita a análise visto que as distribuições a priori e a posteriori estão em uma mesma família de distribuições e uma desvantagem, é que nem sempre a distribuição utilizada representa bem a incerteza a priori.

Se $F = \{p(y|\theta), \theta \in \Theta\}$ é uma classe de distribuições amostrais, então uma classe de distribuições P é conjugada a F se $\forall p(y|\theta) \in F$ e $p(\theta) \in P \Rightarrow p(\theta|y) \in P$.

Uma família de distribuições conjugadas pode ser obtida considerando:

- a. Identificar a classe P de distribuições para θ tal que $l(\theta|y)$ seja proporcional a um membro desta classe,
- b. Verificar se P é fechada por amostragem, isto é, $\forall p_1, p_2 \in P \exists k$ tal que $kp_1p_2 \in P$.

Se além disso existir uma constante k tal que $k^{-1} = \int l(\theta, y)d\theta < \infty$ e todo $p \in P$ é definido como $p(\theta) = kl(\theta|y)$ então P é a família conjugada natural ao modelo amostral gerador de $l(\theta; x)$

Existe também uma outra forma de conjugação de priori, que são as priori conjugadas na família exponencial, bastante utilizadas por incluir grande parte das distribuições contínuas e discretas. Para maiores detalhes desta forma de conjugação, ver (EHLERS, 2007).

2.4.5 Monte Carlo com Cadeias de Markov

O algoritmo de Metropolis-Hastings é um método iterativo usado para gerar amostras de uma densidade de forma indireta, quando não se sabe gerar diretamente dessa densidade (CHIB; GREENBERG, 1995). O algoritmo consiste em escolher um núcleo $q(\cdot|\theta)$, gerar deste núcleo, usando um critério de aceitação do valor gerado para garantir que a amostra obtida é representativa da amostra gerada.

Seja $\pi(\theta)$ a densidade de probabilidade conjunta de interesse com $\theta = (\theta_1, \dots, \theta_k)$ e suponha que as densidade condicionais completas a posteriori $\pi_i(\theta_i) = \pi_i(\theta_i|\theta_{-i})$, $i = 1, 2, \dots, k$, estejam disponíveis.

O algoritmo é dado por:

Passo 1: Inicialize o contador de iterações da cadeia em $j = 1$ e escolha os valores iniciais

$$\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)});$$

Passo 2: Gere um novo valor de θ^* da distribuição de $q(\cdot|\theta)$;

Passo 3: Calcule a probabilidade de aceitação $\alpha(\theta, \theta^*)$ e gere $u \sim U(0, 1)$;

Passo 4: Se $u \leq \alpha$ aceite o novo valor e faça $\theta^{(j+1)} = \theta^*$, caso contrário rejeite e faça $\theta^{(j+1)} = \theta$;

Passo 5: Atualizar o contador de j para $j + 1$ e repetir o passo 2 até a convergência, no qual o novo valor

$$\alpha(\theta, \theta^*) = \min \left\{ 1, \frac{\pi(\theta^*)q(\theta|\theta^*)}{\pi(\theta)q(\theta^*|\theta)} \right\}. \quad (2.37)$$

O conhecimento da densidade de probabilidade $\pi(\theta)$ pode ser parcial, isto é, a menos de uma constante a probabilidade em 2.37 não se altera. A cadeia pode permanecer no mesmo estado por muitas iterações e na prática é comum monitorar a porcentagem das iterações para as quais os novos valores são aceitos.

Para verificar a convergência da cadeia, pode-se utilizar técnicas gráficas como: gráficos de séries temporais, em que a série deve estar em uma faixa, sem apresentar tendências nem oscilações bruscas e gráficos de autocorrelação dos parâmetros, sendo esperado que a primeira correlação seja alta e as demais próximas de zero, indicando que a amostra simulada é não correlacionada. Uma outra avaliação de convergência é feita usando-se técnicas numéricas.

Os diagnósticos de convergência mais utilizados são descritos por Gelman e Rubin (1992), em que consideram a análise de variância para verificar a convergência de cadeias paralelas (com valores iniciais super dispersos) e o critério Geweke et al. (1991), que é baseado em métodos de séries temporais. Neste trabalho, os métodos de MCMC apresentados são utilizados na obtenção de estimativas para quantidades de interesse do modelo.

Para o caso em que as matrizes de transição¹ são simétricas, pode-se simplificar o algoritmo de Metropolis-Hastings para o algoritmo de Metropolis, em que a única diferença é que a probabilidade de aceitação é dada por:

$$\alpha(\theta, \theta^*) = \min \left\{ 1, \frac{\pi(\theta^*)}{\pi(\theta)} \right\}. \quad (2.38)$$

A utilização da distribuição uniforme para servir como matriz de transição é apenas

¹Matriz de transição é um processo estocástico utilizado para estudar fenômenos que passam, a partir de um estado inicial, por uma sequência de ensaios, onde a transição de um determinado estado ocorre segundo uma certa probabilidade

uma entre muitas opções. A distribuição normal centrada em θ_i é uma outra alternativa que pode ser usada.

2.4.5.1 Amostrador de Gibbs

Em muitas situações a geração de uma amostra pode ser complicada ou simplesmente impossível, no entanto, se as condicionais apresentam formas conhecidas (normal, gama, beta), então o método de *Gibbs Sampling* é uma alternativa de geração de amostras. O *Gibbs Sampling* é um esquema iterativo de amostragem de cadeia de Markov cuja matriz de transição é constituída das distribuições condicionais completas, fornecendo, então, uma alternativa de geração baseada em sucessivas gerações das distribuições condicionais completas (CASELLA; GEORGE, 1992).

Seja $\pi(\theta)$ a densidade de probabilidade conjunta de interesse com $\theta = (\theta_1, \dots, \theta_k)$ e suponha que as densidades condicionais completas a posteriori $\pi_i(\theta_i) = \pi_i(\theta_i | \theta_{-i})$, $i = 1, 2, \dots, k$, sejam conhecidas.

Então, o algoritmo é dado por:

Passo 1: Inicialize o contador de iterações da cadeia de $j = 1$ e escolha os valores iniciais

$$\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)});$$

Passo 2: Obtenha um novo valor $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_k^{(j)})$ a partir de $\theta^{(j-1)}$ através de sucessivas gerações de valores

$$\begin{aligned} \theta_1^{(j)} &\sim (\theta_1 | \theta_2^{(j-1)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)}) \\ \theta_2^{(j)} &\sim (\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)}) \\ &\vdots \\ \theta_k^{(j)} &\sim (\theta_k | \theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_{k-1}^{(j)}); \end{aligned}$$

Passo 3: Atualizar o contador de j para $j + 1$ e repetir o passo 2 até a convergência.

À medida que o número de iterações aumenta, a cadeia se aproxima de sua condição de equilíbrio. Podemos assumir que a convergência é atingida em uma iteração quando a Cadeia de Markov possui propriedades de regularidade, tais como, reversibilidade e pode-se garantir que o processo iterativo converge para a distribuição de equilíbrio $\pi(\theta)$.

2.5 Modelos para série temporal número de dias com precipitação

2.5.1 Modelo Clássico

Quando se consideram dados na forma de contagens que apresentam diversas medidas realizadas em um mesmo indivíduo ao longo do tempo, utiliza-se a teoria de modelos INAR. Assim surge o interesse em estudar uma série com número de dias com precipitação através do modelo de Poisson.

2.5.2 Modelo Poisson

Supondo, inicialmente que os Y_{ij} são independentes, com distribuição de Poisson com parâmetros λ_{ij} , isto é, $Y_{ij} \sim Poisson(\lambda_{ij})$, $\lambda_{ij} > 0$.

$$P(Y_{ij} = y_{ij}) = \frac{e^{-\lambda_{ij}} \lambda_{ij}^{y_{ij}}}{y_{ij}!}, y_{ij} = 0, 1, 2, \dots \quad (2.39)$$

e admitindo-se que $\ln(\lambda_{ij}) = x'_{ij}\beta$, tem-se que:

$$E(Y_{ij}|x_{ij}\beta) = \lambda_{ij} = \exp(x'_{ij}\beta)$$

e

$$Var(Y_{ij}|x_{ij}\beta) = \lambda_{ij} = \exp(x'_{ij}\beta).$$

2.5.3 Modelo Poisson-Normal

Considerando-se a distribuição de Poisson com função densidade de probabilidade dada por 2.39 para Y_{ij} e a distribuição normal e para o efeito aleatório b_i para incorporar a correlação entre as medidas longitudinais, tem-se:

$$Y_{ij}|\lambda_{ij} \sim Poisson(\lambda_{ij})$$

$$\ln(\lambda_{ij}) = x'_{ij}\beta + z'_{ij}b_i$$

$$b_i \sim N(0, \mathbf{D})$$

sendo que \mathbf{D} é a matriz de variâncias de b_i . Logo, tem-se:

$$\begin{aligned}
E(Y_{ij}) &= E[E(Y_{ij}|\lambda_{ij})] \\
&= E[\exp(x'_{ij}\beta + z'_{ij}b_i)] \\
&= \exp(x'_{ij}\beta)E[\exp(z'_{ij}b_i)] \\
&= \exp\left(x'_{ij}\beta + \frac{1}{2}z'_{ij}\mathbf{D}z_{ij}\right)
\end{aligned} \tag{2.40}$$

e

$$\begin{aligned}
\text{Var}(Y_{ij}) &= E[\text{Var}(Y_{ij}|\lambda_{ij})] + \text{Var}[E(Y_{ij}|\lambda_{ij})] \\
&= E[\exp(x'_{ij}\beta + z'_{ij}b_i)] + \text{Var}[\exp(x'_{ij}\beta + z'_{ij}b_i)] \\
&= \exp\left(x'_{ij}\beta + \frac{1}{2}z'_{ij}\mathbf{D}z_{ij}\right) + E[(x'_{ij}\beta + z'_{ij}b_i)^2] - E[(x'_{ij}\beta + z'_{ij}b_i)]^2 \\
&= \exp\left(x'_{ij}\beta + \frac{1}{2}z'_{ij}\mathbf{D}z_{ij}\right) + \exp(2x'_{ij}\beta + 2z'_{ij}\mathbf{D}z_{ij}) - \{\exp(x'_{ij}\beta + \frac{1}{2}z'_{ij}\mathbf{D}z_{ij})\} \\
&= \exp\left(x'_{ij}\beta + \frac{1}{2}z'_{ij}\mathbf{D}z_{ij}\right) + \exp(2x'_{ij}\beta + z'_{ij}\mathbf{D}z_{ij})\{\exp(z'_{ij}\mathbf{D}z_{ij}) - 1\}.
\end{aligned} \tag{2.41}$$

3 Material e Métodos

Referenciando-se nos trabalhos anteriores objetivamos estudar o comportamento das séries de número diário de precipitações (valores inteiros não negativos) referente ao município de Garanhuns, agreste pernambucano. Utilizou-se para isso, técnicas para séries temporais de contagem, afim de modelar o comportamento e prever esse número, comparando-o com o modelo clássico de Poisson e contribuindo na elaboração de políticas públicas ambientais.

Os dados foram ajustados em cinco modelos autoregressivos para valores inteiros, dentre os quais quatro tiveram abordagem bayesiana e um com abordagem clássica. Utilizou-se a distribuição Poisson e Poisson-Normal, combinadas a prioris Normal e Beta.

A seleção dos modelos se deu através dos critérios AIC, AICc e BIC, não para escolher a ordem e sim qual modelo melhor se ajustou aos dados. Esses são critérios de informação que baseiam-se fundamentalmente na qualidade do ajuste, através da quantidade de termos de n , mas também penalizam a inclusão de parâmetros extras.

3.1 Descrição dos dados

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE, 2013), o município de Garanhuns, dista 288 km da capital Recife e é constituído de 135.138 habitantes, correspondendo a 22% da população do agreste meridional de Pernambuco. Sua área é de 458.552 km^2 , sendo o nono maior município pernambucano e um dos principais pólos de turismo do Estado (Figura 3.1).



Figura 3.1: Disposição geográfica do Município de Garanhuns-PE

Utilizou-se uma série de tempo obtida da estação meteorológica n° 82893 - Garanhuns - PE, relacionada ao número de dias com precipitação, localizada no nordeste do Brasil, com Latitude: -8,88, Longitude: -36,51, altitude: 822,76 m, a partir de 30/11/1993 à 27/02/2012. Os valores de número de dias de precipitação por mês, do Instituto Nacional de Meteorologia (INMET, 2014).

O objetivo é modelar o quadro de número diário de precipitação no município de Garanhuns. O conjunto observado forma uma série temporal Y_t . Os *softwares* utilizados neste trabalho foram o *software* R, Minitab 16 e Winbugs 1.4.

3.2 Inferência bayesiana para processos INAR(1) de Poisson

Considerando um processo INAR(1), se se pretender obter uma distribuição para a distribuição de $\{Y_t\}$ é necessário que $\{\epsilon_t\}$ também tenha uma distribuição de Poisson. Mais concretamente, $\{Y_t\}$ tem distribuição de Poisson com média $\lambda/(1 - \alpha)$ se somente se $\{\epsilon_t\}$ tem distribuição de Poisson de média λ . Portanto, a distribuição de Poisson tem um papel análogo ao da distribuição do modelo AR. Considera-se o processo autoregressivo de valor inteiro de primeira ordem INAR(1) dado por:

$$y_t = \alpha \circ y_{t-1} + \epsilon_t, t \geq 2, \quad (3.1)$$

onde a operação “ \circ ” é a operação *thinning* binomial, $\alpha \in]0, 1]$ e $\{\epsilon_t\}$ é uma sucessão de variáveis aleatórias de Poisson de parâmetros λ , não correlacionadas e independentes de y_{t-1} .

Sob estas condições e dado y_1 a função verossimilhança da amostra $y = (y_2, \dots, y_n)$ é dada por,

$$\begin{aligned} l(y, \alpha, \lambda | y_1) &= P(Y_2 = y_2, \dots, Y_n = y_n | y_1) = \prod_{t=2}^n P(Y_t = y_t | Y_{t-1} = y_{t-1}) \\ &= \prod_{t=2}^n P(\alpha \circ Y_{t-1} + \epsilon_t = y_t | Y_{t-1} = y_{t-1}) = \prod_{t=2}^n p(y_t | y_{t-1}). \end{aligned}$$

A variável $Y_t | Y_{t-1}$ é a convolução da distribuição binomial de parâmetros Y_{t-1} e α com a distribuição de Poisson de parâmetro λ , portanto a sua função massa de probabilidade é dada por,

$$\begin{aligned} p(y_t | y_{t-1}) &= \sum_{i=0}^{\min(y_t, y_{t-1})} P(\epsilon_t = y_t - i) \times P(\alpha \circ Y_{t-1} = i | Y_{t-1} = y_{t-1}) \\ &= \sum_{i=0}^{\min(y_t, y_{t-1})} e^{-\lambda} \frac{\lambda^{y_t - i}}{(y_t - i)!} \times \sum_{i=0}^{\min(y_t, y_{t-1})} C_i^{y_{t-1}} \alpha^i (1 - \alpha)^{y_{t-1} - i}. \end{aligned}$$

Deste modo, a função de verossimilhança condicional a y_1 é dada por:

$$l(y, \alpha, \lambda | y_1) = \prod_{t=2}^n e^{-\lambda} \sum_{i=0}^{M_t} \frac{\lambda^{y_t - i}}{(y_t - i)!} C_i^{y_{t-1}} \alpha^i (1 - \alpha)^{y_{t-1} - i}, \quad (3.2)$$

onde $M_t = \min(y_t, y_{t-1})$, $t = 2, \dots, n$.

Consideremos a distribuição beta como distribuição a priori para o parâmetro α e a distribuição gama como distribuição a priori para o parâmetro λ , isto é,

$$h(\alpha) \propto \alpha^{a-1} (1 - \alpha)^{b-1} \quad , \quad \alpha \sim Be(a, b), a, b > 0$$

$$h(\lambda) \propto \lambda^{c-1} \exp(-d\lambda) \quad , \quad \lambda \sim Ga(c, d), c, d > 0.$$

A escolha das distribuições de beta e gama para as distribuições a priori dos parâmetros prende-se com o fato de serem conjugadas da binomial e poisson, respectivamente.

Supondo α e λ independentes, a distribuição a priori conjunta é dada por,

$$h(\alpha, \lambda) \propto \lambda^{c-1} \exp(-d\lambda) \alpha^{a-1} (1 - \alpha)^{b-1}, \quad (3.3)$$

onde $\lambda > 0$, $0 < \alpha < 1$ e os hiperparâmetros a, b, c e d são conhecidos e positivos. Nota-se que se $a \rightarrow 0$, $b \rightarrow 0$, $c \rightarrow 0$ e $d \rightarrow 0$ tem-se o caso de uma distribuição a priori não informativa.

Assim, a distribuição a posteriori conjunta é dada por,

$$h(\alpha, \lambda|y) \propto \exp[-(d+n)\lambda] \lambda^{c-1} \alpha^{a-1} (1-\alpha)^{b-1} \times \prod_{t=2}^n \sum_{i=0}^{\min(y_t, y_{t-1})} \frac{\lambda^{y_t-i}}{(y_t-i)!} C_i^{y_{t-1}} \alpha^i (1-\alpha)^{y_{t-1}-i} \quad (3.4)$$

Integrando a distribuição marginal a posteriori para λ em ordem α é dada por:

$$h(\lambda|y) \propto \int \left[\exp[-(d+n)\lambda] \lambda^{c-1} \alpha^{a-1} (1-\alpha)^{b-1} \times \prod_{t=2}^n \sum_{i=0}^{\min(y_t, y_{t-1})} \frac{\lambda^{y_t-i}}{(y_t-i)!} C_i^{y_{t-1}} \alpha^i (1-\alpha)^{y_{t-1}-i} \right] d\alpha. \quad (3.5)$$

Integrando a distribuição marginal a posteriori para α em ordem λ é dada por:

$$h(\alpha|y) \propto \int \left[\exp[-(d+n)\lambda] \lambda^{c-1} \alpha^{a-1} (1-\alpha)^{b-1} \times \prod_{t=2}^n \sum_{i=0}^{\min(y_t, y_{t-1})} \frac{\lambda^{y_t-i}}{(y_t-i)!} C_i^{y_{t-1}} \alpha^i (1-\alpha)^{y_{t-1}-i} \right] d\lambda. \quad (3.6)$$

Como se pode verificar estas integrais são muito complexas e portanto é necessário utilizar a metodologia de Gibbs para obter as estimativas de λ e α . Para tal devem-se calcular as distribuições condicionais completas para os parâmetros λ e α

A distribuição condicional completa para o parâmetro λ é

$$h(\lambda|\alpha, y) \propto \exp[-(d+n)\lambda] \lambda^{c-1} \times \prod_{t=2}^n \sum_{i=0}^{M_t} L(t, i) \lambda^{y_t-i}, \quad (3.7)$$

onde

$$L(t, i) = \frac{1}{(y_t-i)!} C_i^{y_{t-1}} \alpha^i (1-\alpha)^{y_{t-1}-i}, \lambda > 0. \quad (3.8)$$

A distribuição dada em 3.7 é uma combinação linear de funções de densidade de probabilidade de variáveis aleatórias com distribuição gama.

Analogamente a distribuição condicional completa para α é dada por:

$$h(\alpha|\lambda, y) \propto \alpha^{a-1} (1-\alpha)^{b-1} \times \prod_{t=2}^n \sum_{i=0}^{M_t} K(t, i) \alpha^i (1-\alpha)^{y_{t-1}-i}, \quad (3.9)$$

onde

$$K(t, i) = \frac{\lambda^{y_t-i}}{(y_t - i)!} C_i^{y_{t-1}}, 0 < \alpha < 1.$$

A distribuição dada em 3.9 é uma combinação linear de funções densidade de probabilidade de variáveis aleatórias com distribuição beta.

Agora que são conhecidas as distribuições condicionais completas para os parâmetros pode-se utilizar o algoritmo de Gibbs para determinar as estimativas.

3.3 Inferência bayesiana para processos INAR(1) de Poisson-Normal

Considera-se o processo autoregressivo de valor inteiro de primeira ordem INAR(1) dada por:

$$y_t = \alpha \circ y_{t-1} + \epsilon, t \geq 2, \quad (3.10)$$

onde a operação “ \circ ” é a operação *thinning* binomial, $\alpha \in]0, 1]$ e $\{\epsilon\}$ é uma sucessão de variáveis aleatórias binomial-normal com média $\exp(x'_{ij}\beta + \frac{1}{2}z'_{ij}Dz_{ij})$ não correlacionados e independentes de y_{t-1} .

Sob estas condições e dado y_1 a função verossimilhança da amostra $y = (y_2, \dots, y_n)$ é dada por,

$$\begin{aligned} l(y, \alpha, \lambda | y_1) &= P(Y_2 = y_2, \dots, Y_n = y_n | y_1) = \prod_{t=2}^n P(Y_t = y_t | Y_{t-1} = y_{t-1}) \\ &= \prod_{t=2}^n P(\alpha \circ Y_{t-1} + \epsilon_t = y_t | Y_{t-1} = y_{t-1}) = \prod_{t=2}^n p(y_t | y_{t-1}). \end{aligned}$$

A variável $Y_t | Y_{t-1}$ é a convolução da distribuição binomial de parâmetros Y_{t-1} e α com a distribuição de Poisson de parâmetro λ , portanto a sua função massa de probabilidade é dada por,

$$\begin{aligned} p(y_t | y_{t-1}) &= \sum_{i=0}^{\min(y_t, y_{t-1})} P(\epsilon_t = y_t - i) \times P(\alpha \circ Y_{t-1} = i | Y_{t-1} = y_{t-1}) \\ &= \sum_{i=0}^{\min(y_t, y_{t-1})} p(1-p)^{y_{t-1}} \frac{\lambda^{y_t-i}}{(y_t - i)!} \times \sum_{i=0}^{\min(y_t, y_{t-1})} C_i^{y_{t-1}} \alpha^i (1-\alpha)^{y_{t-1}-i}. \end{aligned}$$

Deste modo, a função de verossimilhança condicional a y_1 é dada por:

$$l(y, \alpha, \lambda|y_1) = \prod_{t=2}^n \sum_{i=0}^{M_t} (1-p)^{y_t-i} C_i^{y_t-1} \alpha^i (1-\alpha)^{y_t-1-i}, \quad (3.11)$$

onde $M_t = \min(y_t, y_{t-1})$, $t = 2, \dots, n$.

Considerando novamente a distribuição de beta como distribuição a priori para os parâmetros α e p , tem-se

$$h(p) \propto p^{a-1}(1-p)^{b-1} \quad , \quad p \sim Be(a, b), a, b > 0$$

$$h(\alpha) \propto \alpha^{c-1}(1-\alpha)^{d-1} \quad , \quad \alpha \sim Ga(c, d), c, d > 0.$$

A escolha da distribuição beta e gama para distribuição a priori dos parâmetros prende-se, novamente com o fato de ser a conjugada da binomial e da normal.

Supondo α e p independentes, a distribuição conjunta é dada por,

$$h(\alpha, p) \propto p^{a-1}(1-p)^{b-1} \alpha^{c-1}(1-\alpha)^{d-1}, \quad (3.12)$$

onde $0 < \alpha < 1$ e os hiperparâmetros a, b, c e d são conhecidos e positivos.

Assim, a distribuição conjunta, calculada de acordo com o teorema de Bayes é dada por,

$$h(\alpha, p|y) \propto p^{a-1}(1-p)^{b-1} \alpha^{c-1}(1-\alpha)^{d-1} \times \prod_{t=2}^n \sum_{i=0}^{M_t} p(1-p)^{y_t-i} C_i^{y_t-1} \alpha^i (1-\alpha)^{y_t-1-i} \quad (3.13)$$

Analogamente ao verificado na seção anterior, também não se obtém formas fechadas para as distribuições a posteriori marginais de p e λ daí a necessidade da obtenção das distribuições condicionais completas. Assim a distribuição condicional completa para o parâmetro p é

$$h(p|\alpha, y) \propto n p^{a-1}(1-p)^{b-1} \times \prod_{t=2}^n \sum_{i=0}^{M_t} (1-p)^{y_t-i} N(t, i) \quad (3.14)$$

onde

$$N(t, i) = C_i^{y_t-1} \alpha^i (1-\alpha)^{y_t-1-i}$$

A distribuição dada em 3.14 é uma combinação linear de funções de densidade de probabilidade de variáveis aleatórias com distribuição beta.

Do mesmo modo, a distribuição condicional completa para α é dada por,

$$h(\alpha|p, y) \propto \alpha^{c-1}(1-\alpha)^{d-1} \times \prod_{t=2}^n \sum_{i=0}^{M_t} M(t, i) \alpha^i (1-\alpha)^{y_{t-1}-i}, \quad (3.15)$$

onde

$$M(t, i) = (1-p)^{y_t-i} C_i^{y_{t-1}}, 0 < \alpha < 1.$$

A distribuição dada em 3.15 é uma combinação linear de funções densidade de probabilidade de variáveis aleatórias também com distribuição beta.

Agora que são conhecidas as distribuições condicionais completas para os parâmetros pode-se utilizar o algoritmo de Gibbs para determinar as estimativas.

3.4 Implementação computacional

A implementação computacional foi realizada usando-se o programa WinBUGS 1.4 (SPIEGELHALTER et al., 2002). Para cada modelo, foram geradas cadeias com 5000 iterações para cada parâmetros. As cadeias foram inicializadas em um mesmo ponto e tiveram a convergência monitorada pelo critério de convergência de Gelman e Rubin (1992) existente no programa WinBUGS.

Os parâmetros das distribuições a priori foram escolhidos de forma que as distribuições a priori fossem não informativas ou conjugada beta-gama dependendo do modelo. Em todos os valores utilizados foram $\alpha = 0,25$ e $\lambda = 1$. O modelo Clássico foi analisado no *software* R. Os códigos dos programas nos *software* WinBUGS e no *software* R, podem ser vistos no Apêndice.

3.4.1 WinBUGS

O termo BUGS (*Bayesian Inference Using Gibbs Sampling*), refere-se a inferência Bayesiana utilizando amostrador de Gibbs. Este software é utilizado para a análise bayesiana de modelos estatísticos complexos usando Monte Carlo com Cadeias de Markov (MCMC). O projeto começou em 1989 na unidade de Bioestatística do MRC, em Cam-

bridge, e levou inicialmente o nome de programa BUGS “Classic”, e depois para o *software* WinBUGS desenvolvido em conjunto com o Imperial College School of Medicine at St-Mary’s, em Londres. Portanto, WinBUGS faz parte dos projetos BUGS, que visa tornar os métodos MCMC mais acessível para os que trabalham com Estatística Aplicada. WinBUGS usa uma interface de janelas padrão para controlar a análise, ou pode-se construir o modelo usando uma interface gráfica chamada *Doodle bugs*. WinBUGS é um programa independente, embora ele pode ser chamado a partir de outro *software*.

WinBUGS é uma versão Windows para análise bayesiana de modelos estatísticos complexos usando técnicas de Monte Carlo com Cadeias de Markov (MCMC). Ele permite que os modelos possam ser descrito usando uma versão ligeiramente alterada do idioma BUG, ou usando Doodles (representações gráficas de modelos), que pode, se desejado, ser traduzidos para uma descrição baseada em texto. A linguagem BUGS é mais flexível do que o *Doodles*.

Neste trabalho o WinBUGS foi usado para estimar os parâmetros para a previsão de séries de número de dias com precipitação. Um programa WinBUGS estará completo se estiver composto por três partes: especificação do modelo, entrada de dados, e os valores iniciais. O modelo é baseado em na metodologia do MCMC. A opção para selecionar os valores permite inclusive ser utilizada para a previsão. Portanto, o conhecimento extra sobre os dados podem ser adicionadas e levar a uma melhor previsão.

3.4.2 Diagnóstico de Convergência

Uma das questões mais importantes relativas à aplicação de todas as técnicas MCMC é a questão de quantas repetições iniciais têm de ser descartadas de modo a evitar a possibilidade de viés na estimativa média dos parâmetros causada pelo efeito dos valores iniciais. O parâmetro é analisado, do período de tempo inicial até um número de iteração definido pelo pesquisador, objetivando alcançar a convergência. O tempo que se leva do ponto inicial até o início da convergência é comumente chamado de *burn-in*. Infelizmente, para quase todas as configurações a priori de técnicas utilizando o MCMC, para determinar a taxa de convergência da cadeia, faz-se necessário o *burn-in*. Portanto, é indispensável a realização de algumas análise estatística baseada nos resultados para avaliar a convergência. Para a análise da convergência, usou-se neste trabalho o método de convergência informal.

3.4.3 Método Informal

Talvez a maneira mais fácil de verificar a convergência seja o chamado “método de espessura caneta”, que indica a convergência de uma única cadeia com base em uma análise gráfica da trajetória da cadeia. Se a trajetória da cadeia apresentar o mesmo comportamento qualitativo a partir das iterações após um *burn-in* inicial, então dizemos que houve a convergência. Na prática, isto significa que, se após o *burn-in* inicial todos os valores subsequentes ficam dentro de uma determinada faixa, a cadeia convergiu. Se corre m cadeias paralelas pode-se traçar um histograma dos m valores da n -ésima iteração e comparar esse histograma com m valores de $(n + k)$ iterações, para um k , normalmente entre 10 e 50. A convergência é então aceita se os histogramas não poderem ser distinguidos.

3.5 Medidas de Erros de Previsão

No estudo das técnicas de previsão as medidas de precisão são uma aplicação de extrema importância. Os valores futuros das variáveis tornam-se bastante difíceis de prever dada a complexidade da grande maioria dessas variáveis na vida real. Assim, é fundamental incluir informação acerca da medida em que a previsão pode desviar-se do valor real da variável. Este conhecimento adicional fornece uma melhor percepção sobre o quão precisa pode ser a previsão, ver (STEVENSON, 1996).

A diferença entre o valor real e a previsão do valor dá origem ao erro de previsão:

$$\epsilon_t = A_t - P_t$$

Onde:

ϵ_t = Erro no período t ;

A_t = Valor real no período t ;

P_t = Previsão para o período t .

Erro quadrático médio (EQM)

O erro quadrático médio (EQM) também pode ser usado como uma medida do erro de previsão. O EQM é determinado somando os erros de previsão ao quadrado e dividindo pelo número de erros usados no cálculo, (FERREIRA; VASCONCELOS; ADEODATO, 2005). O erro quadrático médio pode ser expresso pela seguinte equação:

$$EQM = \frac{\sum_{t=1}^n e_t^2}{n}. \quad (3.16)$$

Deseja-se que essa métrica seja a menor possível, em uma previsão perfeita o EQM será zero.

Erro percentual (EPT)

O erro percentual mede a percentagem do erro em relação ao valor real. Calcula-se, subtraindo ao valor real no período t a previsão no respectivo período e divide-se o resultado pelo valor real utilizado anteriormente.

$$EP_t = \frac{(A_t - P_t)}{A_t} \times 100 \quad (3.17)$$

MAPE

De acordo com Ferreira, Vasconcelos e Adeodato (2005) o erro médio percentual absoluto (MAPE - *Mean Absolute Percent Error*).

$$MAPE = \frac{1}{n} \left| \sum_{t=1}^n \frac{(A_t - P_t)}{A_t} \right|, \quad (3.18)$$

Para esta métrica quanto menor o valor, melhor são as previsões geradas pelo modelo. Na previsão perfeita irá assumir valor zero.

3.6 Gráficos de probabilidade

O gráfico de probabilidades é formado por pontos e por uma reta estimada. Os pontos deste gráfico representam percentagens dos dados e são calculados utilizando uma combinação dos métodos não paramétricos e paramétricos. A reta estimada é uma representação gráfica dos percentis, os quais são obtidos utilizando estatística de ordem, estimadores de máxima verossimilhança de uma distribuição de probabilidades selecionada e a função inversa da função de distribuição acumulada desta distribuição selecionada. Considerando que a reta estimada é uma representação dos percentis, primeiro é preciso calcular os percentis para distintas percentagens, com base na distribuição selecionada. Portanto, a transformação de escala, escolhida para linearizar a reta estimada, depende da distribuição dos parâmetros escolhidos. Assim, quanto mais próximos da

linha estimada, melhor a distribuição de probabilidade estima os parâmetros (MARTÍNEZ-ESPINOSA; JÚNIOR; LAHR, 2000).

3.7 Método de Anderson-Darling

Para confirmar o ajuste gráfico, alguns teste de hipóteses não paramétricos podem ser utilizados. Estes testes consideram a forma da distribuição da população em lugar dos parâmetros (ROMEU, 2003). Por este motivo são chamados de testes não-paramétricos. As medidas de ajuste dependem do método de estimação utilizado, sendo o teste de Anderson-Darling, usado para os métodos de máxima verossimilhança e de mínimos quadrados. É uma medida da proximidade dos pontos e da reta estimada no gráfico de probabilidade. O teste de Anderson-Darling é um teste alternativo aos teste de aderência de Chi-quadrado e Kolmogorov-Sminov, o qual tem a vantagem de ser mais sensível que os dois mencionados, pois dá mais peso aos pontos das caudas de distribuição. Assim, valores pequenos da estatística de Anderson-Darling indicam que a distribuição estima melhor os dados (STEPHENS, 1974).

Para estabelecer um critério de rejeição ou não rejeição do modelo (distribuição de probabilidade), é formulada o seguinte teste de hipótese:

$$\begin{cases} H_0, & Y \text{ segue uma distribuição de probabilidade} \\ H_1, & \text{não segue uma determinada distribuição de probabilidade proposta.} \end{cases}$$

A estatística do teste para tomar a decisão é dada por:

$$A^2 = -n - \sum_{i=1}^n \frac{(2i-1)}{n} \ln[F(x_i) + \ln(1 - F(x_{n+1-i}))] \quad (3.19)$$

em que F é a função de distribuição acumulada da distribuição específica. Observe que x_i são os dados ordenados. Os valores críticos ou de rejeição para o teste de Anderson-Darling dependem da distribuição específica que está sendo testada. O teste de Anderson-Darling é um teste unicaudal e a hipótese nula H_0 é rejeitada se o teste estatístico fornecer valor superior ao crítico. Cabe observar que este teste pode ser ajustado, multiplicando-no por uma constante que depende do tamanho da amostra. Estas constantes podem ser encontradas em livro como o (NIST, 2002).

4 Resultados e Discussões

O presente capítulo mostra as análises realizadas na série temporal, número de dias mensais com precipitação (Figura 4.1), nas abordagens clássicas e bayesiana, que tem por finalidade modelar a evolução do número dias com precipitação no município de Garanhuns. As discussões envolvem a modelagem utilizando o modelo auto regressivo de valores inteiros INAR(1). Os *softwares* utilizados nestes contextos foram o software R e o *software* WinBUGS.

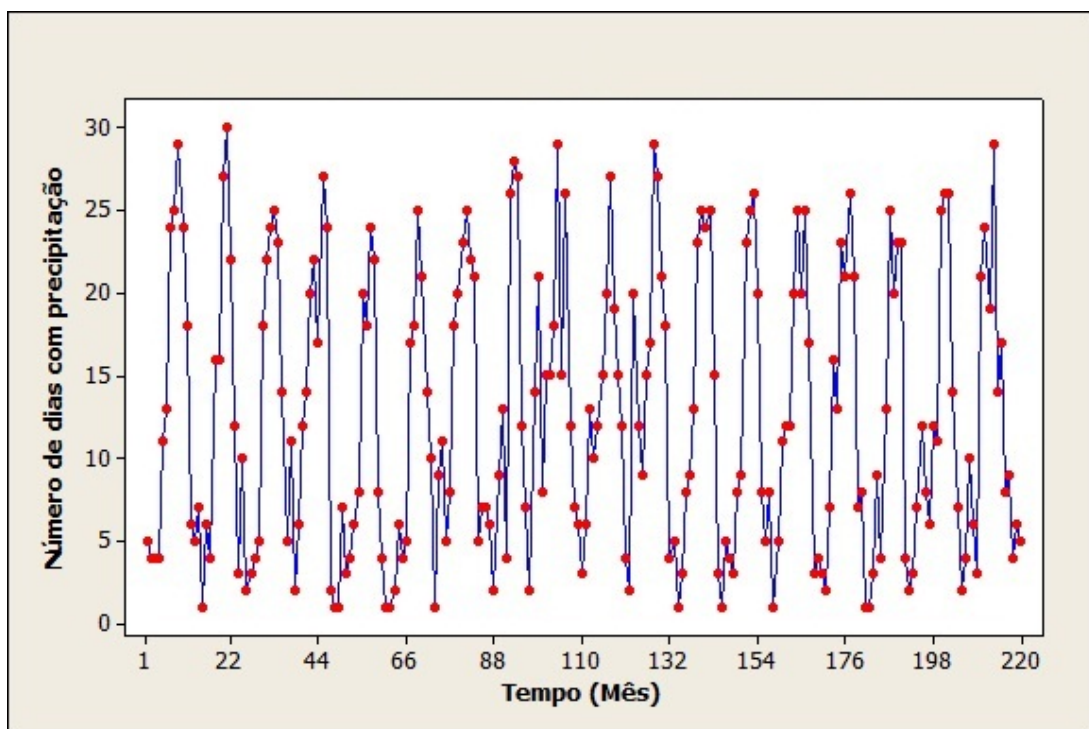


Figura 4.1: Número de dias com precipitação em Garanhuns, 1993 a 2012

Considerou-se que o número de dias mensais com precipitação em um mês está relacionado probabilisticamente ao número de dias com precipitação em um mês anterior, assim como com o número de dias com precipitação em um mês posterior.

4.1 Análise descritiva dos dados

A série completa inicia-se no mês de novembro de 1993 até fevereiro de 2013, totalizando 232 observações. Para o ajustes do modelo retirou-se os 12 últimos meses de observações a fim de comparar com as medidas de previsão ao fim do ajuste.

A Tabela 4.1 contém as medidas descritivas da série do número de dias com precipitação. Tem-se o valor mínimo de 0, ou seja, houve meses em que não ocorreu precipitação. Já em outros meses houve precipitação em todos os dias do mês (máximo = 30), considerando meses com 30 dias.

Tabela 4.1: Medidas descritivas da série de número de dias com precipitação

	Nº Observações	Mínimo	Máximo	Média	Desvio padrão
Série Temporal	220	0	30	12,61	8,53

Os valores das médias por mês, da série temporal número de dias com precipitação, estão apresentadas na Tabela 4.2

Tabela 4.2: Medidas descritivas da série de número de dias com precipitação por mês

Meses	Mínimo	Máximo	Média	Desvio padrão
Janeiro	0	21	6,68	5,78
Fevereiro	2	13	7,00	3,23
Março	3	16	8,50	3,96
Abril	5	21	13,67	4,06
Maió	4	25	19,11	5,30
Junho	17	29	23,11	3,77
Julho	15	30	25,33	3,51
Agosto	14	27	22,50	3,07
Setembro	2	21	12,67	5,12
Outubro	0	12	5,50	2,96
Novembro	0	11	4,37	3,32
Dezembro	0	14	4,37	3,47

Na Figura 4.2 apresentamos a distribuição de média mensal do período analisado, observando um comportamento próximo ao da distribuição normal.

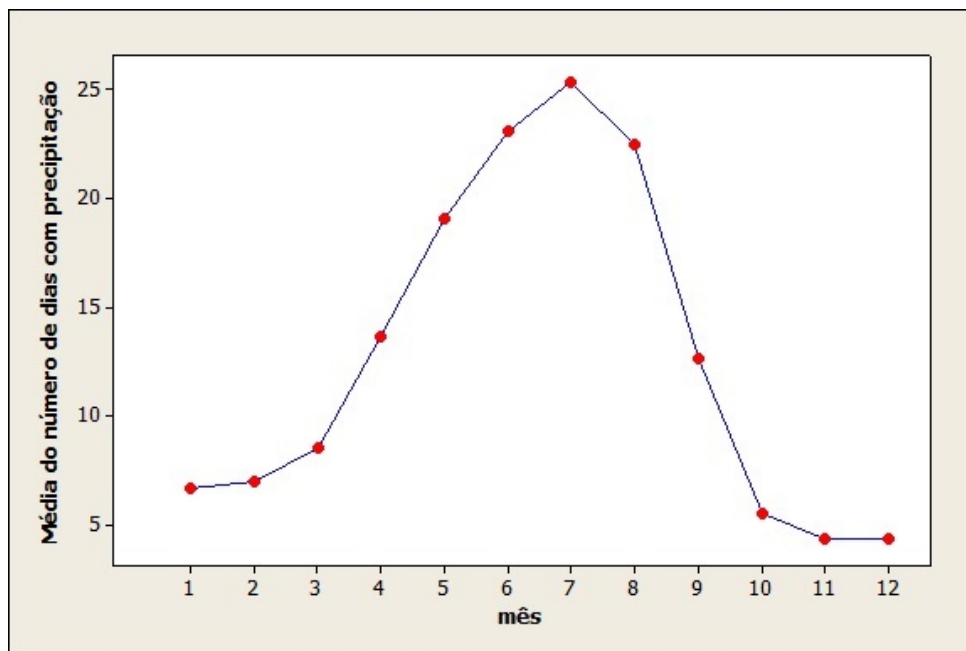


Figura 4.2: Média mensal dos números de dias com precipitação

A função autocorrelação parcial amostral da série número de dias com precipitação estão representados na Figura 4.3.

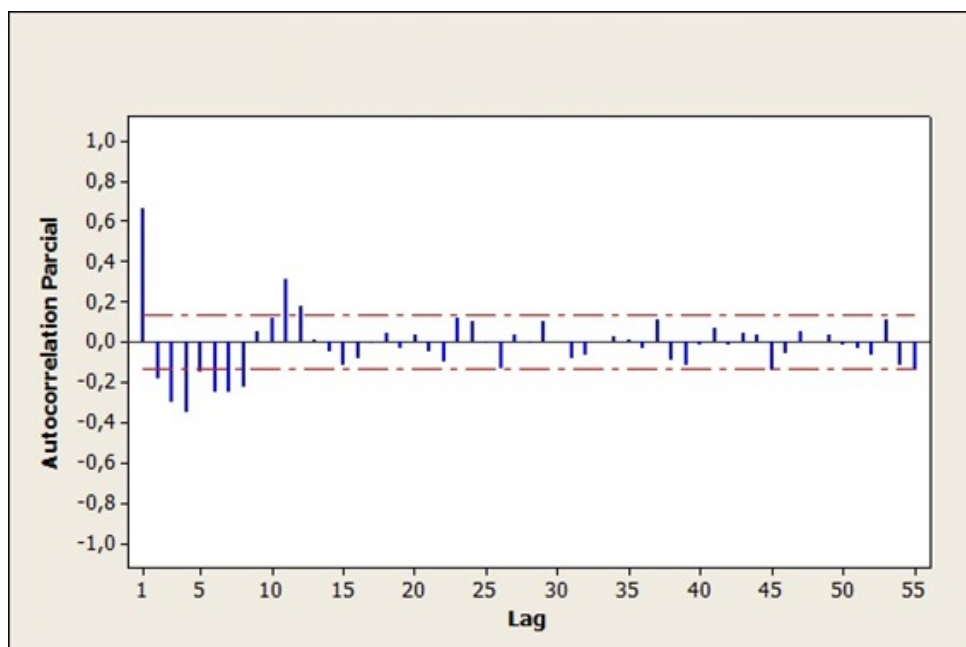


Figura 4.3: Função de autocorrelação parcial, com 5% de limites de significância.

4.2 Modelos ajustados

Nesta tese, analisou-se os dados aplicando o modelo INAR(1) Poisson, e comparou-se essa modelagem com priori Beta e Normal para os parâmetros do modelo. Depois sugerimos uma abordagem Poisson-Normal para o modelo INAR(1), comparando com os resultados encontrados.

Os dados foram ajustados com as seguintes abordagens:

- Clássica - Modelo INAR(1) Poisson (Modelo 1)
- Bayesiana - Modelo INAR(1) Poisson, priori Beta (Modelo 2)
- Bayesiana - Modelo INAR(1) Poisson, priori Normal (Modelo 3)
- Bayesiana - Modelo INAR(1) Poisson-Normal, priori Beta (Modelo 4)
- Bayesiana - Modelo INAR(1) Poisson-Normal, priori Normal (Modelo 5)

Considera-se então um modelo INAR(1) dado por:

$$y_t = \alpha \circ y_{t-1} + \epsilon_t, \geq 2 \quad (4.1)$$

O Modelo 1, terá abordagem clássica. Na abordagem bayesiana usaremos tanto priori não informativa quanto conjugadas, para o INAR (1) Poisson e o INAR(1) Poisson-Normal, assumindo que:

INAR(1) Poisson

- Modelo 2: $\alpha \sim dbeta(10^{-4}, 10^{-4})$ e $\lambda \sim dgamma(10^{-4}, 10^{-4})$ para valores iniciais $a = 0,25$ e $\lambda = 1$, priori conjugada;
- Modelo 3: $\alpha \sim dnorm(10^{-4}, 10^{-4})$ e $\lambda \sim dnorm(10^{-4}, 10^{-4})$ para valores iniciais $\alpha = 0,25$ e $\lambda = 1$, priori não informativa.

INAR(1) Poisson-Normal

- Modelo 4: $\alpha \sim dbeta(10^{-4}, 10^{-4})$ e $\lambda \sim dgamma(10^{-4}, 10^{-4})$ para valores iniciais $a = 0,25$ e $\lambda = 1$, priori conjugada;

- Modelo 5: $\alpha \sim dnorm(10^{-4}, 10^{-4})$ e $\lambda \sim dnorm(10^{-4}, 10^{-4})$ para valores iniciais $\alpha = 0,25$ e $\lambda = 1$, priori não informativa.

Encontrou-se assim os seguintes resultados:

4.2.1 Clássica - Modelo INAR(1) Poisson (Modelo 1)

A modelagem gerou os valores $\alpha = 0,6629$ e $\lambda = 12,3289$, com distribuição de probabilidade apresentada na Figura 4.4 e boxplot da série na Figura 4.5

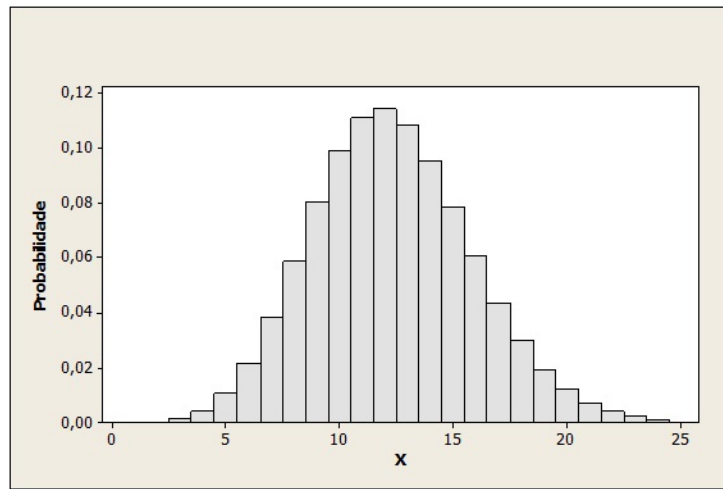


Figura 4.4: Histograma da distribuição de probabilidade dos resultados do Modelo 1.

Na Figura 4.5, 50% dos dados estão entre 5 e 20. 25% dos dados estão abaixo de 5 e 25% acima de 20. Dentre os valores situados entre 5 e 20, existe uma menor dispersão dos dados no intervalo 5 e 12, em comparação ao intervalo 12 e 20.

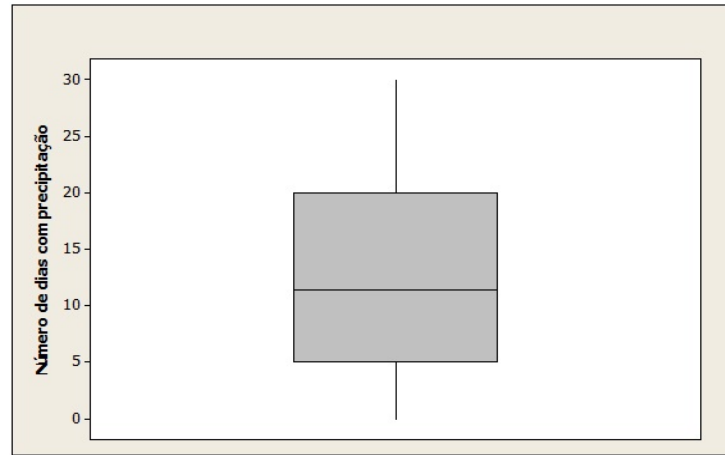


Figura 4.5: Box-plot da série temporal

4.2.2 Bayesiana - Modelo INAR(1) Poisson, priori Beta (Modelo 2)

Resultando um valor para $\alpha = 0,6549$ e $\lambda = 12,6500$, com função marginal a posteriori para cada parâmetro, apresentada na Figura 4.6

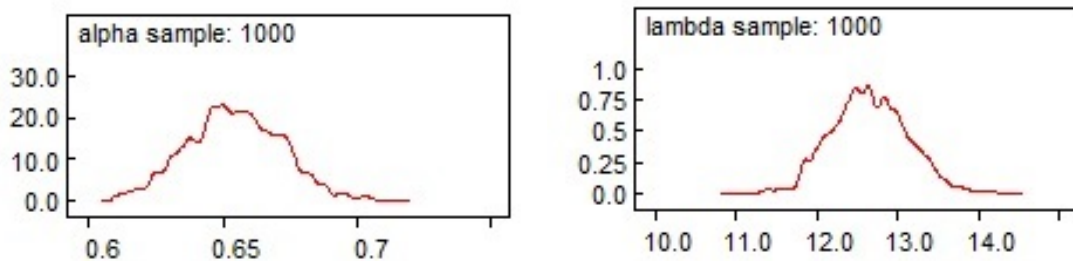


Figura 4.6: Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 2.

Foram realizadas 5000 iterações para os parâmetros do modelo 2, e verificamos na Figura 4.6, que a função marginal a posteriori usou apenas as 1000 iterações finais, descartando assim as 4000 primeiras iterações (*burn-in*). Observa-se que a distribuição dos parâmetros tem uma concentração de informação em torno da média. Considerando os valores iniciais para $\alpha = 0,25$ e $\lambda = 1$ a Figura 4.7 mostra a dinâmica da convergência dos parâmetros do Modelo 2.

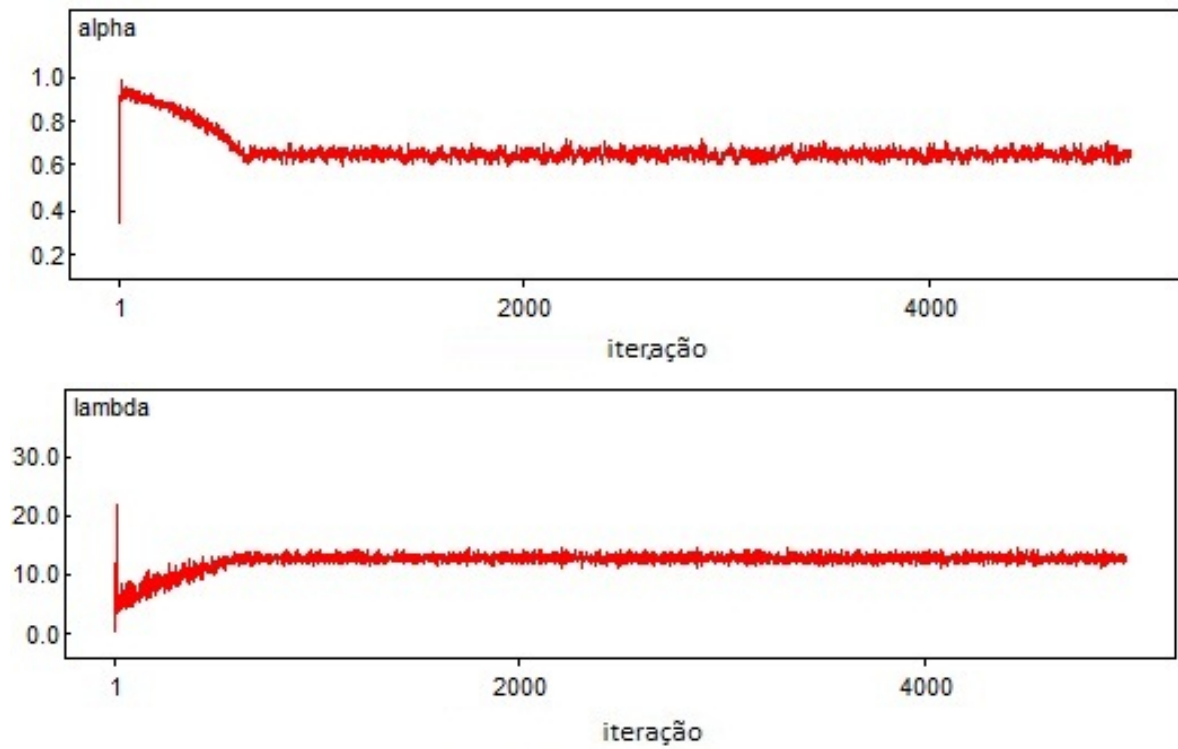


Figura 4.7: Convergência dos parâmetros do Modelo 2

Na Figura 4.8, verificamos o intervalo de variação do parâmetros, nas 1000 iterações finais.

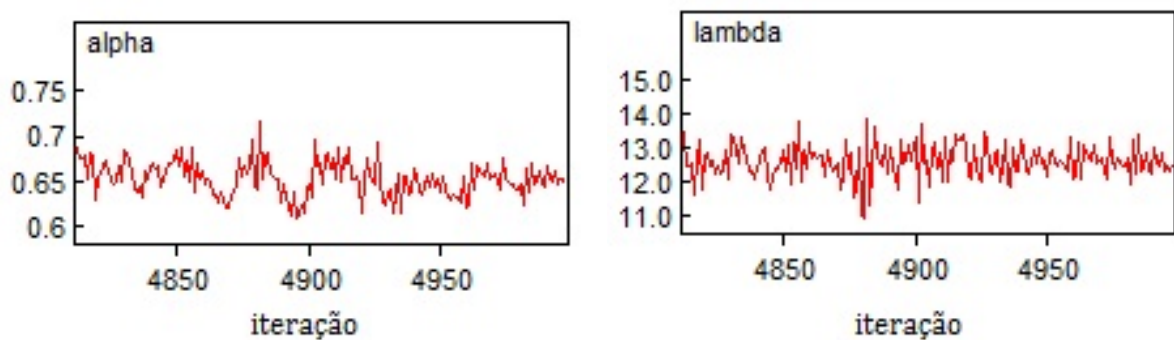


Figura 4.8: Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 2).

Na Figura 4.9, apresentamos as funções de autocorrelação dos parâmetros, onde a suavidade no decaimento indica um comportamento autoregressivo.

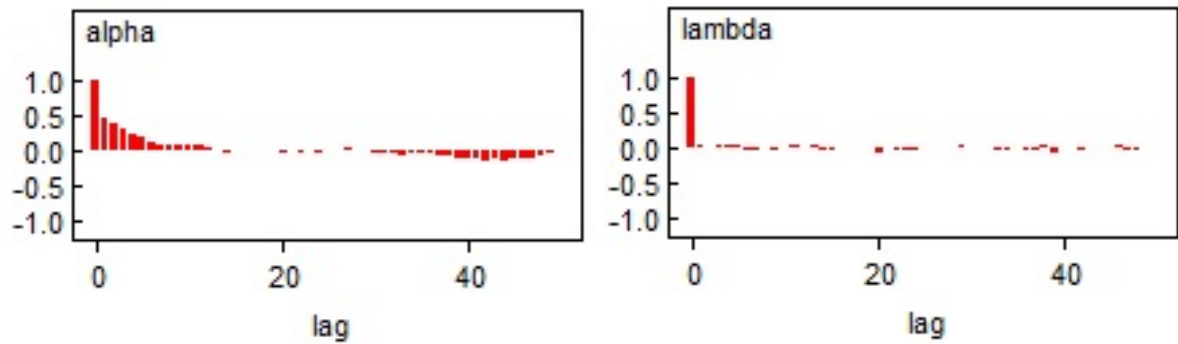


Figura 4.9: Função de autocorrelação parcial para os parâmetros α e λ do Modelo 2.

4.2.3 Bayesiana - Modelo INAR(1) Poisson, priori Normal (Modelo 3)

Foram realizadas 5000 iterações para os parâmetros do modelo 3, e verificamos na Figura 4.10, que a função marginal à posteriori usou apenas as 1000 iterações finais, descartando assim as 4000 primeiras iterações (*burn-in*). Verificamos na Figura 4.10, que a distribuição dos parâmetros tem uma concentração de informação em torno da média. Resultando um valor para $\alpha = 0,6509$ e $\lambda = 12,6200$, com função marginal a posteriori para cada parâmetros, apresentada na Figura 4.10

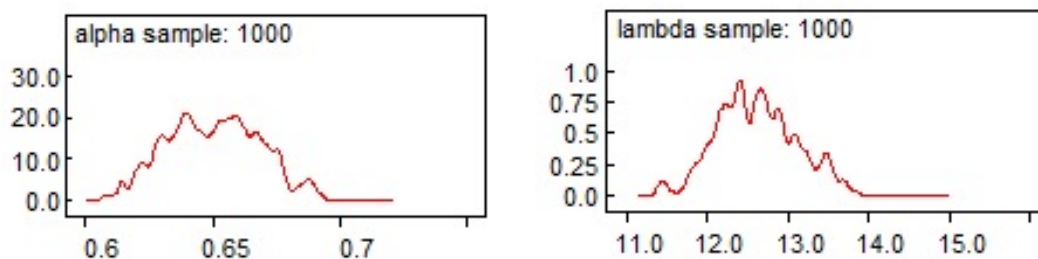


Figura 4.10: Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 3.

Considerando os valores iniciais para $\alpha = 0,25$ e $\lambda = 1$ a Figura 4.11 mostra a dinâmica da convergência dos parâmetros do Modelo 3.

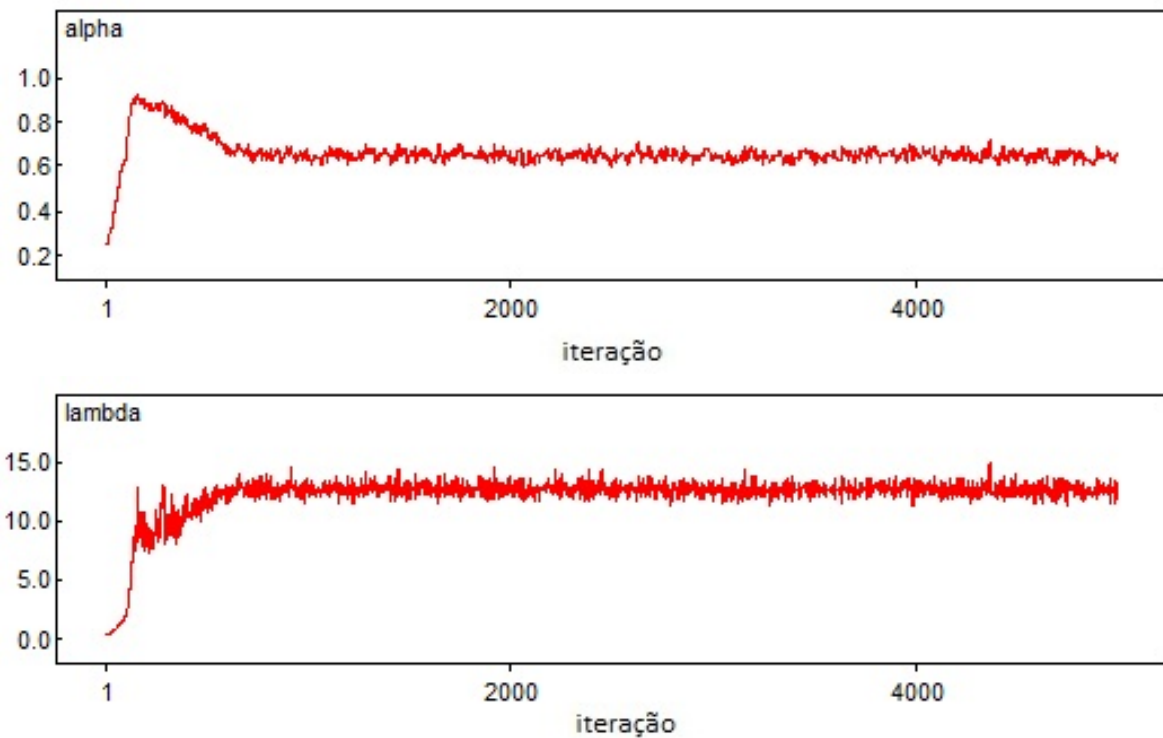


Figura 4.11: Convergência dos parâmetros do Modelo 3

Na Figura 4.12, verificamos o intervalo de variação do parâmetros, nas 1000 iterações finais.

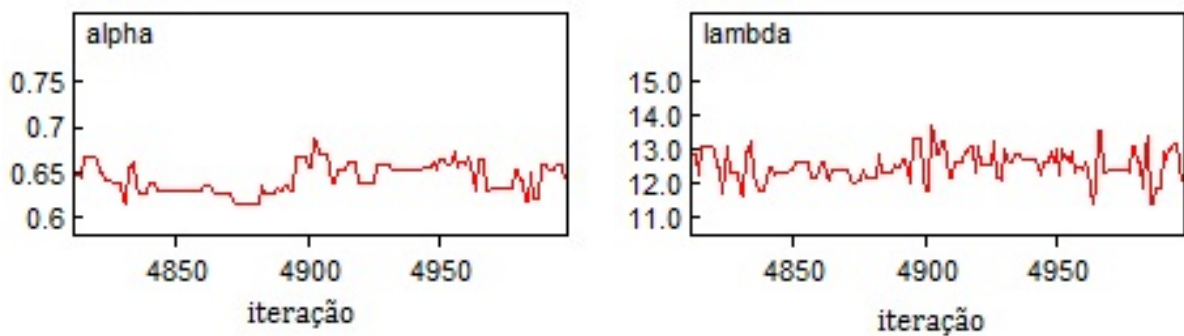


Figura 4.12: Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 3).

Na Figura 4.13 são apresentadas as funções de autocorrelação dos parâmetros, onde a suavidade no decaimento indica um comportamento autoregressivo.

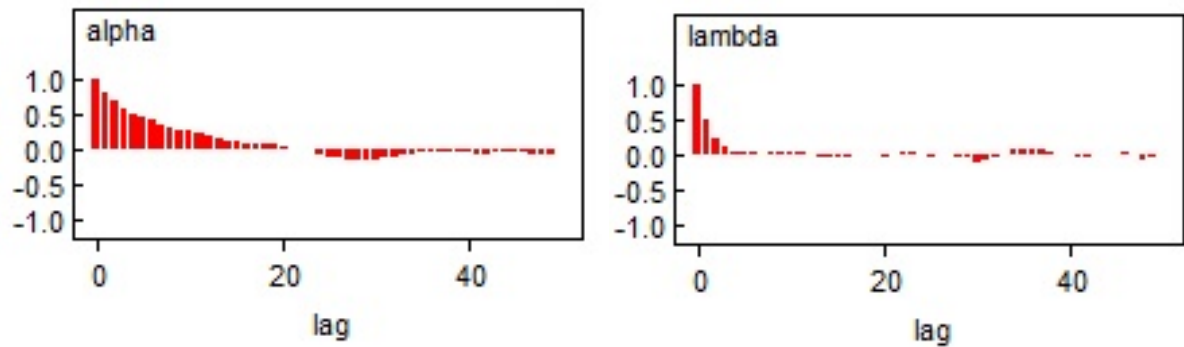


Figura 4.13: Função de autocorrelação parcial para os parâmetros α e λ do Modelo 3.

4.2.4 Bayesiana - Modelo INAR(1) Poisson-Normal, priori Beta (Modelo 4)

Foram realizadas 5000 iterações para os parâmetros do modelo 4, e verificamos na Figura 4.14, que a função marginal á posteriori usou as 4500 iterações finais, descartando assim as 500 primeiras iterações (*burn-in*). Verificamos na Figura 4.14, que a distribuição dos parâmetros tem uma concentração de informação em torno da média. Resultando um valor para $\alpha = 0,6539$ e $\lambda = 12,6600$, com função marginal a posteriori para cada parâmetros, apresentada na Figura 4.14

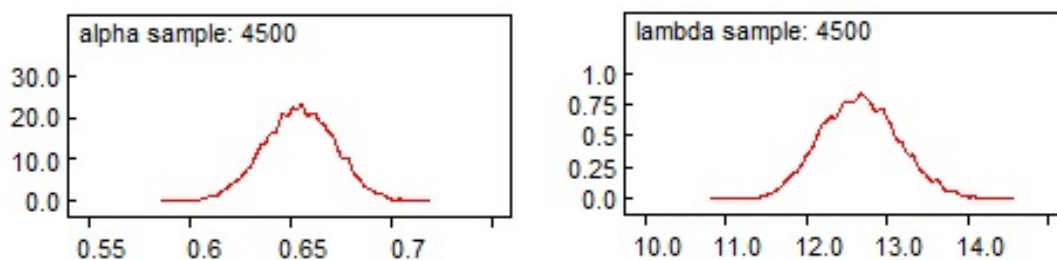


Figura 4.14: Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 4.

Considerando os valores iniciais para $\alpha = 0,25$ e $\lambda = 1$ a Figura 4.15 mostra a dinâmica da convergência dos parâmetros do Modelo 4.

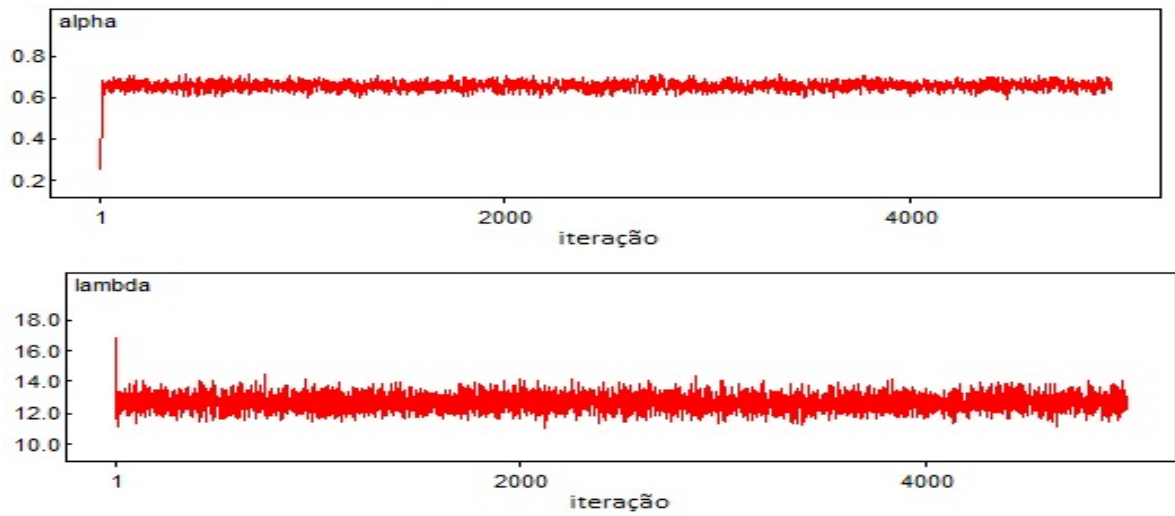


Figura 4.15: Convergência dos parâmetros do Modelo 4.

Na Figura 4.16, verificamos o intervalo de variação do parâmetros, nas 1000 iterações finais.

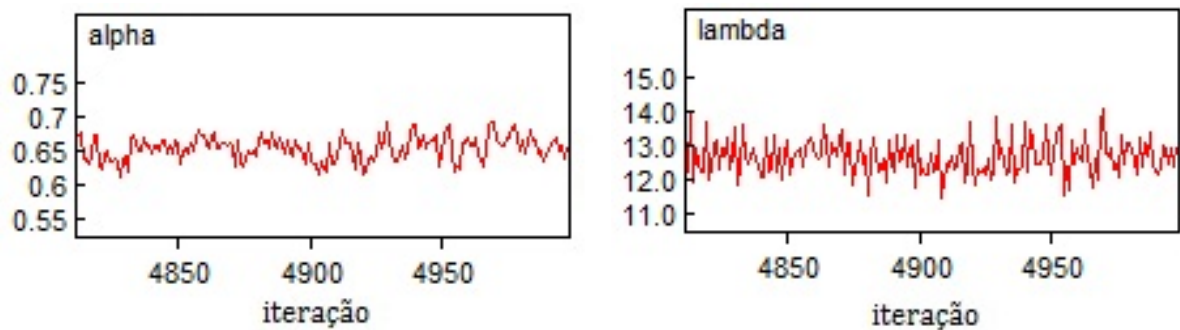


Figura 4.16: Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 4).

Na Figura 4.17 são apresentadas as funções de autocorrelação dos parâmetros, onde a suavidade no decaimento indica um comportamento autoregressivo.

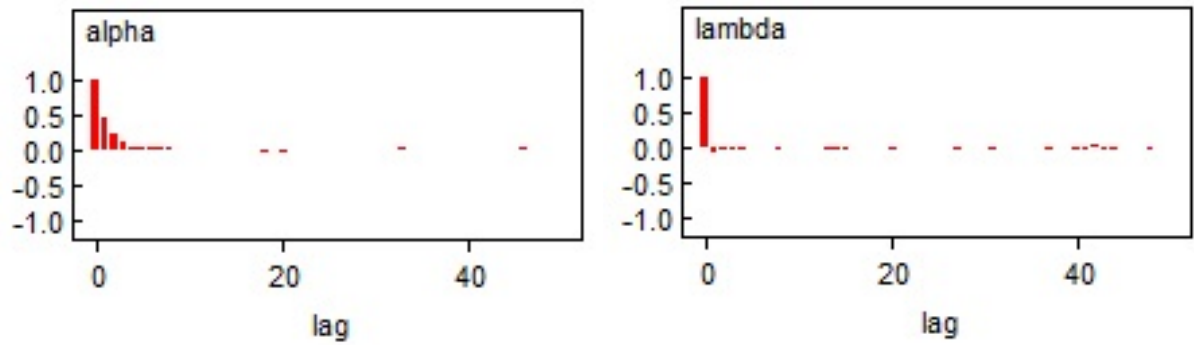


Figura 4.17: Função de autocorrelação parcial para os parâmetros α e λ do Modelo 4.

4.2.5 Bayesiana - Modelo INAR(1) Poisson-Normal, priori Normal (Modelo 5)

Foram realizadas 5000 iterações para os parâmetros do modelo 5, e verificamos na Figura 4.18, que a função marginal a posteriori usou as 1000 iterações finais, descartando assim as 4000 primeiras iterações (*burn-in*). Verificamos na Figura 4.18, que a distribuição dos parâmetros tem uma concentração de informação em torno da média. Resultando um valor para $\alpha = 0,6528$ e $\lambda = 12,6300$, com função marginal a posteriori para cada parâmetros, apresentada na Figura 4.18.

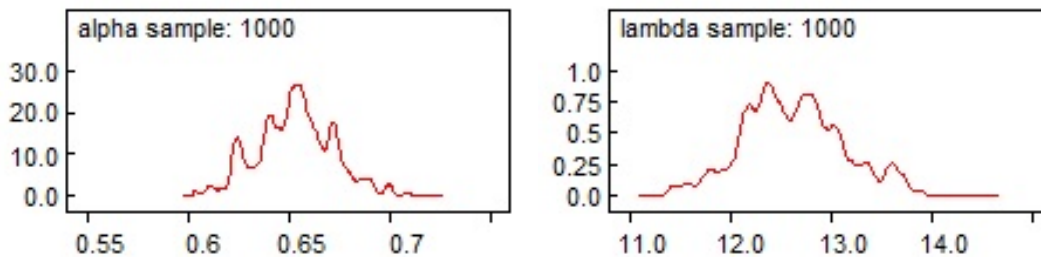


Figura 4.18: Distribuição a posteriori marginal dos parâmetros α e λ do Modelo 5.

Considerando os valores iniciais para $\alpha = 0.25$ e $\lambda = 1$ a Figura 4.19 mostra a dinâmica da convergência dos parâmetros do Modelo 5.

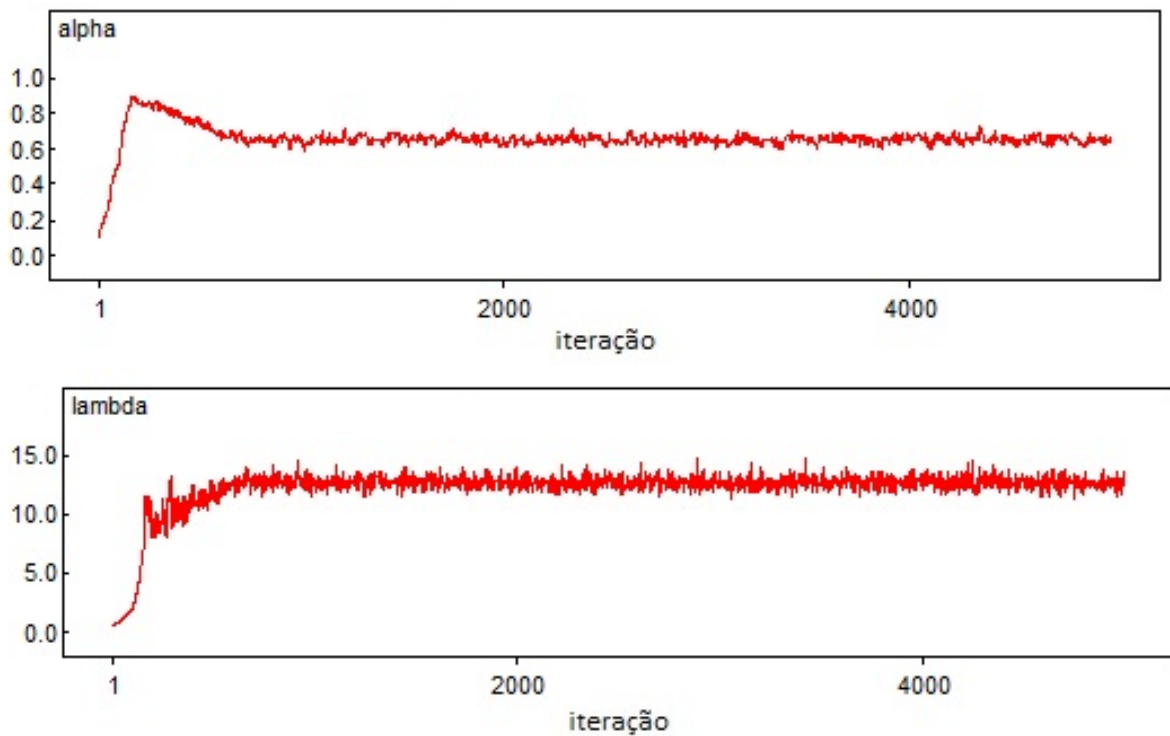


Figura 4.19: Convergência dos parâmetros do Modelo 5.

Na Figura 4.20, verificamos o intervalo de variação do parâmetros, nas 1000 iterações finais.

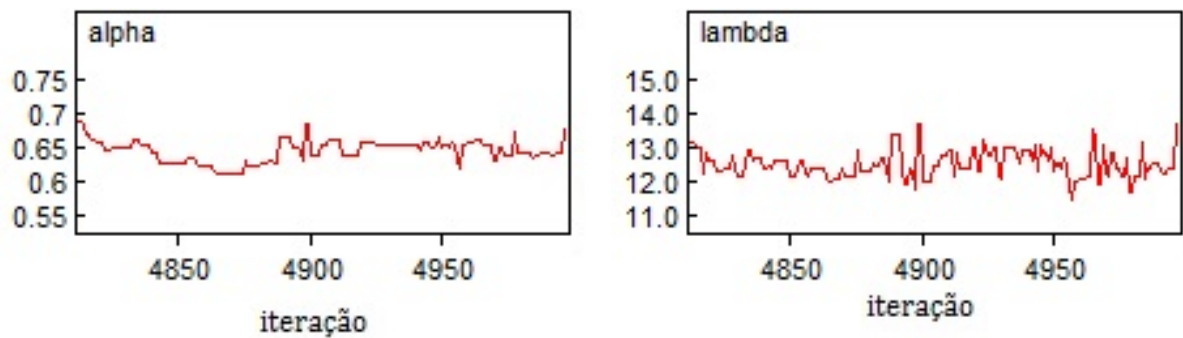


Figura 4.20: Traço da variabilidade dos parâmetros nas 1000 iterações finais (Modelo 5).

Na Figura 4.21 são apresentadas as funções de autocorrelação dos parâmetros, onde a suavidade no decaimento indica um comportamento autoregressivo.

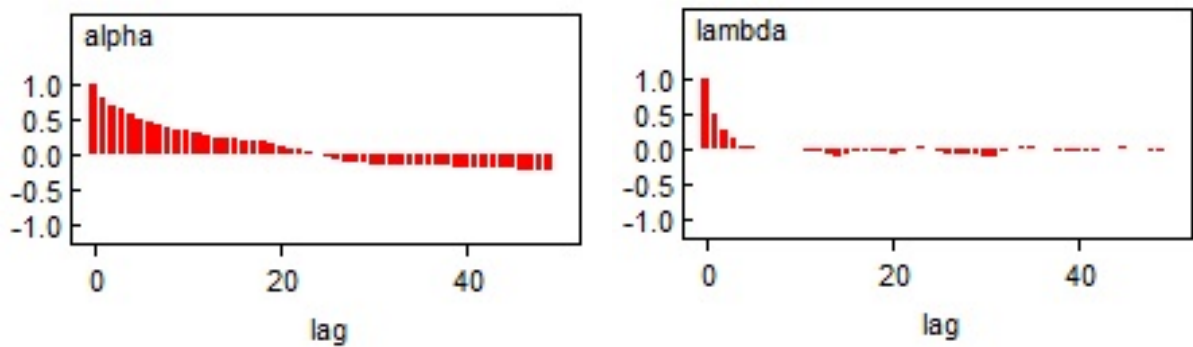


Figura 4.21: Função de autocorrelação parcial para os parâmetros α e λ do Modelo 5.

4.2.6 Adequação dos modelos

Na Tabela 4.3 verificamos os valores AIC calculados para os modelos.

Tabela 4.3: Valores de AIC obtidos nos modelos

Modelos	AIC
Modelo 1	1447,7
Modelo 2	1393,0
Modelo 3	1391,0
Modelo 4	1391,0
Modelo 5	1389,0

Comparando o valor de AIC entre os modelos, os modelos com abordagem bayesiana se ajustou melhor aos dados. Dentre esses modelos, o modelo 1 apresentou AIC = 1447,7, sendo o que apresentou ajuste menos eficiente aos dados. O modelo 5 foi o que apresentou melhor ajuste aos dados, segundo valor de AIC.

De acordo com a Tabela 4.4, que apresenta os valores do critério AICc o modelo 5 foi o indicado com melhor ajuste, visto que obteve menor valor. Este critério demonstra que o modelo 1 (AICc=1692,4) foi o que demonstrou “pior” ajuste aos dados.

Tabela 4.4: Valores de AICc obtidos nos modelos

Modelos	AICc
Modelo 1	1692,4
Modelo 2	1681,3
Modelo 3	1679,2
Modelo 4	1677,8
Modelo 5	1679,1

O critério BIC (Tabela 4.5) indica que o modelo INAR(1)Poisson-Normal utilizando a priori Normal (Modelo 5), foi o que melhor se ajustou aos dados, pois teve o menor valor de BIC (1396,0).

Tabela 4.5: Valores de BIC obtidos nos modelos

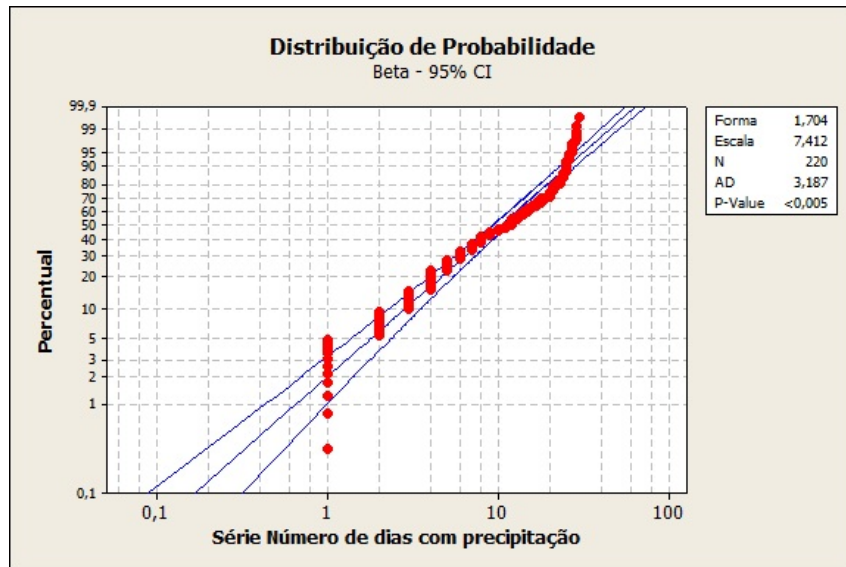
Modelos	BIC
Modelo 1	1424,5
Modelo 2	1407,0
Modelo 3	1401,0
Modelo 4	1401,0
Modelo 5	1396,0

Os critérios de seleção de modelos utilizados nestes trabalho apresentaram resultados coerentes. Em todos os casos, o modelo clássico (Modelo 1) foi o que teve menor adequação ao ajuste dos dados, e, dentre os modelos bayesianos o INAR(1) Poisson-Normal utilizando a priori Normal (Modelo 5) demonstrou melhor ajuste aos dados.

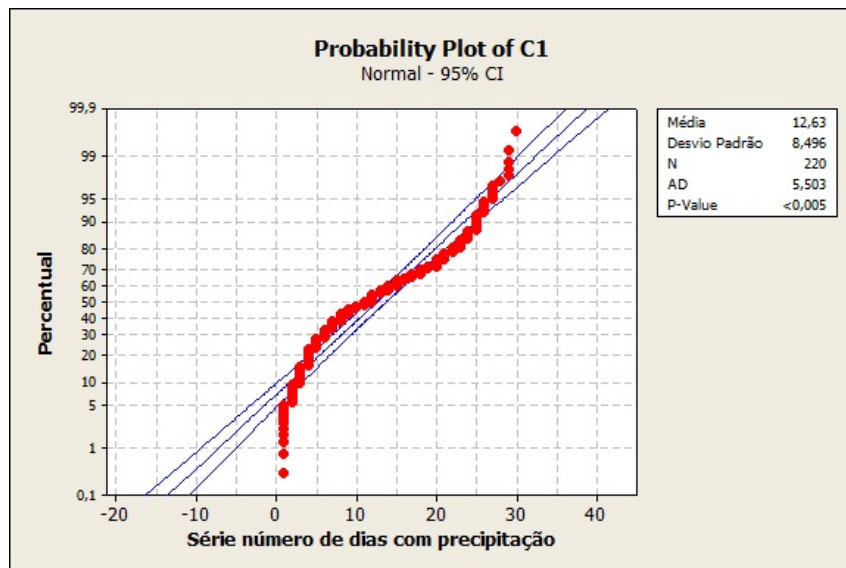
4.3 Distribuição de Probabilidade

Segundo Steutel e Harn (1979), ao adaptar o modelo AR com o operador *thinning*, compara-se o resultado obtido com essa abordagem e o efeito de utilizar a parte inteira do dado contínuo, relatando que o usual era aplicar modelos contínuos mesmo em dados discretos. Considerando a forma de modelar dados discretos antes de Steutel e Harn (1979), aplicamos os modelos Normal e Beta na série e verificamos qual das duas se ajustam melhor aos dados, objetivando analisar se essa abordagem influência no resultado obtido com

a proposta de abordar modelos de dados discretos com modelos mistos. As Figuras 22(a) e 22(b), apresentam as distribuições de probabilidade Beta e Normal, respectivamente.



(a) Distribuições Beta



(b) Distribuições Normal

Figura 4.22: Teste de aderência (Anderson-Darling) para as distribuições Beta e Normal

Usando o teste de Anderson-Darling, encontrou-se para a distribuição Beta o valor de 3,187, e, para distribuição Normal o valor de 5,503. Assim segundo esse teste, a distribuição Beta apresenta um ajuste melhor aos dados se comparado com a distribuição Normal.

4.4 Previsão

Nesta seção, será considerado os cálculos das previsões para o processo autoregressivo de valores inteiros definido em 2.2, seja y_{n+h} a previsão de um valor futuro. Assumindo que o processo INAR(1) é estacionário, após sucessivas substituições podemos escrever y_{n+h} na forma apresentada em 2.12.

Uma vez estimados parâmetros dos modelos utilizados neste trabalho através dos dados valores observados da série (n=220), foi realizado a comparação da previsão dos 12 meses primeiros meses após último mês dos dados utilizados na modelagem da série.

Na Tabela 4.6 é apresentado os valores reais e as previsões encontradas a partir da aplicação de cada modelo, utilizando os tipos de erro: Erro Percentual Total (EPT), Erro Absoluto Médio Percentual (MAPE) e Erro Quadrático Médio (EQM).

Tabela 4.6: Valores previstos pelos modelos (12 passos à frente) e erros de previsão

Valor observado	Previsão Clássica	Previsão Bayesiana Modelo 2	Previsão Bayesiana Modelo 3	Previsão Bayesiana Modelo 4	Previsão Bayesiana Modelo 5
4	3	3	3	4	3
5	2	2	3	4	2
12	3	3	3	5	73
22	7	7	8	9	7
29	14	14	14	15	15
22	19	18	19	20	18
8	14	14	14	13	14
8	5	5	5	4	6
1	5	5	5	4	5
4	0	0	1	2	0
7	2	2	3	5	2
1	4	4	5	6	4
EPT	36,59%	37,18%	32,79%	32,48%	37,40%
MAPE	1,064	1,068	1,095	0,988	1,027
EQM	23,61	23,86	22,07	17,93	20,64

Observando os resultados indicados na Tabela 4.6, verificou-se que o modelo 4 apre-

sentou um menor erro nos métodos de análise de erro utilizado.

4.5 Análises Complementares

Considerando a abordagem bayesiana para estimação dos parâmetros dos modelos 2, 3, 4 e 5, realizou-se algumas análises complementares, dividindo-as em duas partes. Na primeira parte, fixou-se o número de iterações e modificou-se o valor da priori, utilizando $\alpha = 0,15$, $\alpha = 0,25$ e $\alpha = 0,50$. Na segunda parte, foi fixado a priori de α e modificou o número de iterações, utilizando 50000 e 500000.

4.5.1 Parte I

Fixando o número de iterações em 5000 e considerando os casos em que $\alpha = 0,15$ (Figura 4.23) e $\alpha = 0,50$ (Figura 4.24) foi obtido uma amplitude de 0,0019.

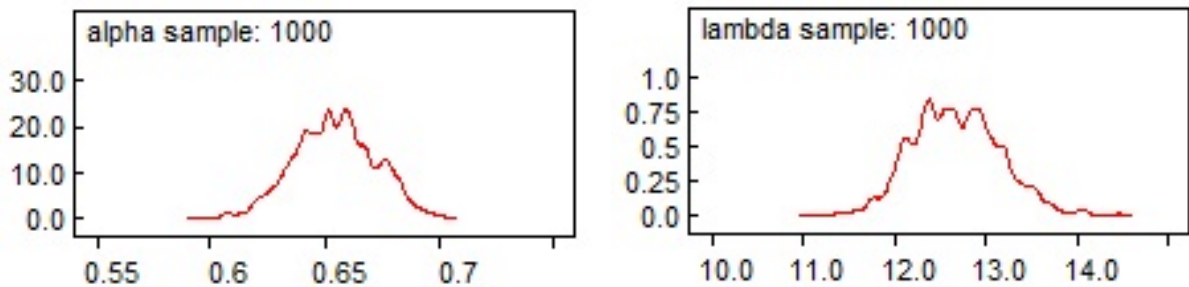


Figura 4.23: Função marginal à posteriori dos parâmetros $\alpha = 0,6539$ e $\lambda = 12,68$, do Modelo 2, com $\alpha = 0,15$.

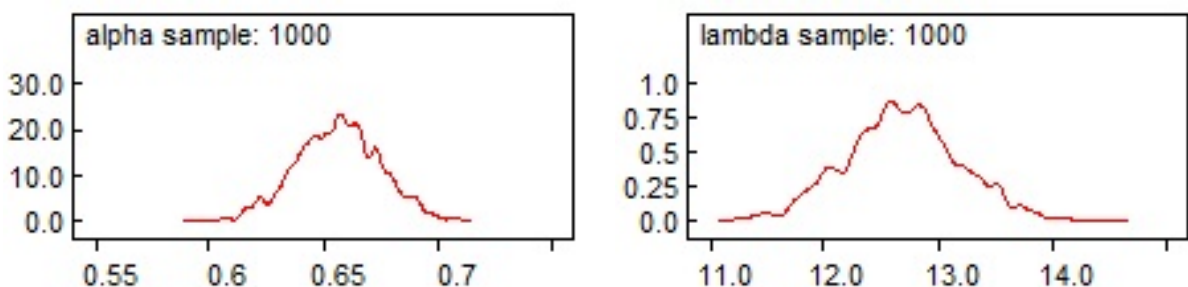


Figura 4.24: Função marginal à posteriori dos parâmetros $\alpha = 0,6558$ e $\lambda = 12,68$, do Modelo 2, com $\alpha = 0,50$.

Fixando o número de iterações em 5000 e considerando os casos em que $\alpha = 0,15$ (Figura 4.25) e $\alpha = 0,50$ (Figura 4.26) foi obtido uma amplitude de 0,0012.

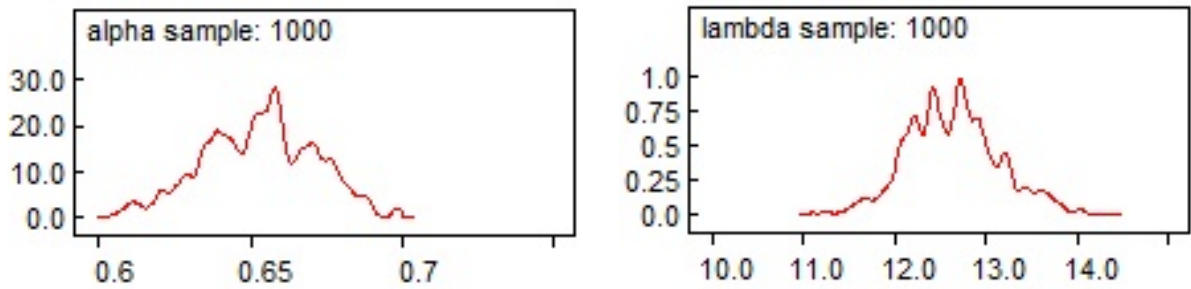


Figura 4.25: Função marginal à posteriori dos parâmetros $\alpha = 0,6534$ e $\lambda = 12,64$, do Modelo 3, com $\alpha = 0,15$.

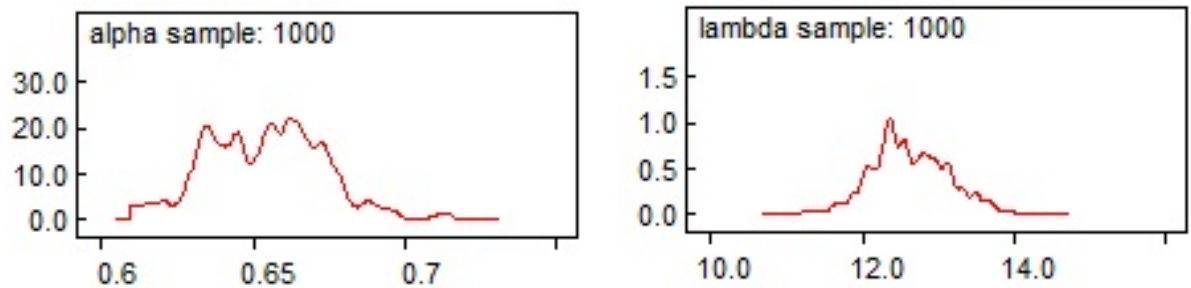


Figura 4.26: Função marginal à posteriori dos parâmetros $\alpha = 0,6546$ e $\lambda = 12,64$, do Modelo 3, com $\alpha = 0,50$.

Fixando o número de iterações em 5000 e considerando os casos em que $\alpha = 0,15$ (Figura 4.27) e $\alpha = 0,50$ (Figura 4.28) foi obtido uma amplitude de 0,0011.

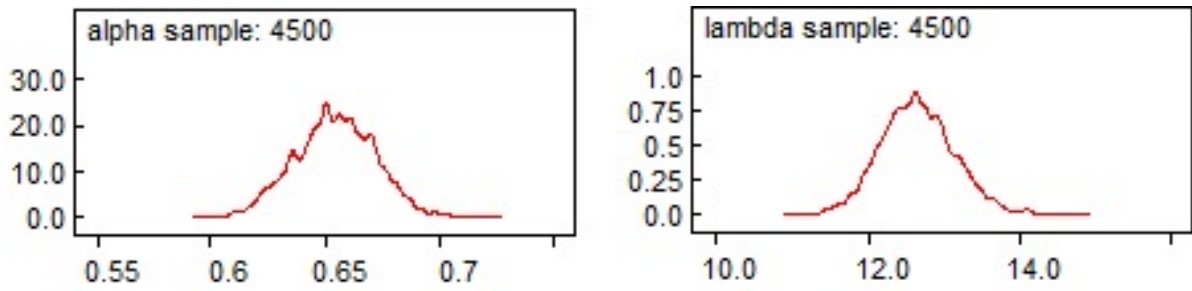


Figura 4.27: Função marginal à posteriori dos parâmetros $\alpha = 0,6536$ e $\lambda = 12,67$, do Modelo 4, com $\alpha = 0,15$.

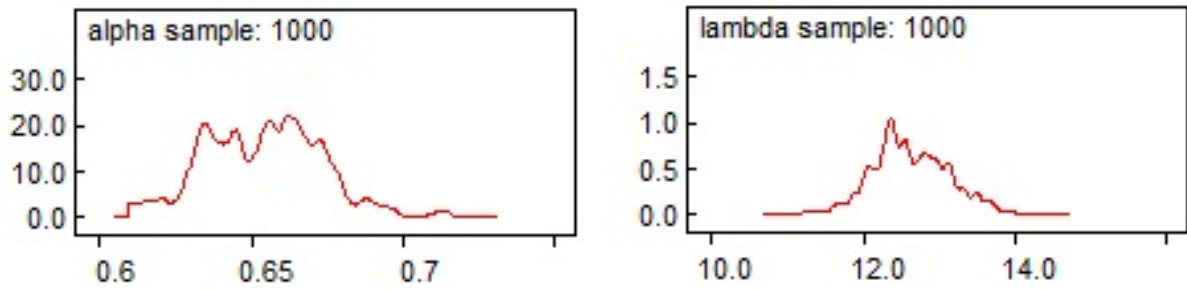


Figura 4.28: Função marginal à posteriori dos parâmetros $\alpha = 0,6547$ e $\lambda = 12,66$, do Modelo 4, com $\alpha = 0,50$.

Fixando o número de iterações em 5000 e considerando os casos em que $\alpha = 0,15$ (Figura 4.29) e $\alpha = 0,50$ (Figura 4.30) foi obtido uma amplitude de 0,0016.

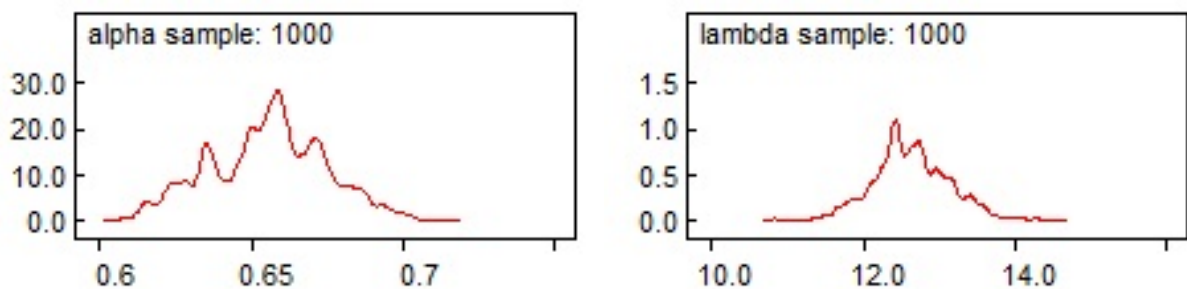


Figura 4.29: Função marginal à posteriori dos parâmetros $\alpha = 0,6560$ e $\lambda = 12,64$, do Modelo 5, com $\alpha = 0,15$.

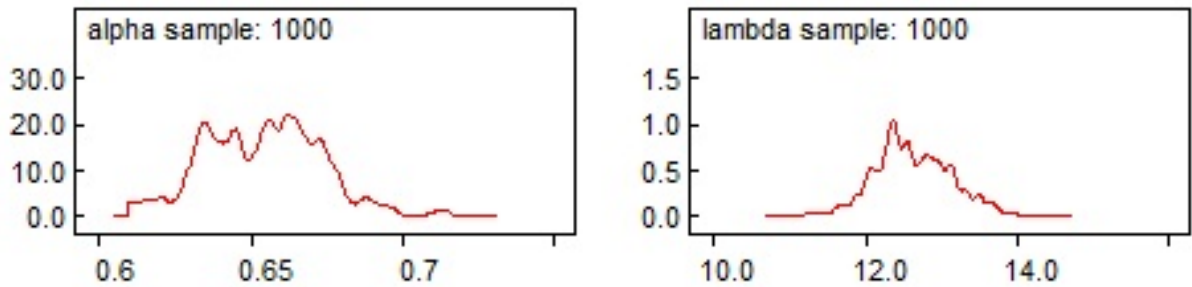


Figura 4.30: Função marginal à posteriori dos parâmetros $\alpha = 0,6544$ e $\lambda = 12,62$, do Modelo 5, com $\alpha = 0,50$.

Foi realizado 5000 iterações para cada modelo com abordagem bayesiana, alterando o valor inicial de α em 0,15 e 0,5, verificamos que o modelo 4 apresentou menor amplitude de variação para o valor a posterior de α e λ (Tabela 4.7).

Tabela 4.7: Amplitude dos parâmetros α e λ variando α

Modelo	Amplitude de α	Amplitude de λ
Modelo 2	0,0019	0,03
Modelo 3	0,0012	0,02
Modelo 4	0,0011	0,01
Modelo 5	0,0016	0,01

Com os resultados apresentados na Tabela 4.7 podemos afirmar que os modelos 4 e 3 foram os que tiveram menor variação na forma.

4.5.2 Parte II

Fixando o valor de α em 0,25 e considerando o número de iterações iguais a 50000 (Figura 4.31) e 500000 (Figura 4.32) foi obtido uma amplitude de 0,0006, em comparação com o α do modelo 2.

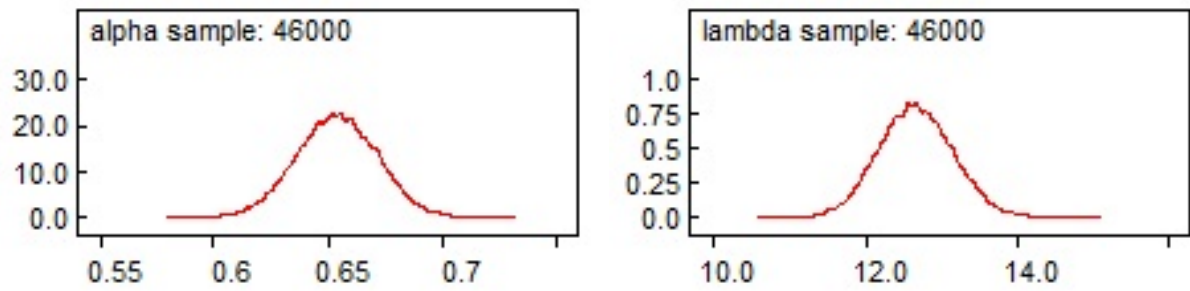


Figura 4.31: Função marginal à posteriori dos parâmetros $\alpha = 0,6542$ e $\lambda = 12,68$, do Modelo 2, com 50000 iterações, sendo 4000 de *Burn-in*.

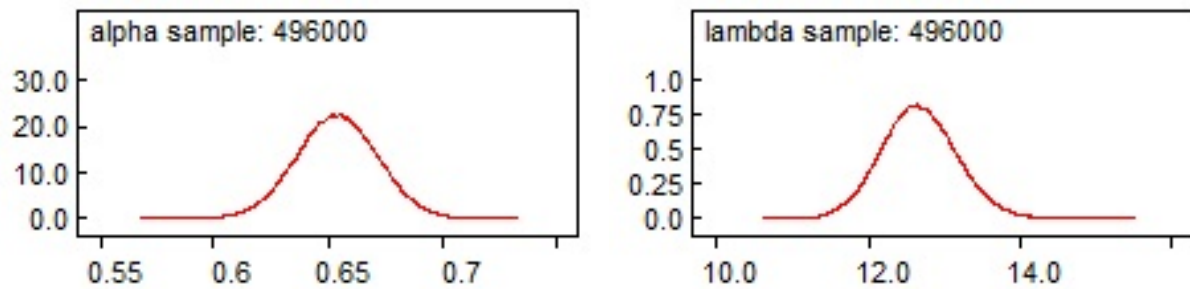


Figura 4.32: Função marginal à posteriori dos parâmetros $\alpha = 0,6543$ e $\lambda = 12,68$, do Modelo 2, com 500000 iterações, sendo 4000 de *Burn-in*.

Fixando o valor de α em 0,25 e considerando o número de iterações iguais a 50000 (Figura 4.33) e 500000 (Figura 4.34) foi obtido uma amplitude de 0,0004, em comparação com o α do modelo 3.

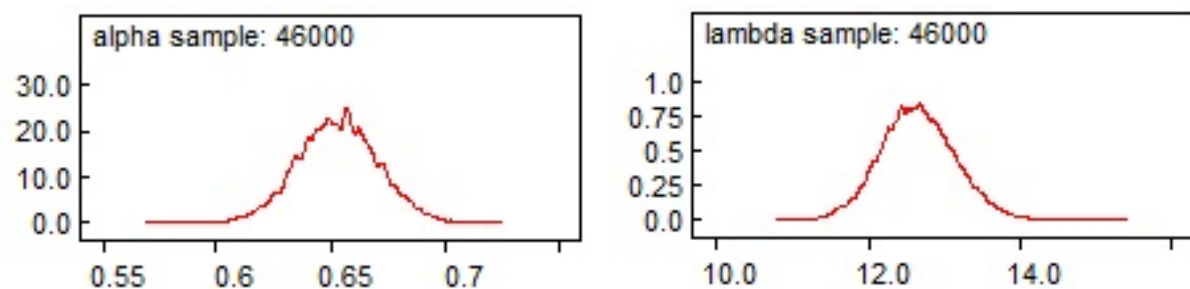


Figura 4.33: Função marginal à posteriori dos parâmetros $\alpha = 0,6505$ e $\lambda = 12,64$, do Modelo 3, com 50000 iterações, sendo 4000 de *Burn-in*.

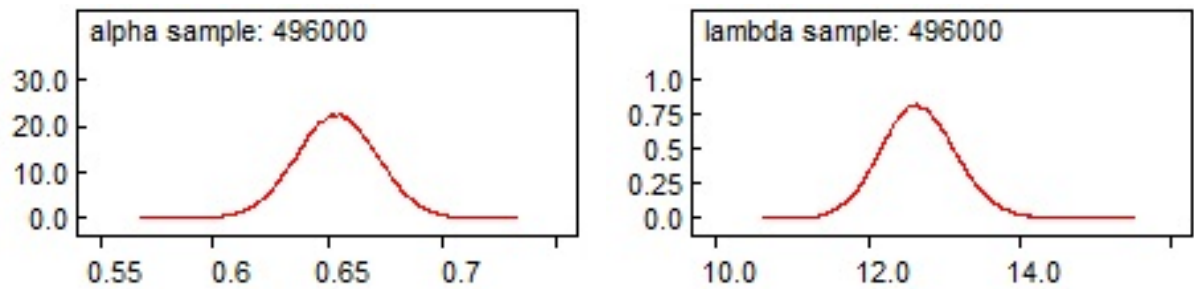


Figura 4.34: Função marginal à posteriori dos parâmetros $\alpha = 0,6505$ e $\lambda = 12,64$, do Modelo 3, com 500000 iterações, sendo 4000 de *Burn-in*.

Fixando o valor de α em 0,25 e considerando o número de iterações iguais a 50000 (Figura 4.35) e 500000 (Figura 4.36) foi obtido uma amplitude de 0,0001, em comparação com o α do modelo 4.

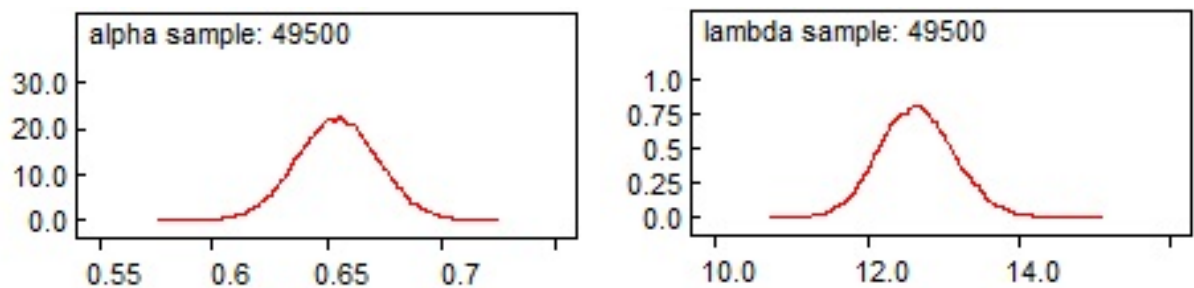


Figura 4.35: Função marginal à posteriori dos parâmetros $\alpha = 0,6540$ e $\lambda = 12,66$, do Modelo 4, com 50000 iterações, sendo 4000 de *Burn-in*.

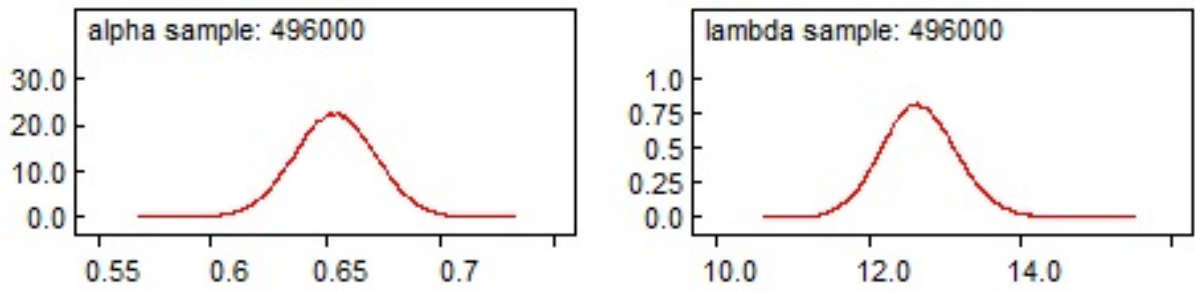


Figura 4.36: Função marginal à posteriori dos parâmetros $\alpha = 0,6540$ e $\lambda = 12,66$, do Modelo 4, com 500000 iterações, sendo 4000 de *Burn-in*.

Fixando o valor de α em 0,25 e considerando o número de iterações iguais a 50000 (Figura 4.37) e 500000 (Figura 4.38) foi obtido uma amplitude de 0,0001, em comparação com o α do modelo 5.

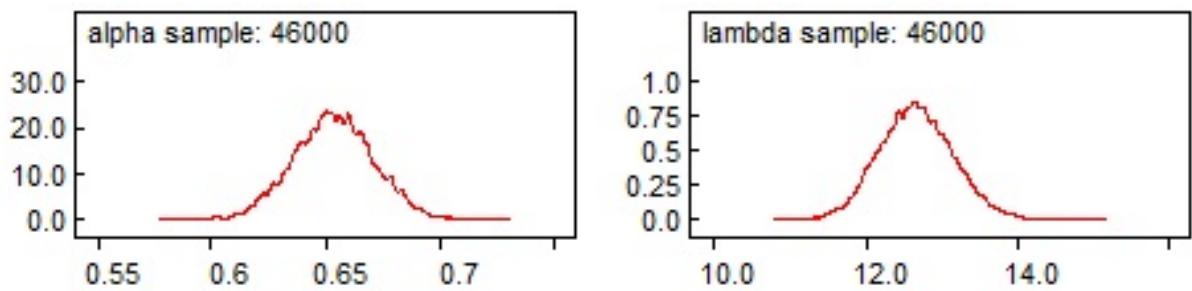


Figura 4.37: Função marginal à posteriori dos parâmetros $\alpha = 0,6528$ e $\lambda = 12,67$, do Modelo 5, com 50000 iterações, sendo 4000 de *Burn-in*.

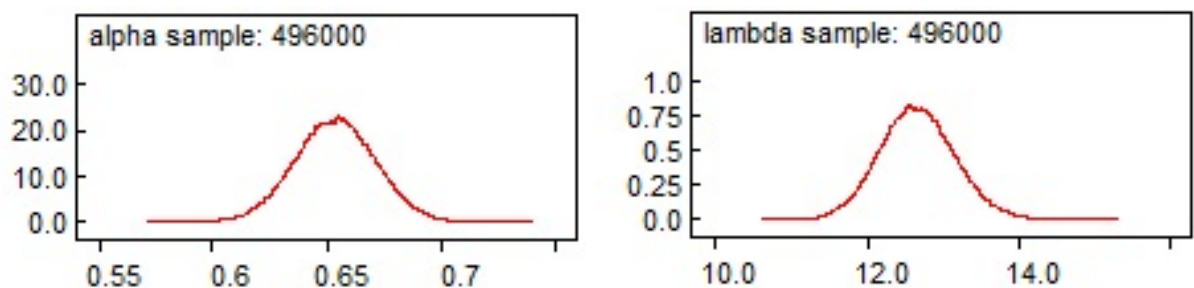


Figura 4.38: Função marginal à posteriori dos parâmetros $\alpha = 0,6529$ e $\lambda = 12,66$, do Modelo 5, com 500000 iterações, sendo 4000 de *Burn-in*.

Fixando o valor de α em 0,25 e realizando 50000 e 500000 iterações para cada modelo com abordagem bayesiana, verificou-se que o modelo 4 apresentou menor amplitude de variação para o valor a posterior de α e λ .

Tabela 4.8: Amplitude dos parâmetros α e λ variando quantidade de iterações

Modelo	Amplitude de α	Amplitude de λ
Modelo 2	0,0006	0,03
Modelo 3	0,0004	0,02
Modelo 4	0,0001	0,00
Modelo 5	0,0001	0,03

5 Conclusão

Este trabalho tratou da utilização de técnicas de modelagem e previsão em séries temporais de valores discretos de ordem 1, com abordagem clássica e abordagem bayesiana. A série é referente ao número de dias com precipitação no município de Garanhuns, agreste meridional de Pernambuco. Os modelos INAR(1) utilizados consideram a característica de contagem dos dados, retornando em suas simulações apenas valores inteiros e positivos, obtendo em suas modelagens e previsões valores que acompanharam as séries de contagem. Os modelos INAR explicaram as características da série apenas com os dados da própria série, baseados nas características probabilísticas das mesmas. A comparação entre os modelos INAR(1) Poisson e INAR(1) Poisson-Normal, foi realizada através da análise de erros de previsão e pelos critérios AIC, AICc e BIC.

No trabalho apresentado por Silva, Pereira e Silva (2009) que compara os modelos INAR(1) Poisson e INAR(1) Binomial Negativa, e concluem que o Poisson com priori Normal e Poisson com priori Beta, apresentam melhores ajustes aos dados. Neste complementamos seus resultados e verificou-se que o modelo INAR(1) Poisson-Normal priori Beta se adequou melhor aos dados, em comparação aos modelos INAR(1) Poisson priori Normal, INAR(1) Poisson priori Beta, INAR(1) Poisson-Normal Priori Normal, sendo também o que apresentou menor esforço computacional na estimação dos parâmetros.

Com a distribuição da média aproximando de uma distribuição Normal, utilizou-se o teste de Anderson-Darling e comparou-se com a distribuição Beta proposta por Silva, Pereira e Silva (2009). Verificou-se que os dados se adequaram melhor a distribuição Beta. Assim, ajustou-se a Poisson com a Normal usando prioris Normal e Beta, tendo em todos os critérios de informação (AIC, AICc e BIC) o Modelo 4 como sendo o melhor modelo ajustado.

Conclui-se que na comparação de abordagem clássica e bayesiana, os modelos bayesianos apresentaram resultados mais consistentes. Na comparação entre os modelos bayesianos, concluímos que os modelos mistos apresentaram melhores ajustes que os utilizados

na literatura. Nas análises complementares (parte I) o Modelo 4 apresentou menor sensibilidade na mudança do valor de α , enquanto na parte II verificou-se que o aumento no número de iterações, após ocorrer a convergência, pouco influência no ajuste dos modelos. Por fim, deve-se enfatizar que o uso de técnicas de simulação MCMC torna o processo de inferência bayesiana mais poderoso e flexível. Além disso pode ser estendido para problemas de dimensão maior.

Recomendações para trabalhos futuros:

- Analisar mais modelos mistos aplicado ao INAR;
- Estender análise para modelos de ordem maior;
- Implementar modelos mistos INAR para médias móveis INMA (Médias Móveis) e SINARIMA (Sazonal).

Referências

- AL-OSH, M. A.; ALZAID, A. A. First-order integer-valued autoregressive (inar (1)) process. **Journal of Time Series Analysis**, Wiley Online Library, v. 8, n. 3, p. 261–275, 1987.
- ANTELMAN, G. **Elementary bayesian statistics**. [S.l.]: Edward Elgar Publishing, 1997.
- APAC. **Agência Pernambucana de Águas e Clima**. 2013. Disponível em: <http://www.apac.pe.gov.br/>. Acessado em: 18/09/2013.
- BERGER, J. O. **Statistical decision theory and Bayesian analysis (2nd edition)**. [S.l.]: Springer, 1995.
- BERGER, J. O.; INSUA, D. R. **Recent developments in Bayesian inference with applications in hydrology**. [S.l.]: Consiglio nazionale delle ricerche, 1996.
- BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. [S.l.]: John Wiley & Sons, 2011.
- BOZDOGAN, H. Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. **Psychometrika**, Springer, v. 52, n. 3, p. 345–370, 1987.
- CARLIN, B. P.; LOUIS, T. A. Bayes and empirical bayes methods for data analysis. **Statistics and Computing**, Springer, v. 7, n. 2, p. 153–154, 1997.
- CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. **The American Statistician**, Taylor & Francis, v. 46, n. 3, p. 167–174, 1992.
- CHIB, S.; GREENBERG, E. Understanding the metropolis-hastings algorithm. **The American Statistician**, Taylor & Francis Group, v. 49, n. 4, p. 327–335, 1995.
- DU, J.-G.; LI, Y. The integer-valued autoregressive (inar (p)) model. **Journal of Time Series Analysis**, Wiley Online Library, v. 12, n. 2, p. 129–142, 1991.
- EHLERS, R. S. **Análise de séries temporais**. Instituto de Ciências Matemáticas e da Computação. Universidade de São Paulo, 2007.
- FERREIRA, T. A. E.; VASCONCELOS, G. C.; ADEODATO, P. J. L. A new evolutionary method for time series forecasting. In: **Proceedings of the 2005 conference on Genetic and evolutionary computation**. New York, NY, USA: ACM, 2005. (GECCO '05), p. 2221–2222.
- FREELAND, R. K.; MCCABE, B. P. Forecasting discrete valued low count time series. **International Journal of Forecasting**, Elsevier, v. 20, n. 3, p. 427–434, 2004.

- GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical science**, JSTOR, p. 457–472, 1992.
- GEWEKE, J. et al. **Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments**. [S.l.]: Federal Reserve Bank of Minneapolis, Research Department, 1991.
- GUIMARÃES, R.; SANTOS, E. G. Principles of stochastic generation of hydrologic time series for reservoir planning and design: A case study. **Journal of Hydrologic Engineering**, v. 16, n. 11, p. 891–898, 2011.
- IBGE. **Instituto Brasileiro de Geografia e Estatística**. 2013. Disponível em: <http://www.ibge.gov.br/home/>. Acessado em: 18/09/2013.
- INMET. **Instituto Nacional de Meteorologia**. 2014. Disponível em: <http://www.inmet.gov.br/portal/>. Acessado em: 10/05/2014.
- JEFFREYS, H. **The theory of probability**. [S.l.]: Oxford University Press, 1998.
- JEFFREYS, S. H. **Theory of Probability: 3d Ed**. [S.l.]: Clarendon Press, 1967.
- JOHNSON, L. A.; MONTGOMERY, D. C. **Forecasting and time series analysis**. [S.l.]: New York, Mc Graw Hill Book Company, 1976.
- LATOURE, A. Existence and stochastic structure of a non-negative integer-valued autoregressive process. **Journal of Time Series Analysis**, Wiley Online Library, v. 19, n. 4, p. 439–455, 1998.
- LEE, J. Y.; LEE, K. K. Use of hydrologic time series data for identification of recharge mechanism in a fractured bedrock aquifer system. **Journal of Hydrologic**, Elsevier Science BV, n. 229, p. 190–201, 2000.
- LIMA, T. d. A. C. **Modelos INAR e RCINAR, estimação e aplicação**. Tese (Doutorado) — Universidade de São Paulo, 2013.
- MARTÍNEZ-ESPINOSA, M.; JÚNIOR, C. C.; LAHR, F. A. R. Métodos paramétricos e não-paramétricos para determinar o valor característico em resultados de ensaio de madeira parametric and non-parametric methods to determine the characteristic value in wood tests results. 2000.
- MARTINS, L. C. et al. Relação entre poluição atmosférica e atendimentos por infecção de vias aéreas superiores no município de São Paulo: avaliação do rodízio de veículos. **Rev Bras Epidemiol**, SciELO Brasil, v. 4, n. 3, p. 220–9, 2001.
- MCKENZIE, E. Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. **Advances in Applied Probability**, JSTOR, p. 679–705, 1986.
- NIST. **Engineering Statistics and Handbook of statistical methods**. [S.l.: s.n.], 2002.
- PAVLOPOULOS, H.; KARLIS, D. Inar(1) modeling of overdispersed count series with an environmental application. **Environmetrics**, v. 19, n. 229, p. 369–393, 2007.

- ROMEU, J. L. Anderson-darling: A goodness of fit test for small samples assumptions. **Selected Topics in Assurance Related Technologies**, v. 10, n. 5, p. 1–6, 2003.
- SHIKLOMANOV, I. et al. The dynamics of river water inflow to the arctic ocean. In: **The Freshwater Budget of the Arctic Ocean**. [S.l.]: Springer, 2000. p. 281–296.
- SILVA, I.; SILVA, M. E. Seleção de ordem em modelos inar. **Actas do X Congresso Anual da SPE**, 2003.
- SILVA, I. M. M. d. Contributions to the analysis of discrete-valued time series. Universidade do Porto. Reitoria, 2005.
- SILVA, L. C.; SUÁREZ, P. ¿ qué es la inferencia bayesiana. **Jano**, v. 58, p. 65–66, 2000.
- SILVA, N.; PEREIRA, I.; SILVA, M. E. Forecasting in inar (1) model. Instituto Nacional de Estatística, 2009.
- SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002.
- STEPHENS, M. A. Edf statistics for goodness of fit and some comparisons. **Journal of the American statistical Association**, Taylor & Francis, v. 69, n. 347, p. 730–737, 1974.
- STEUTEL, F.; HARN, K. V. Discrete analogues of self-decomposability and stability. **The Annals of Probability**, JSTOR, p. 893–899, 1979.
- STEVENSON, W. J. **Production / operations management**. [S.l.]: Chicago: Irwin, 1996. (The Irwin series in production operations management). ISBN 9780256197235.
- TRENBERTH, K. E.; DAI, A. Effects of mount pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering. **Geophysical Research Letters**, Wiley Online Library, v. 34, n. 15, 2007.

Apêndice

Modelo WinBUGS

INAR(1) Poisson priori Beta

```

1 model {
2 C<-10000
3 mu[1]<- beta + alpha*y0
4 y[1] ~ dpois(mu[1])
5 for(t in 2:n) {
6 mu[t] <- beta + alpha*y[t-1]
7   y[t]~ dpois(mu[t])
8 }
9
10 for(i in 2:n){
11   zeros[i]<-0
12   zeros[i]~ dpois(aux[i])
13   aux[i]<- C-loglike[i]
14   loglike[i]<- log(beta) + (y[i]-1)*log(beta+alpha*y[i]) -
15     (beta+alpha*y[i]) - logfact(y[i])
16 like[i]<- exp(loglike[i])
17   }
18   dm<- 2 + g[1] +g[2]
19
20   Deviance<- -2*sum(loglike[2:n])
21
22   AIC <- Deviance + dm*2
23   BIC <- Deviance + dm*log(n)
24   AICC<- n*log(Deviance)*.9 + n*((1+dm/n)/(1-(dm/n)))
25   alpha ~ dbeta(0.000001, 1.0E-6)
26   beta ~ dnorm(0.0, 1.0E-6)
27   lambda<-beta/(1-alpha)
28   var<-beta/pow(1-alpha,3)
29   DI<- 1/((1-alpha)*(1-alpha))
30 y0~ dnorm(0.0,1.0E-6)
31   %#!lambda ~ dgamma(0.01,0.01)
32 }
33

```

```

34 list(alpha=0.25, beta=0.25, y0=20.0)
35
36 list(g=c(1,1), y=c(5,4,4,4,11,13,24,25,29,24,18,6,5,7,0,
37 6,4,16,16,27,30,22,12,3,10,2,3,4,5,18,22,24,25,
38 23,14,5,11,2,6,12,14,20,22,17,27,24,2,0,1,7,3,
39 4,6,8,20,18,24,22,8,4,1,0,2,6,4,5,17,18,25,21,14,10,0,9,
40 11,5,8,18,20,23,25,22,21,5,7,7,6,2,9,13,4,26,28,27,12,7,
41 2,14,21,8,15,15,18,29,15,26,12,7,6,3,6,13,10,12,15,20,27,
42 19,15,12,4,2,20,12,9,15,17,29,27,21,18,4,5,1,3,8,9,13,23,
43 25,24,25,15,3,1,5,4,3,8,9,23,25,26,20,8,5,8,1,5,11,12,12,20,
44 25,20,25,17,3,4,3,2,7,16,13,23,21,26,21,7,8,0,1,3,9,4,13,25,
45 20,23,23,4,2,3,7,12,8,6,12,11,25,26,26,14,7,2,4,10,6,3,21,24,
46 19,29,14,17,8,9,4,6,5), n=220)

```

INAR(1) Poisson priori Normal

```

1 model {
2 C<-10000
3 mu[1]<- beta+ alpha*y0
4 y[1]~ dpois(mu[1])
5
6 for(t in 2:n) {
7 mu[t] <- beta + alpha*y[t-1]
8   y[t]~ dpois(mu[t])
9 }
10
11 for(i in 2:n){
12   b[i]<-0
13   zeros[i]<-0
14   zeros[i]~ dpois(aux[i])
15   aux[i]<- C-loglike[i]
16   loglike[i]<- log(beta) + (y[i]-1)*log(beta+b[i]+alpha*y[i])
17   -(beta+b[i]+alpha*y[i])-logfact(y[i])
18 like[i]<- exp(loglike[i])
19   b[i]~ dnorm(0,1)
20 }
21 dm<- 2 + g[1] +g[2]
22
23 Deviance<- -2*sum(loglike[2:n])
24
25 AIC <- Deviance + dm*2
26 BIC <- Deviance + dm*log(n)
27

```

```

28     alpha ~ dnorm(0.0, 1.0E-6)
29     beta  ~ dnorm(0.0, 1.0E-6)
30     lambda<-beta/(1-alpha)
31     var<-beta/pow(1-alpha,3)
32     DI<- 1/((1-alpha)*(1-alpha))
33 y0~ dnorm(0.0,1.0E-6)
34     %#lambda~dgamma(0.01,0.01)
35
36 }
37
38 list(alpha=0.25, beta=0.5, y0=20.0)
39
40 list(g=c(0,0), y=c(5,4,4,4,11,13,24,25,29,24,18,6,5,7,0,
41 6,4,16,16,27,30,22,12,3,10,2,3,4,5,18,22,24,25,
42 23,14,5,11,2,6,12,14,20,22,17,27,24,2,0,1,7,3,
43 4,6,8,20,18,24,22,8,4,1,0,2,6,4,5,17,18,25,21,14,10,0,9,
44 11,5,8,18,20,23,25,22,21,5,7,7,6,2,9,13,4,26,28,27,12,7,
45 2,14,21,8,15,15,18,29,15,26,12,7,6,3,6,13,10,12,15,20,27,
46 19,15,12,4,2,20,12,9,15,17,29,27,21,18,4,5,1,3,8,9,13,23,
47 25,24,25,15,3,1,5,4,3,8,9,23,25,26,20,8,5,8,1,5,11,12,12,20,
48 25,20,25,17,3,4,3,2,7,16,13,23,21,26,21,7,8,0,1,3,9,4,13,25,
49 20,23,23,4,2,3,7,12,8,6,12,11,25,26,26,14,7,2,4,10,6,3,21,24,
50 19,29,14,17,8,9,4,6,5), n=220)

```

INAR(1) Poisson-Normal priori Beta

```

1 model{
2   # INAR(1) model
3   c<- 10000
4   mu[1] <- beta + alpha*y0
5   y[1] ~ dpois(mu[1])
6
7   for(t in 2:n) {
8     mu[t] <- beta + alpha*y[t-1]
9     y[t] ~ dpois(mu[t])
10  }
11  for(i in 1:n){
12    b[i]<-0
13    logvero[i]<--y[i]*(lambda+b[i])+delta[i]*log(lambda+b[i])
14    b[i]~dnorm(0,1)
15    zeros[i]<-0
16    aux[i]<--logvero[i]+c
17    zeros[i]~dpois(aux[i])

```

```

18         }
19
20 dm<- 2 + g[1] + g[2]
21
22         Deviance<- -2*sum(logvero[1:n])
23
24         AIC <- Deviance + dm*2
25         BIC <- Deviance + dm*log(n)
26 AICC<- n*log(Deviance)*.9 + n*((1+dm/n)/(1-(dm/n)))
27
28 # Prior distribution
29
30 beta~dgamma(0.001,0.001)
31     alpha~dbeta(1,1)
32 tau  ~ dgamma(0.01, 0.01)
33 y0   ~ dnorm(0.0,1.0E-6)
34 lambda~dgamma(0.000001,0.000001)
35 esperanca<- 1/lambda
36 mean<-alpha/(1-beta)
37 var<-alpha/pow(1-beta,3)
38     s2<-1/tau
39     s <-sqrt(s2)
40 }
41
42 list(beta=0.25,alpha=0.25,y0=20.0, lambda=1.0,tau=1.0)
43
44 list(n=220,g=c(1,0),y=c(5,4,4,4,11,13,24,25,29,24,18,6,5,7,0,
45 6,4,16,16,27,30,22,12,3,10,2,3,4,5,18,22,24,25,
46 23,14,5,11,2,6,12,14,20,22,17,27,24,2,0,1,7,3,
47 4,6,8,20,18,24,22,8,4,1,0,2,6,4,5,17,18,25,21,14,10,0,9,
48 11,5,8,18,20,23,25,22,21,5,7,7,6,2,9,13,4,26,28,27,12,7,
49 2,14,21,8,15,15,18,29,15,26,12,7,6,3,6,13,10,12,15,20,27,
50 19,15,12,4,2,20,12,9,15,17,29,27,21,18,4,5,1,3,8,9,13,23,
51 25,24,25,15,3,1,5,4,3,8,9,23,25,26,20,8,5,8,1,5,11,12,12,20,
52 25,20,25,17,3,4,3,2,7,16,13,23,21,26,21,7,8,0,1,3,9,4,13,25,
53 20,23,23,4,2,3,7,12,8,6,12,11,25,26,26,14,7,2,4,10,6,3,21,24,
54 19,29,14,17,8,9,4,6,5),delta=c(0.04,0.04,0.04,0,0,0,0,0,0,
55 0,0.01,0,1,0,0.04,0,0,0,0,0,0,0.17,0,0.5,0.17,0.04,0.01,0,0,
56 0,0,0,0,0.01,0,0.5,0,0,0,0,0,0,0,0,0.5,1,1,0,0.17,0.04,0,0,0,
57 0,0,0,0,0.04,1,1,0.5,0,0.04,0.01,0,0,0,0,0,0,1,0,0,0.01,0,0,
58 0,0,0,0,0,0,0,0,0.5,0,0,0.04,0,0,0,0,0.5,0,0,0,0,0,0,0,0,
59 0,0,0,0,0.17,0,0,0,0,0,0,0,0,0.04,0.5,0,0,0,0,0,0,0,0,0,
60 0.04,0.01,1,0.17,0,0,0,0,0,0,0,0.17,1,0.01,0.04,0.17,0,0,0,

```

```

61 0,0,0,0,0.01,0,1,0.01,0,0,0,0,0,0,0,0.17,0.04,0.17,0.5,0,0,
62 0,0,0,0,0,0,0,1,1,0.17,0,0.04,0,0,0,0,0.04,0.5,0.17,0,0,0,0,
63 0,0,0,0,0,0,0,0.5,0.04,0,0,0.17,0,0,0,0,0,0.04,0,0.01,0.04))

```

INAR(1) Poisson-Normal priori Normal

```

1  model{
2    # INAR(1) model
3    c<- 10000
4    mu[1] <- beta + alpha*y0
5    y[1] ~ dpois(mu[1])
6
7    for(t in 2:n) {
8      mu[t] <- beta + alpha*y[t-1]
9      y[t] ~ dpois(mu[t])
10   }
11
12   for(i in 1:n){
13     b[i]<-0
14     logvero[i]<--y[i]*(lambda+b[i])+delta[i]*log(lambda+b[i])
15     b[i]~dnorm(0,1)
16     zeros[i]<-0
17     aux[i]<--logvero[i]+c
18     zeros[i]~dpois(aux[i])
19   }
20
21   dm<- 2 + g[1] + g[2]
22
23     Deviance<- -2*sum(logvero[1:n])
24
25     AIC <- Deviance + dm*2
26     BIC <- Deviance + dm*log(n)
27   AICC<- n*log(Deviance)*.9 + n*((1+dm/n)/(1-(dm/n)))
28
29   # Prior distribution
30
31   beta ~ dnorm(0.0, 1.0E-6)
32   alpha ~ dnorm(0.0, 1.0E-6)
33   tau ~ dgamma(0.01, 0.01)
34   y0 ~ dnorm(0.0,1.0E-6)
35   lambda~dgamma(0.000001,0.000001)
36   esperanca<- 1/lambda
37   mean<-alpha/(1-beta)

```

```

38 var<-alpha/pow(1-beta,3)
39     s2<-1/tau
40         s <-sqrt(s2)
41 }
42
43 list(beta=0.25,alpha=0.25,y0=20.0, lambda=1.0,tau=1.0)
44
45 list(n=220,g=c(1,0),y=c(5,4,4,4,11,13,24,25,29,24,18,6,5,7,0,
46 6,4,16,16,27,30,22,12,3,10,2,3,4,5,18,22,24,25,
47 23,14,5,11,2,6,12,14,20,22,17,27,24,2,0,1,7,3,
48 4,6,8,20,18,24,22,8,4,1,0,2,6,4,5,17,18,25,21,14,10,0,9,
49 11,5,8,18,20,23,25,22,21,5,7,7,6,2,9,13,4,26,28,27,12,7,
50 2,14,21,8,15,15,18,29,15,26,12,7,6,3,6,13,10,12,15,20,27,
51 19,15,12,4,2,20,12,9,15,17,29,27,21,18,4,5,1,3,8,9,13,23,
52 25,24,25,15,3,1,5,4,3,8,9,23,25,26,20,8,5,8,1,5,11,12,12,20,
53 25,20,25,17,3,4,3,2,7,16,13,23,21,26,21,7,8,0,1,3,9,4,13,25,
54 20,23,23,4,2,3,7,12,8,6,12,11,25,26,26,14,7,2,4,10,6,3,21,24,
55 19,29,14,17,8,9,4,6,5),delta=c(0.04,0.04,0.04,0,0,0,0,0,0,0,0,
56 0.01,0,1,0,0.04,0,0,0,0,0,0,0.17,0,0.5,0.17,0.04,0.01,0,0,0,0,0,
57 0,0.01,0,0.5,0,0,0,0,0,0,0,0.5,1,1,0,0.17,0.04,0,0,0,0,0,0.04,
58 1,1,0.5,0,0.04,0.01,0,0,0,0,0,0,1,0,0,0.01,0,0,0,0,0,0,0,0.5,
59 0,0,0.04,0,0,0,0,0.5,0,0,0,0,0,0,0,0.17,0,0,0,0,0,0,0,
60 0,0,0.04,0.5,0,0,0,0,0,0,0.04,0.01,1,0.17,0,0,0,0,0,0.17,
61 1,0.01,0.04,0.17,0,0,0,0,0,0.01,0,1,0.01,0,0,0,0,0,0.17,0.04,
62 0.17,0.5,0,0,0,0,0,0,0,1,1,0.17,0,0.04,0,0,0,0,0.04,0.5,0.17,0,0,
63 0,0,0,0,0,0,0,0.5,0.04,0,0,0.17,0,0,0,0,0,0,0.04,0,0.01,0.04))

```