

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
DEPARTAMENTO DE ESTATÍSTICA

**Estimação de Máxima Verossimilhança para Processo de
Nascimento Puro Espaço-Temporal com Dados
Parcialmente Observados**

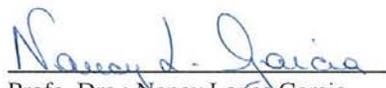
Daniela Bento Fonsechi Goto
Dissertação de Mestrado orientada pela
Profa. Dra. Nancy Lopes Garcia

Campinas - SP
2008

Estimação de Máxima Verossimilhança para Processo de Nascimento
Puro Espaço-Temporal com Dados Parcialmente Observados

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Daniela Bento Fonseca Goto e aprovada pela comissão julgadora.

Campinas, 10 de setembro de 2008


Profa. Dra.: Nancy Lopes Garcia
Orientador

Banca Examinadora:

1. Nancy Lopes Garcia – IMECC - UNICAMP
2. Profa. Dra. Alexandra Mello Schmidt – DME - UFRJ
3. Prof. Dr. Jesus Enrique Garcia – IMECC - UNICAMP.

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de MESTRE em Estatística.

**FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DO IMECC DA UNICAMP
Bibliotecária: Maria Júlia Milani Rodrigues – CRB8a 2116**

Goto, Daniela Bento Fonseca

G712e Estimação de máxima verossimilhança para processo de
nascimento puro espaço-temporal com dados parcialmente observados /
Daniela Bento Fonseca Goto -- Campinas, [S.P. :s.n.], 2008.

Orientadora : Nancy Lopes Garcia

Dissertação (mestrado) - Universidade Estadual de Campinas,
Instituto de Matemática, Estatística e Computação Científica.

1. Método da projeção. 2. Estimador de máxima verossimilhança.
3. Dados faltantes (Estatística). . I. Garcia, Nancy Lopes. II.
Universidade Estadual de Campinas. Instituto de Matemática, Estatística
e Computação Científica. III. Título.

Título em inglês: Maximum likelihood estimation for space-time pure birth process with missing data.

Palavras-chave em inglês (Keywords): 1. Projection method. 2. Maximum likelihood estimation. 3. Missing data (Statistics).

Área de concentração: Inferência em Processos Estocásticos

Titulação: Mestre em Estatística

Banca examinadora:

Profa. Dra. Nancy Lopes Garcia (IMECC-UNICAMP)

Profa. Alexandra Mello Schmidt (DME-UFRJ)

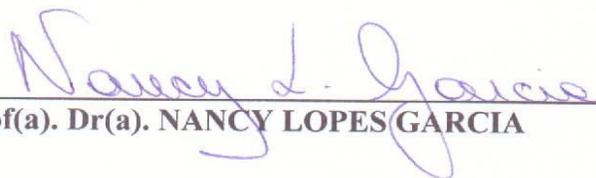
Prof. Dr. Jesus Enrique Garcia (IMECC-UNICAMP)

Data da defesa: 10/09/2008

Programa de pós-graduação: Mestrado em Estatística

Dissertação de Mestrado defendida em 10 de setembro de 2008 e aprovada

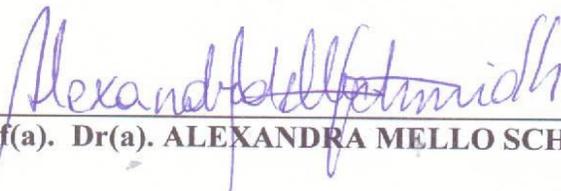
Pela Banca Examinadora composta pelos Profs. Drs.



Prof(a). Dr(a). NANCY LOPES GARCIA



Prof(a). Dr(a). JESUS ENRIQUE GARCIA



Prof(a). Dr(a). ALEXANDRA MELLO SCHMIDT

Ao meu amado marido, Samuel Goto
e aos meus queridos pais, Jackson e Márcia.

Agradecimentos

Primeiramente à Deus e ao meu Mentor que sempre estiveram do meu lado me dando força e inspiração para a realização deste projeto.

À minha mãe Márcia, ao meu pai Jackson e à minha irmã Thaís que desde o início acreditaram na minha escolha e me apoiaram me dando amor, carinho e incentivo, possibilitando, desta forma, a realização deste trabalho.

Ao meu amado marido Samuel pelo imenso amor, apoio e pela enorme paciência. Mesmo longe, sempre estive ao meu lado com sábias palavras de incentivo e carinho. Agradeço, também, pela sua ajuda com minhas dúvidas em programação.

Aos meus queridos amigos da graduação pelos momentos de descontração e por todo o incentivo.

Aos queridos e inesquecíveis amigos que fiz durante o mestrado, em especial, à Camila Estevam que compartilhou momentos difíceis e felizes, uma grande amiga que ficará para sempre.

À professora Dra. Nancy Lopes Garcia pela sua excelente orientação, desde a graduação, e sua enorme paciência em me explicar sobre o misterioso e complicado estudo da probabilidade. Extremamente dedicada e compreensiva, sempre disposta a conversar, apoiar e aconselhar.

À todos os professores do IMECC que contribuíram para minha formação.

À professora Denise Duarte, da UFMG, por gentilmente ter emprestado sua sala para que eu pudesse trabalhar enquanto estive em Belo Horizonte.

Ao IMECC e ao departamento de Estatística pelo espaço físico concedido.

À toda minha família (Bento, Fonsechi, Fukujima e Goto) que esperaram ansiosamente para o término deste projeto.

Ao tio Gordão que ajudou nas simulações.

À Alexandra Mello Schmitd e Jesus Enrique Garcia por participarem da banca examinadora.

À CAPES pela bolsa concedida.

Resumo

O objetivo desta dissertação é estudar estimação de máxima verossimilhança para processos de nascimento puro espacial para dois diferentes tipos de amostragem: a) quando há observação permanente em um intervalo $[0, T]$; b) quando o processo é observado após um tempo T fixo. No caso b) não se conhece o tempo de nascimento dos pontos, somente sua localização (dados faltantes). A função de verossimilhança pode ser escrita para o processo de nascimento puro não homogêneo em um conjunto compacto através do método da projeção descrito por Garcia and Kurtz (2008), como projeção da função de verossimilhança. A verossimilhança projetada pode ser interpretada como uma esperança e métodos de Monte Carlo podem ser utilizados para estimar os parâmetros. Resultados sobre convergência quase-certa e em distribuição são obtidos para a aproximação do estimador de máxima verossimilhança. Estudos de simulação mostram que as aproximações são adequadas.

Abstract

The goal of this work is to study the maximum likelihood estimation of a spatial pure birth process under two different sampling schemes: a) permanent observation in a fixed time interval $[0, T]$; b) observation of the process only after a fixed time T . Under scheme b) we don't know the birth times, we have a problem of missing variables. We can write the likelihood function for the non-homogeneous pure birth process on a compact set through the method of projection described by Garcia and Kurtz (2008), as the projection of the likelihood function. The fact that the projected likelihood can be interpreted as an expectation suggests that Monte Carlo methods can be used to compute estimators. Results of convergence almost surely and in distribution are obtained for the approximations to the maximum likelihood estimator. Simulation studies show that the approximations are appropriate.

Sumário

1	Introdução	1
2	Processos pontuais obtidos através de transformações de processos de Poisson	5
2.1	Processo de Poisson	5
2.2	Método da projeção e transformação de processos de Poisson	6
2.2.1	Projeções através de conjuntos aleatórios	7
2.3	Verossimilhança para processos pontuais	8
2.4	Dados faltantes	9
2.4.1	Algoritmo EM	9
2.4.2	Estimador de máxima verossimilhança via método de Monte Carlo	10
3	Processo de nascimento puro	11
3.1	Método da projeção para processos de nascimento puro	11
3.2	Verossimilhança para um processo observado após um período de tempo fixo	13
4	Simulações	17
4.1	Dados Completos	17
4.2	Processo após um tempo fixo T	22
4.2.1	Algoritmo Metropolis-Hastings	22
4.2.2	Simulação com $\beta_0 = 0$	23
4.2.3	Simulação Metropolis-Hastings com $\beta_0 = 1$ e proposta uniforme	23
4.2.4	Simulação Metropolis-Hastings com $\beta = 1$ e proposta beta	27
5	Conclusão	31
A	Processos pontuais univariados - Algumas definições	33
B	Gráficos	34

Lista de Figuras

2.1	Projeção de um processo de Poisson.	7
3.1	Projeção de um processo de Poisson.	12
4.1	Nascimentos até tempo 1.	19
4.2	Nascimentos até tempo 2.	19
4.3	Nascimentos até tempo 3.	19
4.4	Nascimentos até tempo 4.	19
4.5	Nascimentos até tempo 5.	20
4.6	Histograma das estimativas de β para os dados completos.	21
4.7	Histograma da estimativa dos β para dados incompletos.	24
4.8	Diagnóstico para as cadeias 1 à 4 - proposta uniforme.	25
4.9	Acf para as cadeias 1 à 4 - proposta uniforme.	25
4.10	Acf para as cadeias 1 à 4 com 250 dados - proposta uniforme.	26
4.11	Histograma da estimativa dos β para dados incompletos - proposta uniforme.	26
4.12	Diagnóstico para as cadeias 1 à 4 - proposta beta.	27
4.13	Acf para as cadeias 1 à 4 - proposta beta.	28
4.14	Acf para as cadeias 1 à 4 com 250 dados - proposta beta.	28
4.15	Histograma da estimativa dos β para dados incompletos - proposta beta.	29
B.1	Diagnóstico para as cadeias 5 à 8 - proposta uniforme.	34
B.2	Diagnóstico para as cadeias 9 à 12 - proposta uniforme.	35
B.3	Diagnóstico para as cadeias 13 à 16 - proposta uniforme.	35
B.4	Diagnóstico para as cadeias 17 à 20 - proposta uniforme.	36
B.5	Diagnóstico para as cadeias 21 à 24 - proposta uniforme.	36
B.6	Diagnóstico para as cadeias 25 à 28 - proposta uniforme.	37
B.7	Diagnóstico para as cadeias 29 à 32 - proposta uniforme.	37
B.8	Diagnóstico para as cadeias 33 à 36 - proposta uniforme.	38
B.9	Diagnóstico para as cadeias 37 à 40 - proposta uniforme.	38

B.10 Diagnóstico para as cadeias 5 à 8 - proposta beta.	39
B.11 Diagnóstico para as cadeias 9 à 12 - proposta beta.	39
B.12 Diagnóstico para as cadeias 13 à 16 - proposta beta.	40
B.13 Diagnóstico para as cadeias 17 à 20 - proposta beta.	40
B.14 Diagnóstico para as cadeias 21 à 24 - proposta beta.	41
B.15 Diagnóstico para as cadeias 25 à 28 - proposta beta.	41
B.16 Diagnóstico para as cadeias 29 à 32 - proposta beta.	42
B.17 Diagnóstico para as cadeias 33 à 36 - proposta beta.	42
B.18 Diagnóstico para as cadeias 37 à 40 - proposta beta.	43
B.19 Acf para as cadeias 5 à 8 - proposta uniforme.	44
B.20 Acf para as cadeias 9 à 12 - proposta uniforme.	44
B.21 Acf para as cadeias 13 à 16 - proposta uniforme.	45
B.22 Acf para as cadeias 17 à 20 - proposta uniforme.	45
B.23 Acf para as cadeias 21 à 24 - proposta uniforme.	46
B.24 Acf para as cadeias 25 à 28 - proposta uniforme.	46
B.25 Acf para as cadeias 29 à 32 - proposta uniforme.	47
B.26 Acf para as cadeias 33 à 36 - proposta uniforme.	47
B.27 Acf para as cadeias 37 à 40 - proposta uniforme.	48
B.28 Acf para as cadeias 5 à 8 com 250 dados- proposta uniforme.	48
B.29 Acf para as cadeias 9 à 12 com 250 dados- proposta uniforme.	49
B.30 Acf para as cadeias 13 à 16 com 250 dados- proposta uniforme.	49
B.31 Acf para as cadeias 17 à 20 com 250 dados- proposta uniforme.	50
B.32 Acf para as cadeias 21 à 24 com 250 dados- proposta uniforme.	50
B.33 Acf para as cadeias 25 à 28 com 250 dados- proposta uniforme.	51
B.34 Acf para as cadeias 29 à 32 com 250 dados- proposta uniforme.	51
B.35 Acf para as cadeias 33 à 36 com 250 dados- proposta uniforme.	52
B.36 Acf para as cadeias 37 à 40 com 250 dados- proposta uniforme.	52
B.37 Acf para as cadeias 5 à 8 - proposta beta.	53
B.38 Acf para as cadeias 9 à 12 - proposta beta.	53
B.39 Acf para as cadeias 13 à 16 - proposta beta.	54
B.40 Acf para as cadeias 17 à 20 - proposta beta.	54
B.41 Acf para as cadeias 21 à 24 - proposta beta.	55
B.42 Acf para as cadeias 25 à 28 - proposta beta.	55
B.43 Acf para as cadeias 29 à 32 - proposta beta.	56
B.44 Acf para as cadeias 33 à 36 - proposta beta.	56
B.45 Acf para as cadeias 37 à 40 - proposta beta.	57

B.46 Acf para as cadeias 5 à 8 com 250 dados - proposta beta.	57
B.47 Acf para as cadeias 9 à 12 com 250 dados- proposta beta.	58
B.48 Acf para as cadeias 13 à 16 com 250 dados - proposta beta.	58
B.49 Acf para as cadeias 17 à 20 com 250 dados - proposta beta.	59
B.50 Acf para as cadeias 21 à 24 com 250 dados - proposta beta.	59
B.51 Acf para as cadeias 25 à 28 com 250 dados - proposta beta.	60
B.52 Acf para as cadeias 29 à 32 com 250 dados - proposta beta.	60
B.53 Acf para as cadeias 33 à 36 com 250 dados - proposta beta.	61
B.54 Acf para as cadeias 37 à 40 com 250 dados - proposta beta.	61

Lista de Tabelas

- 4.1 Análise descritiva da estimativa de β 21
- 4.2 Análise descritiva da estimativa de β para dados incompletos. 23
- 4.3 Análise descritiva da estimativa de β para dados incompletos - proposta uniforme. 27
- 4.4 Análise descritiva da estimativa de β para dados incompletos - proposta beta. 29

Capítulo 1

Introdução

Estudos que envolvem processos pontuais espaço-temporais de nascimento puro vêm crescendo devido ao interesse e à necessidade de compreender os eventos que ocorrem durante um intervalo de tempo, numa determinada região do espaço. Diggle (2003) estudou processos pontuais espaciais analisando estimativas da taxa de intensidade, Cox e Isham (1980) analisaram propriedades probabilísticas destes processos, Ogata(1998) fez estudos não paramétricos de processos pontuais espaciais, Cressie(1993) estudou estatística espacial usando dados geoestatísticos.

Pode-se citar aqui alguns exemplos de processos de nascimento puro que são de grande interesse à sociedade, como ocorrência de terremotos, epidemia de uma determinada doença na cidade, aparecimento de doença na agricultura, estudo da evolução de bactérias, etc. Há o interesse pelos cientistas, políticos e toda sociedade em se estimar a intensidade destes processos para entender como está ocorrendo o desenvolvimento do processo, por exemplo, entender como a ferrugem asiática (doença que ocorre em plantação de soja, veja www.cnpso.embrapa.br) contamina todo o plantio de soja em uma determinada região. Esta doença tem taxa de contaminação, por exemplo, $\lambda(x, n)$ e uma vez que a planta é contaminada, não há mais recuperação, portanto o controle deve acontecer antes da contaminação. Assim, utilizando a estatística espacial, é possível criar uma estratégia de proteção no plantio da soja, podendo implementar um sistema de alerta para detecção de possíveis novos pontos de contaminação e a previsão do comportamento espacial da doença.

Um processo pontual espaço-temporal é caracterizado como processo de nascimento puro quando se observa apenas novas ocorrências de determinado evento no espaço, que são chamados de nascimentos. Keiding (1974), Beyer, Keiding and Simonsen (1976), Moran (1951) estudaram estimação por máxima verossimilhança para processos de nascimento puro no caso unidimensional. No caso de processos pontuais espaciais os trabalhos são frequentemente na área seismológica, Ecologia, Geografia, Epidemiologia, entre outros, veja por exemplo Stoyan e Stoyan (1998), Kerscher (2000), Boots, Okabe and Thomas (2003), Geyer (1999), Geyer and Moller (1994). Alguns livros clássicos

em processos pontuais espaciais são Ripley (1981 e 1988), Diggle (1983), Stoyan, Kendall and Mecke (1995), Moller and Waagepetersen(2003).

Infelizmente dados completos (conjunto de dados que contém toda a informação da posição e do tempo de nascimento) são raros. É muito comum conhecer apenas as posições das ocorrências dos eventos, mas não os tempos de nascimentos destas ocorrências. Por isso existe a necessidade de se estudar processos de nascimento puro para dados parcialmente observados (Little e Rubin (2002)). Para se trabalhar com este tipo de problema o algoritmo EM (Dempster, Laird and Rubin (1977)) é muito usado para encontrar o estimador de máxima verossimilhança do modelo pois a cada iteração o algoritmo calcula a esperança da log-verossimilhança incluindo os dados que faltam como se eles tivessem sido observados e, assim, calcula o estimador de máxima verossimilhança maximizando a esperança da log-verossimilhança encontrada. Os passos são repetidos até atingir a convergência. Recentemente, com o avanço da tecnologia e com a disponibilidade de melhores computadores, métodos de Monte Carlo também são muito utilizados para trabalhar com problemas de dados faltantes, principalmente quando a forma fechada da verossimilhança completa não é conhecida. Neste caso, simulações utilizando o algoritmo Metropolis-Hastings podem ser implementados para gerar as realizações da variável não observada.

Um processo pontual espaço-temporal é um modelo de pontos indistinguíveis, distribuídos aleatoriamente em algum espaço e em determinado intervalo de tempo. Este trabalho baseia-se na idéia de que um processo pontual de nascimento puro em um compacto $K \subset \mathbb{R}^2$ pode ser obtido como projeção de um processo de Poisson em $\mathbb{R}^2 \times [0, \infty)$. A idéia é construir um processo de Poisson em $\mathbb{R}^2 \times [0, \infty)$, construir um subconjunto aleatório em $\mathbb{R}^2 \times [0, \infty)$ e projetar em \mathbb{R}^2 os pontos do processo de Poisson que pertencem ao subconjunto. Desta forma pode-se facilmente obter a função de verossimilhança e alguns resultados assintóticos. O método da projeção obtém processos pontuais não triviais através de processos de Poisson de dimensão maiores e foi introduzido por Kurtz (1989) e estudado por Garcia(1995a e 1995b).

O objetivo deste trabalho é obter estimação paramétrica pelo método de máxima verossimilhança sob dois diferentes tipos de amostragem para o processo de nascimento puro espaço-temporal:

(A) Observação permanente em um intervalo de tempo fixo $[0, T]$;

(B) Observação após um período T fixado.

Note que no item (B) são conhecidos apenas as localizações dos pontos observados, mas desconhecem-se os tempos de nascimentos, havendo, desta forma, o problema de dados faltantes.

No Capítulo 2 será explicado como é feita a transformação do processo de Poisson e como são obtidas projeções através de conjuntos aleatórios que foi caracterizado por Garcia (1995a). Serão descritos o método da projeção introduzido por Kurtz (1989), a função de verossimilhança para

processos pontuais utilizando o conceito da derivada de Radon-Nikodym, dados faltantes e métodos para trabalhar com este tipo de conjunto de dados.

No Capítulo 3 será descrito o método da projeção para processo de nascimento puro espaço-temporal e a função de verossimilhança sob o tipo de amostragem descrito no item (A) será obtida. Na Seção 3.2 a função de verossimilhança para o processo sob (B) será introduzida, a qual será dada como projeção da verossimilhança completa. O método Monte Carlo baseado em cadeias de Markov (do inglês MCMC) será descrito para obter uma aproximação para o estimador de máxima verossimilhança para o parâmetro β , parâmetro da função intensidade do processo de nascimento puro.

No Capítulo 4 serão apresentados simulações para mostrar a adequação dos métodos descritos.

Capítulo 2

Processos pontuais obtidos através de transformações de processos de Poisson

Processos de Poisson são os processos pontuais mais simples de se obter e mais interessantes. Na prática são muito utilizados para modelarem vários fenômenos naturais e, também, são os mais populares para modelos de contagem e, além disto, os processos de Poisson são a base para construção de processos mais complexos.

2.1 Processo de Poisson

Suponha deseja-se contar o número de eventos que ocorrem em determinada região do espaço em um certo intervalo de tempo, seja a região espaço-temporal $A \subseteq \mathbb{R}^d \times [0, \infty)$. Seja $N(A)$ o número de pontos na região A , um processo pontual $N : \mathcal{B}(\mathbb{R}^d \times [0, \infty)) \rightarrow \mathbb{N}$ é dito ser um processo de Poisson com medida média μ se as suposições (A1) e (A2) são satisfeitas:

- (A1) O número de eventos em regiões distintas do espaço são variáveis aleatórias independentes, ou seja, se A_1, \dots, A_m são disjuntos então $N(A_1), \dots, N(A_m)$ são variáveis aleatórias independentes;
- (A2) O número de eventos em certa região A do espaço tem distribuição de Poisson, ou seja, para $k > 0$,

$$P[N(A) = k] = e^{(-\mu(A))} \frac{\mu(A)^k}{k!}.$$

Definição 2.1 *A intensidade de um processo de Poisson N , com medida média, μ absolutamente*

contínua com respeito a medida de Lebesgue, é a função λ que satisfaz a relação

$$\mu(A) := \int \int_A \lambda(x, y) dx dy. \quad (2.1)$$

Definição 2.2 A função log-verossimilhança do processo de Poisson N_T na região $A \subseteq \mathbb{R}^d$, no intervalo de tempo $[0, T]$, é dada por

$$l(\theta) = \int_A \int_T \log \mu(s, t; \theta) dN_T(s, t) - \int_A \int_T \mu(s, t; \theta) ds dt \quad (2.2)$$

Veja Daley and Vere-Jones (1988) para mais detalhes.

2.2 Método da projeção e transformação de processos de Poisson

O método da projeção foi introduzido por Kurtz (1989) com o objetivo de construir processos pontuais não triviais em \mathbb{R}^d , como por exemplo processo de Cox, classe de distribuições de Gibbs através de projeções de processos de Poisson de dimensões maiores, através de subconjuntos aleatórios no espaço maior, projetando-se os pontos do processo de Poisson, que se localizam dentro destes subconjuntos em subespaços de dimensão menores. Esta construção pode ser utilizada para gerar processos de Cox, estados de Gibbs e processos de nascimento e morte com taxas variáveis. Garcia (1995a e 1995b) utilizou este método para construir processos de nascimento e morte com taxa variável de nascimento e morte e no comportamento assintótico para modelos de uma epidemia. Para uma revisão sobre o assunto veja Garcia e Kurtz (2008).

Seja o intervalo $[0, \infty)$, neste caso qualquer processo pontual simples N , tal que $N([0, t]) - \Lambda(t)$ é um martingal, pode ser obtido como uma mudança no tempo aleatório $N([0, t]) = Y(\Lambda(t))$ onde Y é um processo de Poisson com taxa unitária (Veja proposição 13.4.III, Daley and Vere-Jones, 1988).

O método da projeção, que foi introduzido por Kurtz (1989), pode ser visto como uma generalização da representação da mudança no tempo aleatório. O processo constrói processos pontuais através de projeções do processo de Poisson. Estes processos projetados são feitos cuidadosamente de tal maneira a herdarem muitas das boas propriedades dos processos de Poisson. Este método dá um acoplamento natural entre os processos pontuais, o qual nos fornece uma maneira para comparar resultados e teoremas limites. Leis dos grandes números e teoremas central do limite para processos de Poisson podem ser explorados para obter resultados correspondentes nos processos pontuais em estudo. Ferrari e Garcia (1997) aplicaram o método da projeção para redes com perdas.

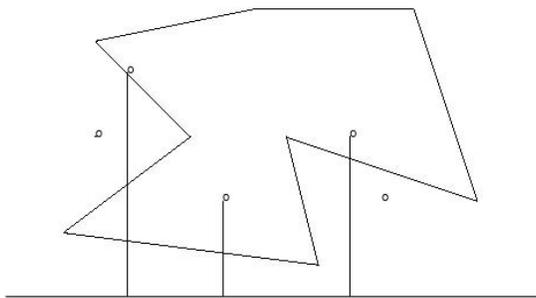


Figura 2.1: Projeção de um processo de Poisson.

2.2.1 Projeções através de conjuntos aleatórios

Seja N um processo de Poisson em \mathbb{R}^{d+1} com medida média μ . Seja Γ um conjunto aleatório de \mathbb{R}^{d+1} (em geral não independente de N). Um processo pontual N_Γ pode ser definido em \mathbb{R}^d por $N_\Gamma(B) = N(\Gamma \cap B \times \mathbb{R})$.

Por exemplo, se o conjunto Γ for independente de N condicionado em Γ , N_Γ é um processo de Poisson com medida média $\mu(\Gamma \cap B \times \mathbb{R})$ e a medida média $\mu_\Gamma(B) \equiv \mathbb{E}[N_\Gamma(B)]$ para N_Γ é

$$\mu_\Gamma(B) = \mathbb{E}[\mu(\Gamma \cap B \times \mathbb{R})]. \quad (2.3)$$

A igualdade em (2.3) nem sempre é verdadeira. Contudo há uma classe de conjuntos para os quais esta identidade é satisfeita, a classe de *Conjuntos de Parada*.

Definição 2.1 *Seja $\{\mathcal{F}_A\}$ uma família crescente de σ -álgebras indexadas por Borelianos, $A \in \mathcal{B}(E)$ (se $A \subset B$ então $\mathcal{F}_A \subset \mathcal{F}_B$). Então, um conjunto aleatório $\Gamma \subset E$ é **conjunto de parada** com respeito a $\{\mathcal{F}_A\}$ se $\{\Gamma \subset A\} \in \mathcal{F}_A$ para todo $A \in \mathcal{C}(E)$, onde $\mathcal{C}(E)$ é classe de conjuntos fechados.*

Teorema 2.1 *Seja N um processo de Poisson em \mathbb{R}^{d+1} com medida média μ com respeito a $\{\mathcal{F}_A\}$ e N_Γ é um processo pontual em \mathbb{R}^d obtido projetando-se os pontos de N que estão em Γ em \mathbb{R}^d . Se Γ é um conjunto de parada, então a medida média para N_Γ satisfaz*

$$\mathbb{E}[N_\Gamma(B)] = \mu_\Gamma(B) = \mathbb{E}(\mu(\Gamma \cap B \times \mathbb{R})). \quad (2.4)$$

Garcia (1995a) estudou uma caracterização sobre quais processos podem ser obtidos pelo método da projeção e verificou, em particular, que o processo de nascimento puro pode ser obtido através deste método, como será introduzido no próximo capítulo.

2.3 Verossimilhança para processos pontuais

A função de verossimilhança para processos pontuais será especificada nesta seção através da densidade (derivada de Radon Nikodym) com respeito ao processo de Poisson unitário. Seja P a lei de um processo pontual absolutamente contínuo com respeito a Q , a lei do processo de Poisson unitário, então a distribuição do processo pontual é caracterizada pela derivada de Radon-Nikodym (ou medida de Gibbs) e dada por:

$$dP(n) = L(n)dQ(n) \quad (2.5)$$

onde $L(n)$ é a função verossimilhança do processo da lei P com respeito a Q .

A Equação (2.5) pode ser reescrita como:

$$L(n) = \frac{dP}{dQ}(n). \quad (2.6)$$

Se o processo observado for denotado por $n = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$, a verossimilhança do processo é $L((X_1, Y_1), \dots, (X_N, Y_N))$ e se as observações Y_1, \dots, Y_N são faltantes, então define-se

$$L(X_1, \dots, X_N) = \mathbb{E}^Q[L((X_1, Y_1), \dots, (X_N, Y_N)) | X_1, \dots, X_N]. \quad (2.7)$$

A Equação (2.7) pode ser obtida através da derivada de Radon-Nikodym.

Derivadas de Radon-Nikodym

Considere um espaço de probabilidade (Ω, \mathcal{F}, Q) e as variáveis aleatórias X e Y com valores nos espaços mensuráveis (S_1, \mathcal{B}_1) e (S_2, \mathcal{B}_2) respectivamente. Suponha que f seja uma função mensurável em $S_1 \times S_2$ satisfazendo $\mathbb{E}^Q[f(X, Y)] = 1$ e defina $dP = f(X, Y)dQ$, ou seja, P é absolutamente contínua com respeito a Q e $f(X, Y)$ é a derivada de Radon-Nikodym. Então

$$\mathbb{E}^P[h(X)] = \mathbb{E}^Q[h(X)f(X, Y)] = \mathbb{E}^Q[h(X)\mathbb{E}^Q[f(X, Y)|X]]. \quad (2.8)$$

Sejam P^X e Q^X as restrições de P e Q em $\sigma(X)$. Então (2.8) implica que P^X é absolutamente contínua com respeito a Q^X (notação: $P^X \ll Q^X$) com a derivada de Radon-Nikodym dada por:

$$g(X) = \mathbb{E}^Q[f(X, Y)|X] = \int f(X, y)\nu(dy, X) \quad (2.9)$$

onde $\nu(B, x)$ é a distribuição condicional de Y dado $X = x$.

Para mais detalhes sobre verossimilhança para processos pontuais obtida através da derivada de Radon-Nikodym veja Cox e Isham (1980), Isham (1981) Ripley e Kelly (1977).

2.4 Dados faltantes

Em processos pontuais há um problema muito comum que é a não observação de determinada variável. Por exemplo, suponha que deseja-se estudar a ocorrência de determinada doença, muitas vezes se conhece a localização da contaminação, mas não se sabe quando foi contaminado. Diz-se que o tempo do processo é um dado faltante.

Dados faltantes (Little and Rubin 2002) aparecem naturalmente (dados que deveriam ser observados mas não foram) ou são intencionalmente escolhidos (modelos que incluem variáveis aleatórias que não foram observadas, chamadas de variáveis latentes ou efeitos aleatórios). Nestes casos o modelo é especificado pelo conjunto de dados completo, (x, y) , e $f_{\theta}(x, y)$ é a função de verossimilhança completa. Quando há problema de dados faltantes, dizemos que, por exemplo, y falta e x foi observado. O estimador de máxima verossimilhança maximiza a densidade marginal $f_{\theta}(x)$ que é chamada de verossimilhança dos dados observados. Esta densidade marginal é especificada implicitamente pelo modelo completo, $f_{\theta}(x) = \int f_{\theta}(x, y)dy$ e muitas vezes não se conhece sua forma fechada. Isto faz com que a inferência com dados faltantes seja difícil.

Um dos métodos mais utilizados para se trabalhar com problemas de dados faltantes é o algoritmo EM, introduzido por Dempster, Laird and Rubin (1977). É um método bastante utilizado, pois ele garante que o estimador de máxima verossimilhança será encontrado, se existir.

Métodos de Monte Carlo também são propostos para aproximar a verossimilhança dos dados observados em um modelo com dados faltantes, principalmente quando a forma fechada da verossimilhança completa não é conhecida.

2.4.1 Algoritmo EM

Um dos métodos mais utilizados para se trabalhar com dados faltantes é o algoritmo EM, introduzido por Dempster, Laird and Rubin (1977). O algoritmo EM é um procedimento de máxima verossimilhança que trabalha com a relação entre os parâmetros desconhecidos do modelo e com os dados faltantes. Considere que a variável X é observada e a variável Y é faltante, então o algoritmo realiza os seguintes passos:

1. **A esperança Passo E** Dado um conjunto de parâmetros estimados, o passo E calcula a esperança condicional da log-verossimilhança dos dados completos, dadas as variáveis observadas;

2. **A maximização Passo M** Dado o conjunto de dados completo, o passo M estima o parâmetro maximizando a log-verossimilhança completa do passo E.

É necessário então conhecer a função verossimilhança completa para poder calcular a seguinte esperança:

$$Q(\theta) = \mathbb{E}_Y[\log f(x, y|\theta)|x], \quad (2.10)$$

e, assim, tem-se $\theta_n = \arg \max_{\theta} Q(\theta)$. Em outras palavras, θ_{n+1} é o valor que maximiza (M) a esperança condicional (E) da log-verossimilhança completa, dado as variáveis observadas sob o valor anterior do parâmetro.

Os dois passos acima são executados até atingir a convergência.

2.4.2 Estimador de máxima verossimilhança via método de Monte Carlo

Métodos de Monte Carlo também tem sido bastante usados para encontrar o estimador de máxima verossimilhança quando há dados parcialmente observados, principalmente quando a forma fechada da verossimilhança não é conhecida.

Sung and Geyer (2007) sugerem simular dados faltantes independentemente dos dados observados, usando Monte Carlo ordinário (amostra independente). Ott (1979), Kong, Liu, and Wong (1994) propõem simular os dados faltantes dependendo dos dados observados também utilizando Monte Carlo ordinário, outros como Lange and Sobel (1991), Thompson and Guo (1991), Gelfand and Carlin (1993), Geyer (1994), Thompson (2003) sugerem simular os dados faltantes dependendo dos dados observados, utilizando Método de Monte Carlo baseados em cadeias de Markov (do inglês, MCMC).

Seja $f_{\theta}(x, y)$ a densidade conjunta com y faltante e x observado, então a constante de normalização para a distribuição condicional de y dado x é a verossimilhança $f_{\theta}(x)$. Por conveniência será usado a razão de verossimilhança com respeito à função de verossimilhança com o parâmetro fixo Ψ , então a log-verossimilhança é dada por

$$l(\theta) = \log \left\{ \frac{f_{\theta}(x)}{f_{\Psi}(x)} \right\} = \log \left[E_{\Psi} \left\{ \frac{f_{\theta}(X, Y)}{f_{\Psi}(X, Y)} \mid X = x \right\} \right]. \quad (2.11)$$

A aproximação natural de Monte Carlo da log-verossimilhança em (2.11)

$$\log \left\{ \frac{1}{n} \sum_{i=1}^n \frac{f_{\theta}(x, Y_i)}{f_{\Psi}(x, Y_i)} \right\} \quad (2.12)$$

onde Y_1, Y_2, \dots são realizações da distribuição condicional de Y dado $X = x$, geralmente simulado usando o algoritmo Metropolis-Hastings quando a constante de normalização é desconhecida.

Capítulo 3

Processo de nascimento puro

Um processo de nascimento puro é um processo de Markov a tempo contínuo onde os estados representam o tamanho e a localização da população em um determinado tempo e espaço, e as transições de um estado para o outro são chamadas de nascimentos. No caso do nascimento puro há apenas a transição do estado l para o estado $l + 1$. Um processo de nascimento puro tem muitas aplicações em demografia, epidemiologia, geologia, biologia.

A seguir será apresentado um exemplo de processo de nascimento puro no caso *não espacial*.

Considere células reproduzindo de acordo com as seguintes regras:

- (i) Uma célula no tempo t tem probabilidade de replicar $\lambda h + o(h)$ no tempo $(t, t + h)$;
- (ii) Eventos em células diferentes são independentes.

Se l transições são feitas no intervalo $(0, t)$, diz-se que o processo está no estado E_l . Em um processo de nascimento puro não espacial ocorrem as seguintes mudanças: $E_l \rightarrow E_{l+1} \rightarrow E_{l+2} \dots$

Considere agora uma plantação de soja onde foi detectada ocorrência da ferrugem asiática, pode-se dizer que o processo de contaminação para outras plantas é um processo de nascimento puro espaço-temporal. Denote este processo por \tilde{N} , na região $K \subseteq \mathbb{R}^d$. Este processo é especificado em termos da função não-negativa $\lambda : \mathbb{R}^d \times \mathcal{N}(K) \rightarrow [0, \infty)$. A função λ significa que se a configuração de pontos no tempo t é $n \in \mathcal{N}(K)$, então a probabilidade de que um ponto seja adicionado à configuração numa vizinhança do ponto x , que tem área ΔA , no próximo intervalo de tamanho Δt , é aproximadamente $\lambda(x, n)\Delta A\Delta t$.

3.1 Método da projeção para processos de nascimento puro

Um processo de nascimento puro é facilmente obtido através do método da projeção que foi introduzido por Kurtz (1989) e estudado por Garcia (1995a). Como descrito no Capítulo 2, a idéia

do método é a construção de processos pontuais em \mathbb{R}^d através de processo de Poisson de dimensão maior, construindo um subconjunto aleatório de espaço dimensional maior e projetando os pontos do processo de Poisson num subconjunto de dimensão menor.

Seja N um processo de Poisson em $K \times [0, \infty)$ com medida média de Lebesgue. É necessário construir uma família de conjuntos aleatórios Γ_t e assim o processo pontual \tilde{N}_t , tal que:

$$\begin{cases} \tau(t, x) = \int_0^t \lambda(x, \tilde{N}_s) ds \\ \tau(0, x) = 0 \\ \tilde{N}_t(B) = N(\Gamma_t \cap B \times [0, \infty)) \\ \Gamma_t = \{(x, y); x \in R^d, 0 \leq y \leq \tau(t, x)\}, \end{cases} \quad (3.1)$$

onde Γ_t é um conjunto de parada com respeito à filtragem $\mathcal{F}_A = \sigma\{N(B); B \subset A\}$.

Garcia (1995a) mostra que o Sistema (3.1) é um procedimento simples para simular processos de nascimento puro, se K é limitado este sistema contém um número finito de equações diferenciais e existe uma única solução para o sistema. O processo projetado \tilde{N}_t é o processo de nascimento puro espacial com intensidade espaço-temporal dada por $\lambda(x, \tilde{N}_t)$, e mais,

$$\tilde{N}_t(B) - \int_0^t \int_B \lambda(x, \tilde{N}_s) dx ds \quad (3.2)$$

é um \mathcal{F}_{Γ_t} -martingal, ou seja, \tilde{N}_t pode ser obtido como mudança no tempo aleatório.

A Figura 3.1 mostra a projeção de um processo de Poisson num subespaço menor.

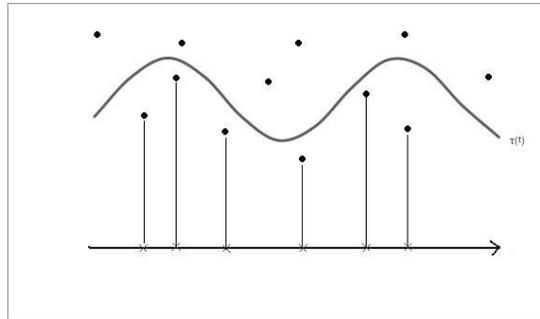


Figura 3.1: Projeção de um processo de Poisson.

Seja $\tilde{N}(B \times [0, t]) = \tilde{N}_t(B)$. A distribuição de \tilde{N} restrito a $K \times [0, T]$ é absolutamente contínua em relação à distribuição do processo de Poisson em $K \times [0, T]$ com a medida média de Lebesgue. A verossimilhança foi obtida por Kurtz (1989) e é dada por

$$L_T(n) = \exp \left\{ \int_{K \times [0, T]} \log \lambda(x, n_{s-}) n(dx \times ds) - \int_{K \times [0, T]} (\lambda(x, n_s) - 1) dx ds \right\}. \quad (3.3)$$

Por exemplo, suponha que $\lambda(x, n_s) = \exp \left\{ - \int \rho(\beta, x - y) n_s(dy) \right\}$ com $\rho'(\beta, x)$ limitada inferiormente, seja completamente conhecida. Sejam (x_i, y_i) , $i = 1, \dots, \tilde{N}_T(K)$ os pontos observados de \tilde{N} em $K \times [0, T]$. Neste caso, L_T pode ser escrita como:

$$L_T(n) = \exp \left\{ - \sum_{i < j} \rho(\beta, x_i - x_j) - \int_{[0, T]} \int_K (\exp \left\{ - \sum_i \rho(\beta, x - x_i) \mathbf{I}_{\{y_i < s\}} \right\} - 1) dx ds \right\} \quad (3.4)$$

e encontrar o valor de β que maximiza $L_T(n)$ é simples.

3.2 Verossimilhança para um processo observado após um período de tempo fixo

Este capítulo será dedicado a processos em que os nascimentos são observados somente após o instante T fixo, observando-se, desta maneira, apenas as localizações dos pontos, mas não os tempos de nascimento. Para estimação do parâmetro temos um problema de dados faltantes. O algoritmo EM (Dempster et al. 1977) não pode ser utilizado neste caso pois requer expressão fechada para a esperança condicional no passo E. Monte Carlo EM (Wei and Tanner, 1990; Guo and Thompson, 1991) pode ser utilizado, porém ele não fornece uma estimativa para o erro. Como a verossimilhança para os dados completos é conhecida, máxima verossimilhança por Monte Carlo pode ser usada (Thompson and Guo, 1991) e o erro Monte Carlo pode ser estimado (Geyer, 1994).

A verossimilhança completa com intensidade igual a $\lambda(x, n_s) = \exp \left\{ - \int \rho(\beta, x - y) n_s(dy) \right\}$ pode ser escrita como:

$$L_T(n) = \exp \left\{ - \sum_{i < j} \rho(\beta, x_i - x_j) - \int_{[0, T]} \int_K (\exp \left\{ - \sum_i \rho(\beta, x - x_i) \mathbf{I}_{\{y_i < s\}} \right\} - 1) dx ds \right\}. \quad (3.5)$$

Neste caso o processo de Poisson com medida média de Lebesgue em $K \times [0, T]$ tem a propriedade de que sua projeção em K é um processo de Poisson com medida média $Tm(K)$ (onde $m(K)$ é a medida de Lebesgue). Assim, seja N uma variável aleatória com distribuição de Poisson com parâmetro $m(K) \times T$, então dado $N = n$, $(X_1, Y_1), \dots, (X_n, Y_n)$ são independentes, identicamente distribuídos (i.i.d.) $U(K \times [0, T])$, portanto, dados $N = n$ e \mathbf{X} , sob \mathbb{Q} , as variáveis Y_1, \dots, Y_n são i.i.d. $U([0, T])$, então a verossimilhança “projetada” baseada em (2.9) é dada por

$$\hat{L}_T(\beta, \mathbf{x}) = \frac{1}{T^k} \int_{[0,T]^k} L_T(n) dy, \quad (3.6)$$

a qual, neste caso em particular, tem a forma:

$$\begin{aligned} \hat{L}_T(\beta, \mathbf{x}) &= \\ &= T^{-k} \int_0^T \cdots \int_0^T \exp\left\{-\sum_{i<j} \rho(\beta, x_i - x_j)\right. \\ &\quad \left.- \int_{[0,T]} \int_K (\exp\{-\sum_i \rho(\beta, x - x_i) \mathbf{I}_{\{y_i < s\}}\} - 1) dx ds\right\} dy_1 \dots dy_k. \end{aligned} \quad (3.7)$$

A Equação (3.7) é a função de verossimilhança para distribuição do processo \tilde{N}_T , fornecendo assim a base para estimação de máxima verossimilhança para β .

É possível observar que (3.7) pode ser reescrita como

$$L_T(\beta, \mathbf{x}) = L_T(\beta_0, \mathbf{x}) \left[\int_{[0,T]^k} \frac{L(\beta, \mathbf{x}, \mathbf{y})}{L(\beta_0, \mathbf{x}, \mathbf{y})} g(\mathbf{y}|\mathbf{x}, \beta_0) d\mathbf{y} \right] \left[\int_{[0,T]^k} g(\mathbf{y}|\mathbf{x}, \beta_0) d\mathbf{y} \right]^{-1}, \quad (3.8)$$

onde $g(\mathbf{y}|\mathbf{x}, \beta_0)$ é a distribuição condicional de \mathbf{Y} dados $\mathbf{X} = \mathbf{x}$ e $\beta = \beta_0$. Então se $\mathbf{Y}^{(j)}, j = 1, \dots, m$ são amostrados de uma cadeia de Markov ergódica com distribuição estacionária $g(\mathbf{y}|\mathbf{x}, \beta_0)$, uma aproximação de Monte Carlo para (3.8) é

$$L_m(\beta, \mathbf{x}) = L_T(\beta_0, \mathbf{x}) \frac{1}{m} \sum_{j=1}^m \frac{L(\beta, \mathbf{x}, \mathbf{Y}^{(j)})}{L(\beta_0, \mathbf{x}, \mathbf{Y}^{(j)})}. \quad (3.9)$$

Note que é necessário gerar uma amostra da distribuição

$$g(\mathbf{y}|\mathbf{x}, \beta_0) \propto \frac{L(\beta_0, \mathbf{x}, \mathbf{y}^{(j)})}{L(\beta_0, \mathbf{x})} \quad (3.10)$$

que é conhecido a menos de uma constante de normalização. O vetor \mathbf{Y} pode ser simulado através do algoritmo de Metropolis-Hastings (Metropolis *et al.*, 1953; Hastings, 1970).

A log-verossimilhança projetada é dada por:

$$l(\beta, \mathbf{x}) = \log L_T(\beta, \mathbf{x}) - \log L_T(\beta_0, \mathbf{x})$$

e sua aproximação Monte Carlo é

$$l_m(\beta, \mathbf{x}) = \log L_m(\beta, \mathbf{x}) - \log L_T(\beta_0, \mathbf{x}).$$

Denote por $\hat{\beta}_m = \arg \max l_m(\beta)$ e $\hat{\beta} = \arg \max l(\beta)$, o estimador de máxima verossimilhança para β .

Teorema 3.1 *Se o espaço paramétrico é compacto e $\hat{\beta}$ é único, então $\hat{\beta}_m \rightarrow \hat{\beta}$ quase certamente quando $m \rightarrow \infty$.*

Teorema 3.2 *Sob as condições do Teorema 3.1, existe uma constante positiva $\sigma^2(\hat{\beta})$ dependendo de $\hat{\beta}$ tal que*

$$\frac{\sqrt{m} (\hat{\beta}_m - \hat{\beta})}{\sigma(\hat{\beta})} \xrightarrow{D} N(0, 1),$$

quando $m \rightarrow \infty$.

As demonstrações dos Teoremas 3.1 e 3.2 são baseadas nos resultados de Geyer(1994) e podem ser encontradas em Garcia e Kurtz (1998).

Capítulo 4

Simulações

Este capítulo foi dedicado para as simulações das amostragens descritas no Capítulo 3. Todas as simulações foram feitas no *software* livre R (www.r-project.org). Para estimar o parâmetro do modelo através do método da máxima verossimilhança utilizou-se a função *optimize* a qual exige o pacote *stats* do R. *Optimize* encontra o máximo (ou o mínimo) da função desejada. Para diagnóstico do MCMC é necessário o pacote *CODA*.

4.1 Dados Completos

Esta seção será dedicada para simulação de um processo de nascimento puro utilizando o método da projeção como descrito na Seção 3.1 e o parâmetro será estimado pelo método da máxima verossimilhança.

Suponha um processo de nascimento puro com intensidade espaço-temporal dada por

$$\lambda(x, n) = \exp \left\{ \int (\beta^\alpha (|x - y| + C)^{-\alpha}) n(dy) \right\}, \quad (4.1)$$

com $\alpha > 0$. A intensidade foi escolhida desta forma pois assim pode-se abranger casos atrativos (repulsivo utilizando $-\beta$), ou seja, esta intensidade funciona muito bem para casos de epidemia, quando há ocorrência de uma doença contagiosa na população ela se espalha com rapidez para pessoas do mesmo convívio.

Neste caso, usando o método da projeção dado por (3.1), é muito fácil a simulação do processo de nascimento puro no conjunto compacto K . Primeiro gera-se $(\xi_1, \tau_1), \dots, (\xi_m, \tau_m)$ realizações de um processo de Poisson de taxa unitária em $K \times [0, S]$. As realizações do processo de nascimento puro $(x_1, y_1), \dots, (x_m, y_m)$ em $K \times [0, T]$, com $S \geq T$, são obtidas resolvendo o Sistema (3.1), interpretando cada ponto (ξ, τ) do processo de Poisson como uma localização individual em ξ que nasce no tempo

y satisfazendo $\tau(\xi, y) = \tau$. Neste caso, seja

$$\mathbf{C}_0 = \{(\xi_1, \tau_1), \dots, (\xi_m, \tau_m)\}, \quad (4.2)$$

e defina

$$y_1 = \min\{\tau; (\xi, \tau) \in \mathbf{C}_0 \text{ para algum } \xi\}; \quad (4.3)$$

$$y_1^* = \min\{\tau; (\xi, \tau) \in \mathbf{C}_0 \text{ para algum } \xi\}; \quad (4.4)$$

$$x_1 = \xi \text{ se } (\xi, y_1^*) \in \mathbf{C}_0; \quad (4.5)$$

$$\mathbf{C}_1 = \mathbf{C}_0 \setminus y_1^*. \quad (4.6)$$

Para $2 \leq k \leq m$ define-se recursivamente

$$y_k = \min\{(\tau - y_{k-1}^*) \exp\{-\beta^\alpha \sum_{j=1}^{k-1} (|\xi - x_j| + C)^{-\alpha} + y_{k-1}\}; (\xi, \tau) \in \mathbf{C}_{k-1}\} \quad (4.7)$$

$$y_k^* = (y_k - y_{k-1}) \exp\{\beta^\alpha \sum_{j=1}^{k-1} (|\xi - x_j| + C)^{-\alpha}\} + y_{k-1}^* \quad (4.8)$$

$$x_k = \xi \text{ se } (\xi, y_k^*) \in \mathbf{C}_{k-1}; \quad (4.9)$$

$$\mathbf{C}_k = \mathbf{C}_{k-1} \setminus y_k^*. \quad (4.10)$$

Este processo é repetido m vezes até obter $\mathbf{C}_m = \emptyset$. Se o tempo de nascimento do processo gerado for maior que T , ignora-se todos os nascimentos após T , senão o programa é executado novamente até obter tempo de nascimento menor ou igual a T .

As Figuras (4.1) à (4.5) representam os pontos de um processo de nascimento puro divididos em 5 tempos. Note que em cada gráfico apenas nascem novos pontos e não morre nenhum, desta maneira, na Figura (4.5), há todos os pontos que nasceram no intervalo de tempo de $[0, T]$. Também é possível notar uma tendência dos pontos se aglomerarem nas posições 3, 4 e 5. Este fato acontece pela função intensidade escolhida.

Seja a função intensidade dada por (4.1), portanto a equação da verossimilhança completa dada por (3.3) é da forma:

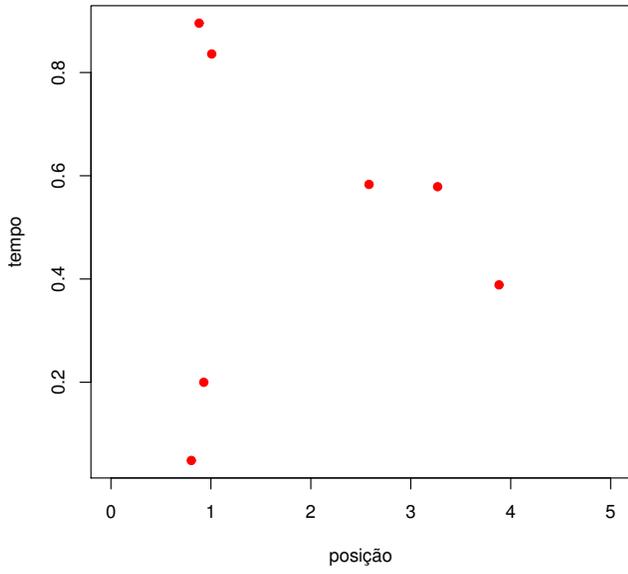


Figura 4.1: Nascimento até tempo 1.

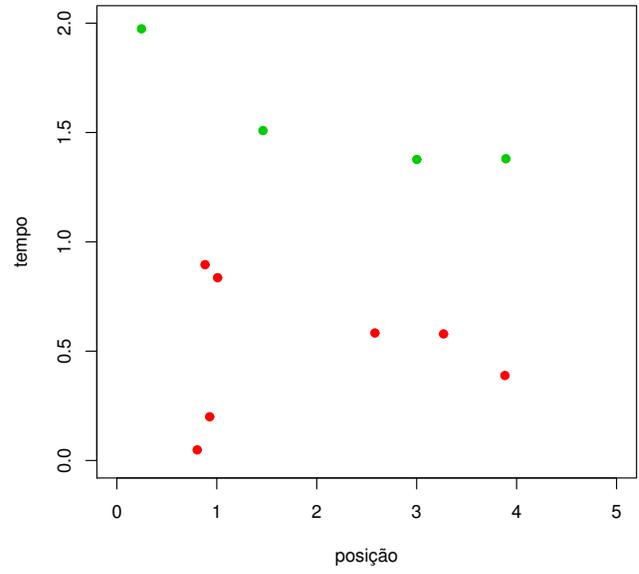


Figura 4.2: Nascimento até tempo 2.

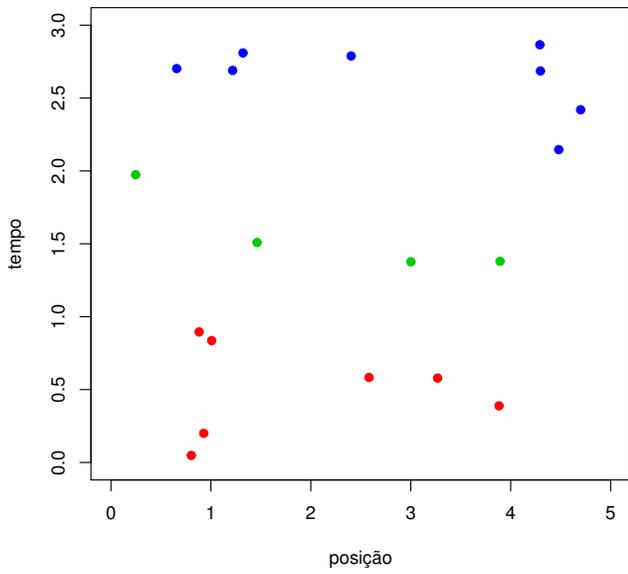


Figura 4.3: Nascimento até tempo 3.

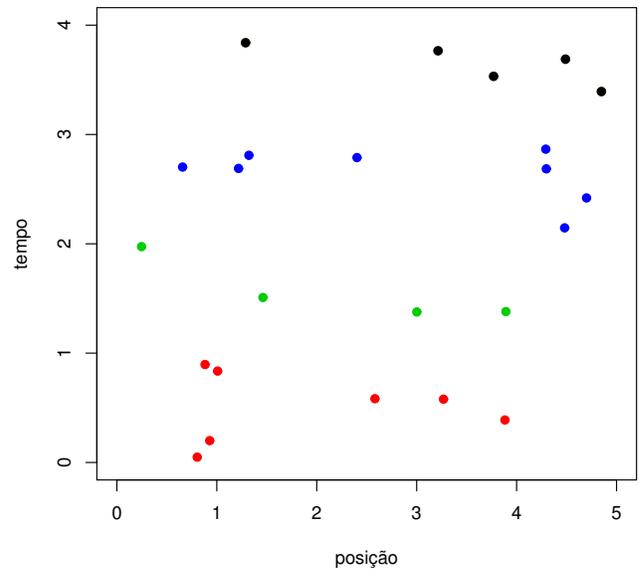


Figura 4.4: Nascimento até tempo 4.

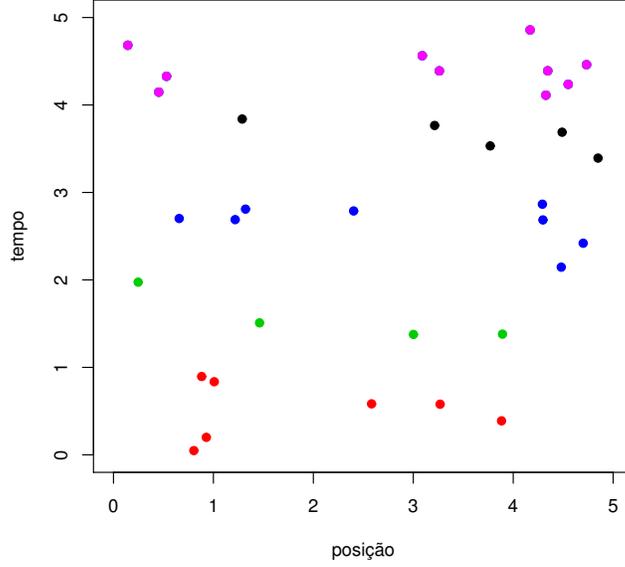


Figura 4.5: Nascimentos até tempo 5.

$$L_T(n) = \exp\left\{\beta^\alpha \sum_{i < j} (|x_{(i)} - x_{(j)}| + C)^{-\alpha}\right\} \quad (4.11)$$

$$- \int_{[0, T]} \int_K (\exp\{\beta^\alpha \sum_i (|x - x_{(i)}| + C)^{-\alpha} I_{\{y_{(i)} < s\}}\} - 1) dx ds \Big\}$$

$$= \exp\left\{\beta^\alpha \sum_{i < j} (|x_{(i)} - x_{(j)}| + C)^{-\alpha}\right\} \quad (4.12)$$

$$- \int_K \left(\sum_{l=1}^{m+1} (y_{(l)} - y_{(l-1)}) \exp\{\beta^\alpha \sum_{i=1}^{l-1} (|x - x_{(i)}| + C)^{-\alpha}\}\right) dx + \int_K \int_{[0, T]} 1 ds dx \Big\}$$

onde $y_{(1)} < y_{(2)} < \dots < y_{(m)}$ são os tempos de nascimentos ordenados e $x_{(1)}, x_{(2)}, \dots, x_{(m)}$ são as respectivas posições.

Para a simulação do processo de nascimento foi usado $\beta = 1, \alpha = 2$ e $C = 5, T = 5$ e $K = 5$.

Na Figura 4.6 pode-se ver o histograma das estimativas do parâmetro β e a Tabela 4.1 apresenta as estatísticas descritivas de $\hat{\beta}$ para os dados completos. Foram estimados 340 valores de β . Nota-se que a cauda do histograma é bastante pesada, no entanto a mediana é próxima do valor $\beta = 1$ e as estimativas estão em volta deste valor, concentrando 50% dos valores em torno de 0,84 e 1,06. O teste de normalidade Shapiro-Wilk (*shapiro.test* do *software* R) evidencia a rejeição da hipótese de

normalidade da estimativa do parâmetro. De acordo com Rathbun (1994) o estimador de máxima verossimilhança para processos pontuais espaço-temporais tem distribuição assintoticamente normal se o processo atender algumas condições de regularidade e $T \rightarrow \infty$. Neste projeto o valor de T é igual a 5, então não necessariamente atende as condições necessárias para que o estimador de máxima verossimilhança convirja para a distribuição normal.

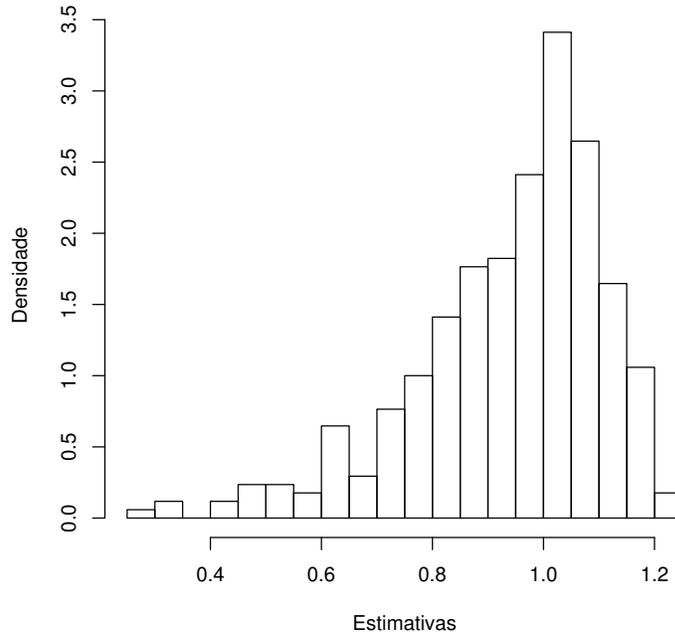


Figura 4.6: Histograma das estimativas de β para os dados completos.

Tabela 4.1: Análise descritiva da estimativa de β .

Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo	Variância
0,28	0,85	0,98	0,94	1,06	1,24	0,03

Uma amostra de nascimento puro com 34 pontos foi selecionada contendo a informação da posição e do tempo de nascimento (dados completos), a estimativa do parâmetro deste conjunto de dados completo é 0,81 e este valor será considerado para comparação com a aproximação Monte Carlo sugerida na Seção 3.2. As informações sobre o tempo foram retiradas e utilizou-se apenas as posições para execução da simulações da Seção 3.2(dados incompletos).

4.2 Processo após um tempo fixo T

Para a simulação no caso em que observa-se o processo até um tempo fixo T , como descrito na Seção 3.2 onde foi apresentado problema de dados faltantes, será utilizado algoritmo de geração de amostra, pois o tempo de nascimento não foi observado. Deseja-se gerar amostras do tempo de nascimento o qual tem distribuição condicionada em x e β proporcional à (3.10) e neste trabalho a densidade de interesse, (3.10), é conhecida à menos de uma constante, por esta razão utilizou-se o algoritmo Metropolis-Hastings (Metropolis *et al.*, 1953; Hastings, 1970).

4.2.1 Algoritmo Metropolis-Hastings

Para gerar amostras do tempo foi utilizado o algoritmo Metropolis-Hastings. O tempo de nascimento não foi observado, mas sabe-se que tem distribuição condicionada em x e β proporcional à (3.10), que é conhecida a menos de uma constante. O algoritmo Metropolis-Hastings é bastante utilizado neste caso e também quando a densidade é difícil, ou até mesmo, impossível de ser amostrada. A idéia é amostrar um valor de \mathbf{y} da densidade $q(\mathbf{x}/\mathbf{y})$ (chamada de *densidade proposta*). A escolha da proposta deve ser feita de maneira que a geração dos valores é possível ou mais fácil. O valor é aceito ou não com uma dada probabilidade. Este método de correção garante a convergência da cadeia para a distribuição de equilíbrio. Os passos deste método de amostragem são:

1. Inicialize o contador de iterações $t=0$ e especifique um valor inicial (neste projeto, um vetor inicial) $\mathbf{y}^{(0)}$.
2. Gere um novo valor, \mathbf{y}' da distribuição proposta $q(.|\mathbf{y})$.
3. Calcule a probabilidade de aceitação, $\alpha(\mathbf{y}, \mathbf{y}') = \min \left\{ 1, \frac{g(\mathbf{y}')q(\mathbf{y}/\mathbf{y}')}{g(\mathbf{y})q(\mathbf{y}'/\mathbf{y})} \right\}$, onde g é a distribuição de interesse dada por (3.10).
4. Gere $u \sim U(0, 1)$.
5. Se $u \leq \alpha$ então aceite o novo valor e faça $\mathbf{y}^{(t+1)} = \mathbf{y}'$, caso contrário rejeite e faça $\mathbf{y}^{(t+1)} = \mathbf{y}$.
6. Incremente o contador de t para $t + 1$ e volte para o passo 2

É necessário implementar o algoritmo até que a cadeia atinja convergência e a teoria de MCMC garante que a cadeia de Markov irá produzir uma amostra da distribuição de interesse se a cadeia rodar por um tempo suficientemente longo, o problema é determinar quão longo é suficiente para garantir a convergência das cadeias.

Geweke (1992), Gelman e Rubin (1992), Brooks e Gelman (1998) sugerem procedimentos para verificação da convergência, mas uma maneira simples e bastante eficiente é a análise de séries temporais de estatísticas derivadas da cadeia de Markov, como somas, médias, medianas etc. Considera-se que a cadeia convergiu quando a série das estatísticas estabiliza-se. Quando tem-se mais de uma cadeia independente (iniciadas com valores diferentes) pode-se verificar se as estatísticas convergem para o mesmo ponto de estabilidade.

4.2.2 Simulação com $\beta_0 = 0$

Na Equação (3.10) a escolha de β_0 é livre e em um processo de nascimento puro quando o parâmetro é igual a uma constante o processo pontual é um processo de Poisson, contudo optou-se primeiramente por escolher $\beta_0 = 0$, desta forma $\lambda(x, n)$ é igual a 1, assim os nascimentos são independentes e os tempos de nascimentos têm distribuição uniforme em $[0, T]$. O parâmetro β foi estimado usando a aproximação Monte Carlo, como descrito em (3.11) e $\hat{\beta}_m \rightarrow \hat{\beta}$ quase certamente, onde $\hat{\beta}$ é o estimador de máxima verossimilhança.

Pela Figura 4.7 nota-se que usando a aproximação Monte Carlo obtém-se estimativas em torno de $\hat{\beta} = 0,81$ e pelo teste de normalidade Shapiro-Wilk não há evidências para rejeitar a hipótese de normalidade (p -valor = 0,12). Note na Tabela (4.2) as estatísticas descritivas para as estimativas do parâmetro.

Tabela 4.2: Análise descritiva da estimativa de β para dados incompletos.

Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo	Variância
0,40	0,60	0,80	0,78	0,90	1,1	0,03

4.2.3 Simulação Metropolis-Hastings com $\beta_0 = 1$ e proposta uniforme

Para gerar tempos de nascimento escolheu-se um novo valor de β_0 na Equação (3.10), $\beta_0 = 1$. Com esta mudança o tempo de nascimento não tem mais distribuição uniforme, contudo utilizou-se o algoritmo Metropolis-Hastings para a geração dos tempos de nascimentos. Gerou-se 40 cadeias com proposta uniforme em $[0, T]$ de tamanho 10 mil, eliminou-se os 5 mil dados iniciais, como sugerem os gráficos de diagnósticos, representada nesta seção pela Figura 4.8 (para as outras cadeias os gráficos de diagnósticos estão no Apêndice B).

Neste projeto há 40 cadeias multivariadas, ou seja, em cada cadeia há 10 mil vetores com 34 pontos em cada vetor, calculou-se a mediana de cada vetor, obtendo desta maneira 10 mil medianas em cada cadeia e com esta estatística foi feito o gráfico *cumuplot* que constrói o gráfico da evolução

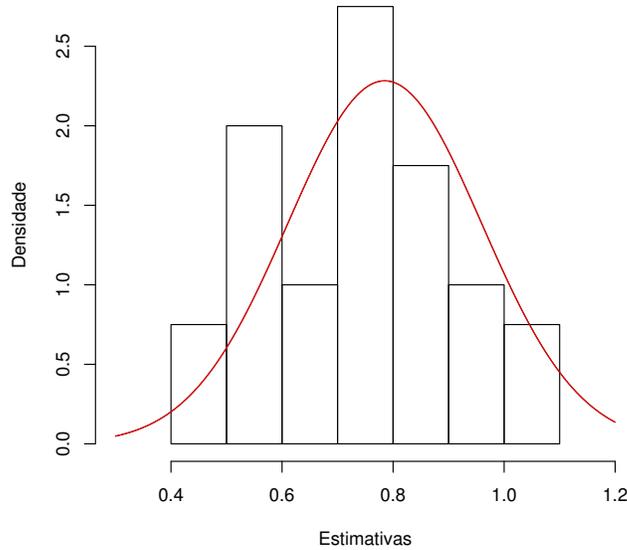


Figura 4.7: Histograma da estimativa dos β para dados incompletos.

dos quantis amostrais como função da iteração. A linha preta dos gráficos é o quantil 0,50 e as linhas tracejadas são os quantis 0,025 e 0,975 (eixo x é o número da iteração e o eixo y é o quantil amostral acumulado). Nota-se que em todos os gráficos o valor do quantil amostral como função da iteração vai para o mesmo valor três e próximo a iteração 5 mil o gráfico apresenta tendência constante. Após a análise de diagnóstico das séries os dados das cadeias foram selecionados de 20 em 20 de tal maneira a eliminar a dependência entre os elementos sorteados da cadeia, como pode ser notado nas figuras da função de autocorrelação, Figura 4.9 e Figura 4.10, ou seja, 250 dados foram utilizados para estimar o parâmetro do modelo. (As figuras de todas as cadeias podem ser vistas no Apêndice B.)

Foram obtidas 40 estimativas de β_m através do método de máxima verossimilhança utilizando a aproximação Monte Carlo, como descrito em (3.9). As estimativas podem ser vistas na Figura 4.11, nota-se a concentração dos valores em torno do $\hat{\beta} = 0,81$ e pelo teste da normalidade *Shapiro-Wilk* não há evidências para rejeitar a hipótese de normalidade dos dados (p-valor 0,31), mostrando desta forma a adequação do método proposto. Na Tabela (4.3) estão as estatísticas descritivas das estimativas, nota-se que a média e a mediana estão bem próximas ao valor 0,81 que é a estimativa dos dados completos.

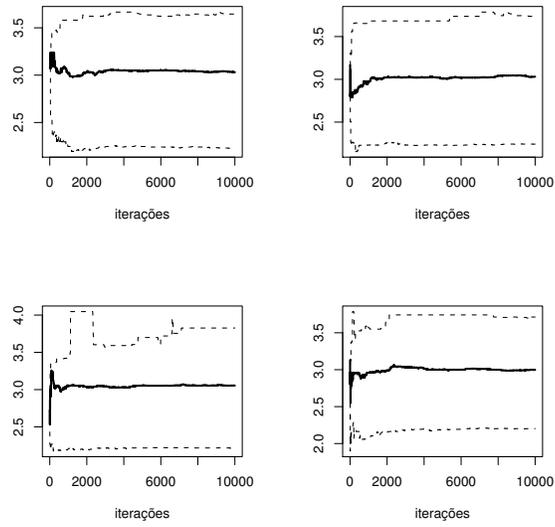


Figura 4.8: Diagnóstico para as cadeias 1 à 4 - proposta uniforme.

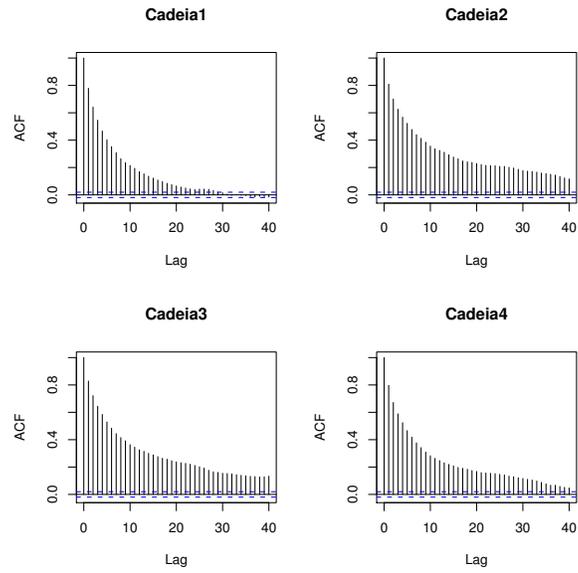


Figura 4.9: Acf para as cadeias 1 à 4 - proposta uniforme.

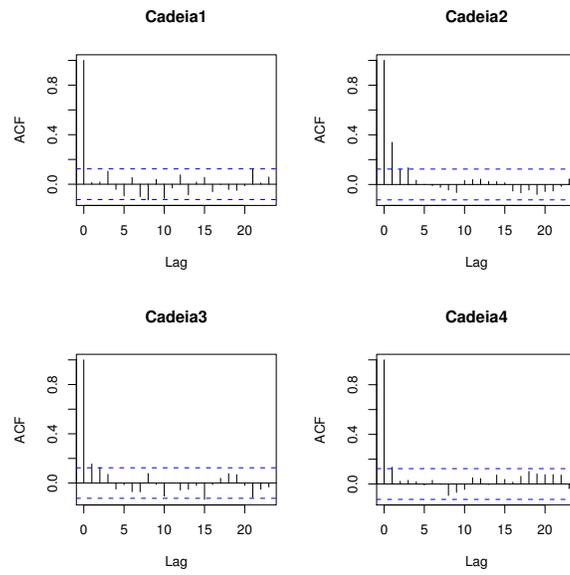


Figura 4.10: Acf para as cadeias 1 à 4 com 250 dados - proposta uniforme.

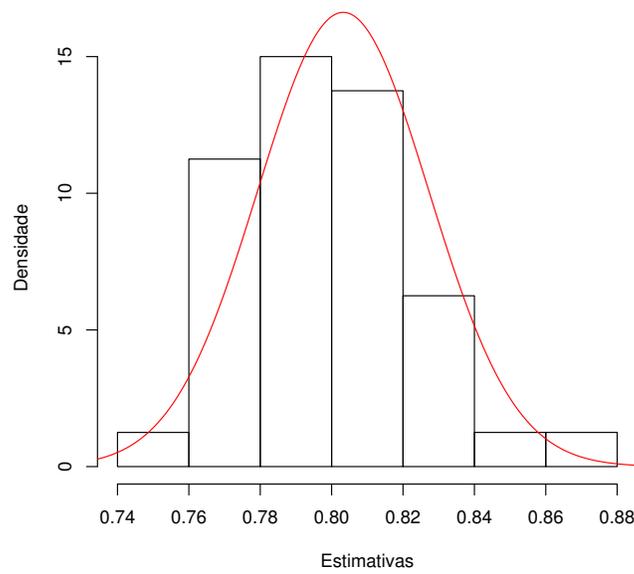


Figura 4.11: Histograma da estimativa dos β para dados incompletos - proposta uniforme.

Tabela 4.3: Análise descritiva da estimativa de β para dados incompletos - proposta uniforme.

Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo	Variância
0,75	0,79	0,80	0,80	0,81	0,83	0,0005

4.2.4 Simulação Metropolis-Hastings com $\beta = 1$ e proposta beta

Optou-se também por gerar amostras dos tempos de nascimento utilizando proposta $beta(a, b)$ assimétrica com $a = 0,77$ e $b = 0,64$ multiplicada por T . A assimetria da beta foi considerada pois os dados gerados de tal forma a abranger casos atrativos, assim, com a proposta $beta(a, b)$, pode-se atribuir maiores probabilidades para tempos maiores e foi multiplicada por T pois os tempos das ocorrências estão dispostos entre 0 e T e não entre 0 e 1. Do mesmo modo como feito com a proposta uniforme, foram geradas 40 cadeias de tamanho 10 mil e eliminou-se os 5 mil dados iniciais, como sugere a Figura 4.12. Os dados foram selecionados de 20 em 20 eliminando desta maneira a dependência da cadeia, como pode ser notado nas Figuras 4.13 e 4.14.

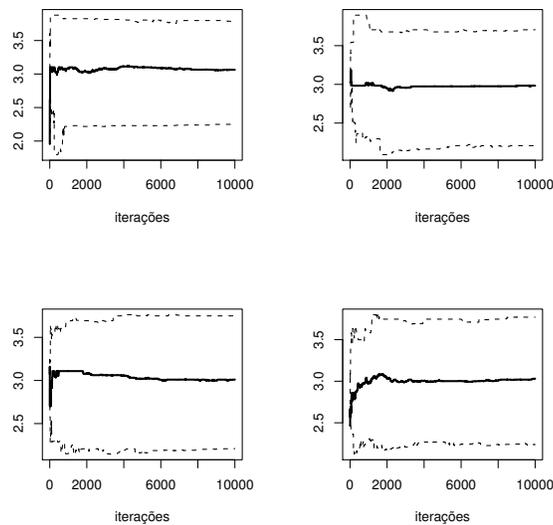


Figura 4.12: Diagnóstico para as cadeias 1 à 4 - proposta beta.

Foram obtidas 40 estimativas de β_m através do método de máxima verossimilhança utilizando a aproximação Monte Carlo, como descrito em (3.9). As estimativas podem ser vistas na Figura 4.15, nota-se a concentração dos valores em torno do $\hat{\beta} = 0,81$ e pelo teste da normalidade *Shapiro-Wilk* não há evidências para rejeitar a hipótese de normalidade dos dados (p -valor 0,27), novamente

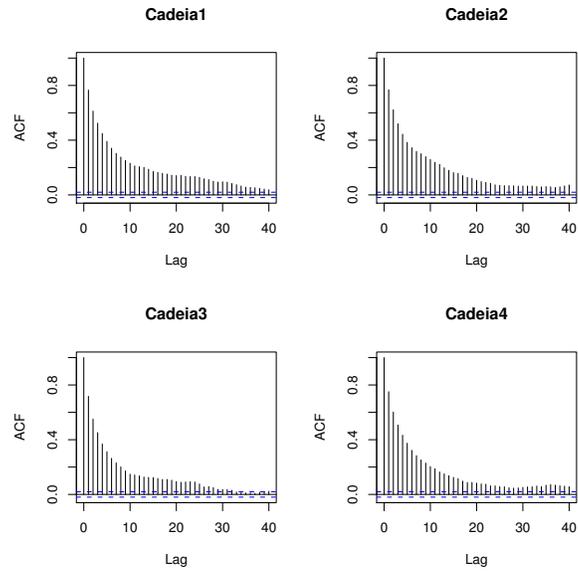


Figura 4.13: Acf para as cadeias 1 à 4 - proposta beta.

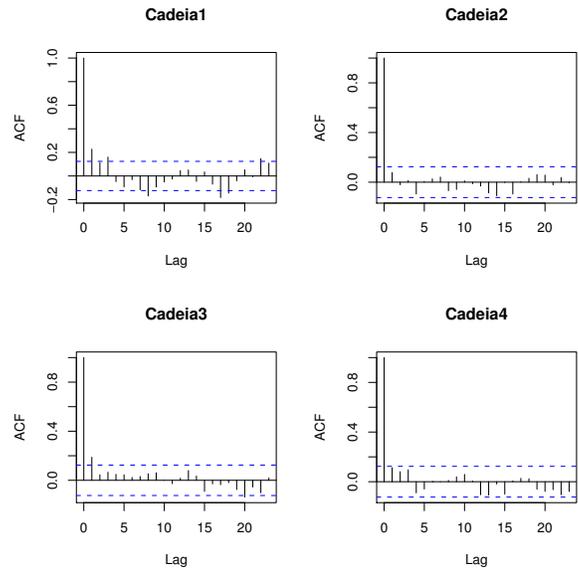


Figura 4.14: Acf para as cadeias 1 à 4 com 250 dados - proposta beta.

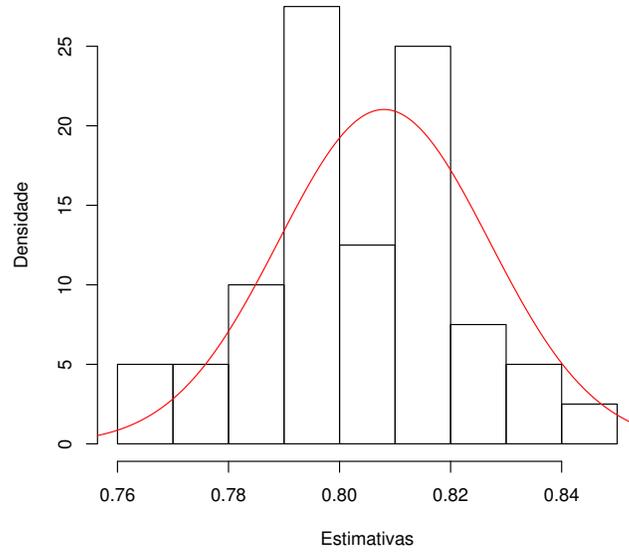


Figura 4.15: Histograma da estimativa dos β para dados incompletos - proposta beta.

mostrando a adequação do modelo proposto na teoria. Na Tabela (4.4) estão as estatísticas descritivas das estimativas, nota-se que a média e a mediana estão bem próximas ao valor 0,81 que é a estimativa dos dados completos.

Tabela 4.4: Análise descritiva da estimativa de β para dados incompletos - proposta beta.

Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo	Variância
0,76	0,80	0,81	0,80	0,82	0,85	0,00036

Capítulo 5

Conclusão

A proposta desta dissertação foi trabalhar com estimação paramétrica de máxima verossimilhança para processo de nascimento puro espacial com dados parcialmente observados e a utilização do método da projeção para encontrar a função de verossimilhança. Mostrou-se como obter a função de verossimilhança quando não são conhecidos os tempos de nascimento, ou seja, quando há problema de dados faltantes. Neste caso a função de verossimilhança pôde ser obtida como projeção da verossimilhança completa.

A possibilidade da escolha do valor de $\beta_0 = 0$ na simulação fornece maior facilidade na geração dos dados pois neste caso os tempos de nascimento tem distribuição uniforme. A simulação dos tempos de nascimento para o caso em que fixa $\beta_0 = 0$ mostrou-se adequada pois as estimativas utilizando a aproximação de Monte Carlo convergem para o valor estimado com os dados completos e tem distribuição assintoticamente normal.

Já no caso em que foi escolhido $\beta_0 = 1$ o método de simulação Metropolis-Hastings foi utilizado para gerar os dados não observados e este método mostrou-se muito eficiente para as duas propostas utilizadas: uniforme e beta. A aproximação Monte Carlo para encontrar o estimador de máxima verossimilhança é adequada para processo de nascimento puro espacial com dados parcialmente observados, pois as estimativas utilizando a aproximação ficou em torno do valor estimado quando tem-se os dados completos e verificou-se que estas estimativas são assintoticamente normais.

Trabalhos Futuros

Há diversas generalizações possíveis deste trabalho. Por exemplo, se fizermos uma pequena modificação no esquema de amostragem, observando o processo após um número k (fixo) de nascimentos, a expressão da verossimilhança projetada fica muito mais simples. Neste caso a verossimilhança completa do processo é da forma:

$$L_k(\mathbf{x}, \mathbf{y}) = \exp\left\{-\sum_{i < j} \rho(\beta, x_i - x_j) - \int_K \sum_{j=2}^k (y_{(j)} - y_{(j-1)}) (\exp\{-\sum_{\sigma_i \leq j-1} \rho(\beta, x - x_{\sigma_i})\} - 1) dx\right\} \quad (5.1)$$

onde σ é a permutação de $\{1, 2, \dots, k\}$ tal que $(x_{\sigma_1}, y_{(1)}), (x_{\sigma_2}, y_{(2)}), \dots, (x_{\sigma_k}, y_{(k)})$ seja a amostra observada onde $y_{(1)}, \dots, y_{(k)}$ são estatísticas de ordem de y_1, \dots, y_k . Sob o processo de Poisson unitário (Q) sabemos que $Y_{(1)}, Y_{(2)} - Y_{(1)}, \dots, Y_{(k)} - Y_{(k-1)}$ são variáveis aleatórias i.i.d. $\exp(m(K))$ e esta observação simplifica a expressão da verossimilhança projetada para:

$$\begin{aligned} \hat{L}_k(x_1, \dots, x_k) &= \mathbb{E}^Q[L_k(n)|\mathbf{x}] \\ &= \exp\left\{-\sum_{i < j} \rho(\beta, x_i - x_j)\right\} \left[\frac{1}{k!} \sum_{\sigma} \prod_{j=2}^k \frac{m(K)}{\int_K (\exp\{-\sum_{\sigma_i \leq j-1} \rho(\beta, x - x_{\sigma_i})\}) dx} \right] \quad (5.2) \end{aligned}$$

Também, neste trabalho nos concentramos somente em um parâmetro, a saber β . Entretanto, podemos também pensar na estimação paramétrica dos parâmetros α e C na expressão (3.8).

Além disso, seria muito interessante poder deixar que os dados nos deem a forma da taxa de nascimento. Isto é, não fazer nenhuma hipótese sobre a forma da função λ e/ou ρ e estimá-las de forma não paramétrica utilizando expansão em funções de base como splines ou wavelets.

Apêndice A

Processos pontuais univariados - Algumas definições

Definição A.1 Dado um espaço de probabilidade completo (Ω, Φ, P) , uma família $\{\Phi_t, t \geq 0\}$ de sub- σ -álgebras de Φ é dita ser uma **filtragem** se esta é crescente, isto é, Φ_t é σ -álgebra $\Phi_t \subseteq \Phi$, para todo $t \geq 0$ e $\Phi_s \subseteq \Phi_t$, para todo $s < t$.

Definição A.2 Seja (Ω, Φ, P) um espaço de probabilidade completo e $\{\Phi_t, t \geq 0\}$ uma filtragem. Uma variável aleatória $\tau : \Omega \rightarrow [0, \infty]$ é chamada **Tempo de Markov (ou tempo de parada)** relativo a filtragem $\{\Phi_t, t \geq 0\}$ se, para todo $t \geq 0$, tivermos

$$\{\omega : \tau(\omega) \leq t\} \in \Phi_t. \quad (\text{A.1})$$

Definição A.3 Seja (Ω, Φ, P) um espaço de probabilidade completo e $\{\Phi_t, t \geq 0\}$ uma filtragem contínua à direita (isto é, $\Omega_t = \bigcap_{s \geq t} \Omega_s$) tal que Ω_0 contém todos os conjuntos de probabilidade zero de Ω . Um **processo pontual simples** em $[0, \infty)$ é uma sequência $T = \{\tau_n, n \geq 1\}$ de tempos de Markov com relação a filtragem $\{\Phi_t, t \geq 0\}$ que satisfaz, com probabilidade 1, as seguintes propriedades:

1. $\tau_1 > 0$;
2. $\tau_n < \tau_{n+1}$, se $\tau_n < \infty$;
3. $\tau_n = \tau_{n+1}$, se $\tau_n = \infty$

Apêndice B

Gráficos

As Figuras B.1 à B.9 são os gráficos de diagnóstico de convergência das cadeias para proposta uniforme e as Figuras B.10 à B.18 são os gráficos de diagnóstico de convergência das cadeias para proposta beta.

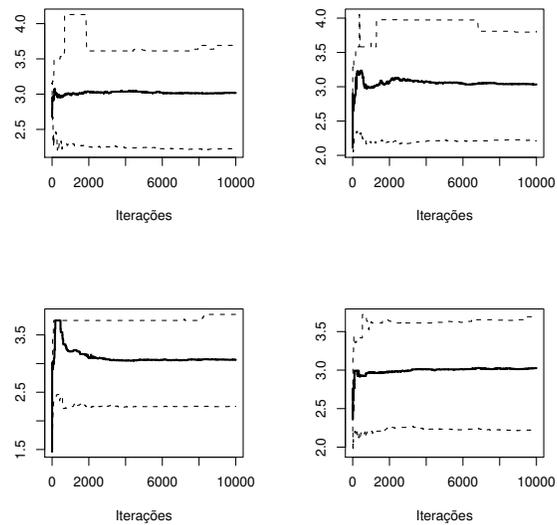


Figura B.1: Diagnóstico para as cadeias 5 à 8 - proposta uniforme.

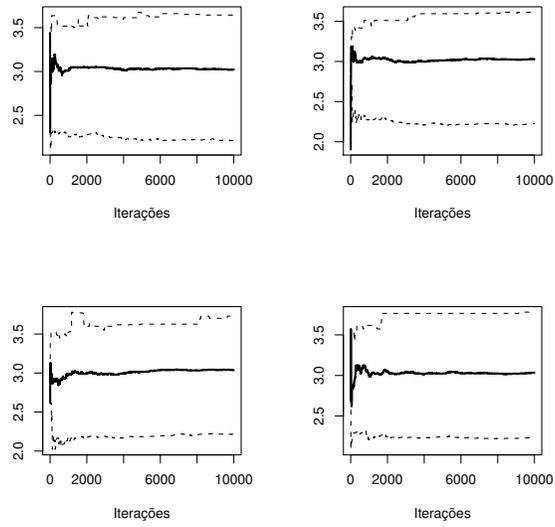


Figura B.2: Diagnóstico para as cadeias 9 à 12 - proposta uniforme.

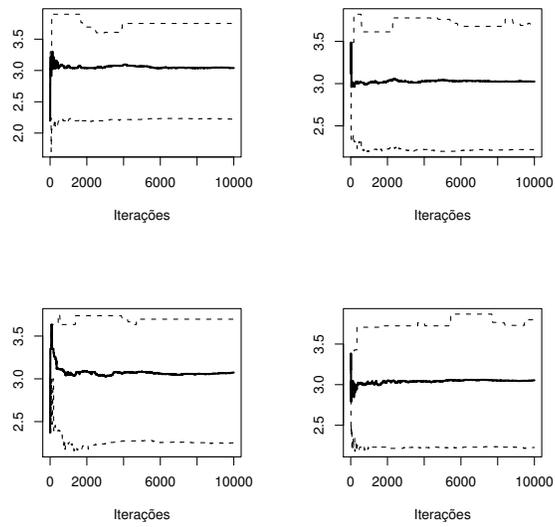


Figura B.3: Diagnóstico para as cadeias 13 à 16 - proposta uniforme.

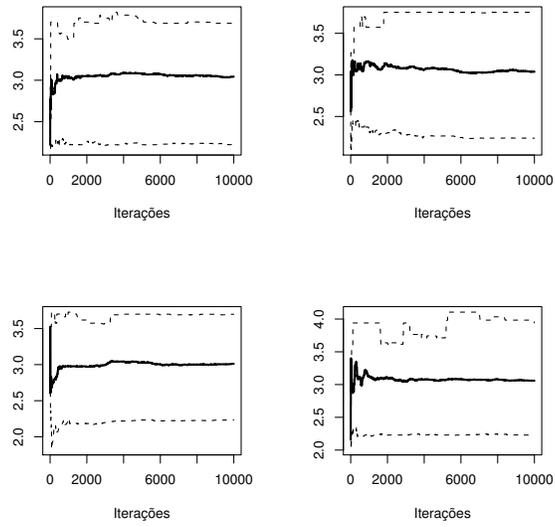


Figura B.4: Diagnóstico para as cadeias 17 à 20 - proposta uniforme.

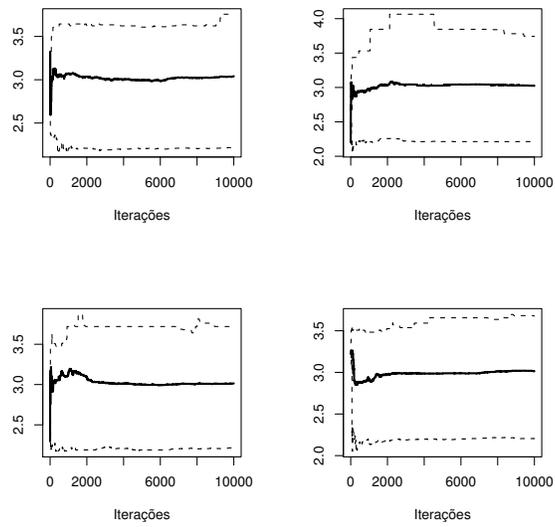


Figura B.5: Diagnóstico para as cadeias 21 à 24 - proposta uniforme.

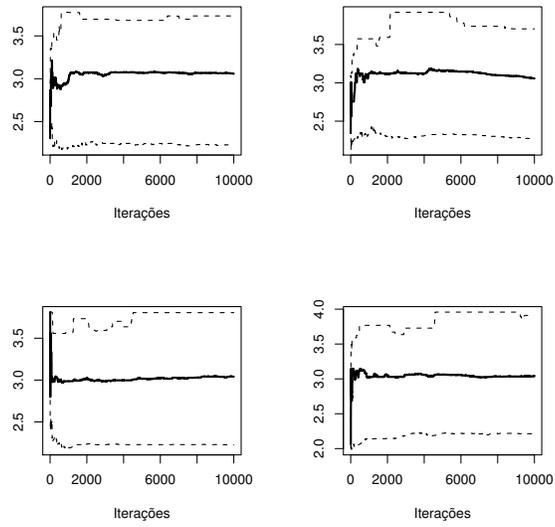


Figura B.6: Diagnóstico para as cadeias 25 à 28 - proposta uniforme.

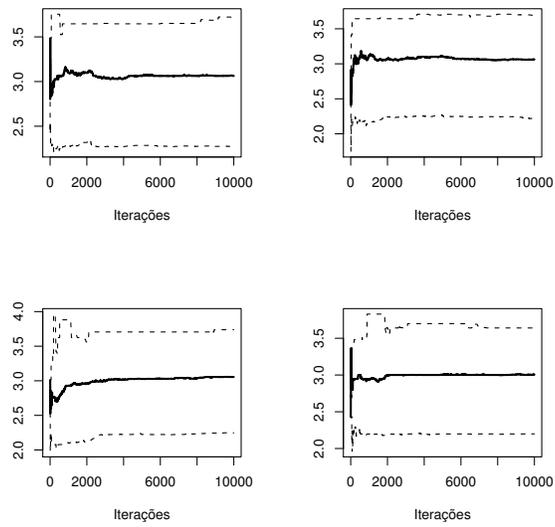


Figura B.7: Diagnóstico para as cadeias 29 à 32 - proposta uniforme.

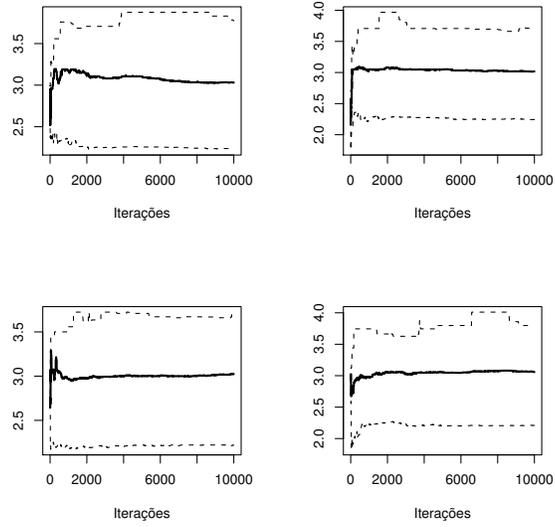


Figura B.8: Diagnóstico para as cadeias 33 à 36 - proposta uniforme.

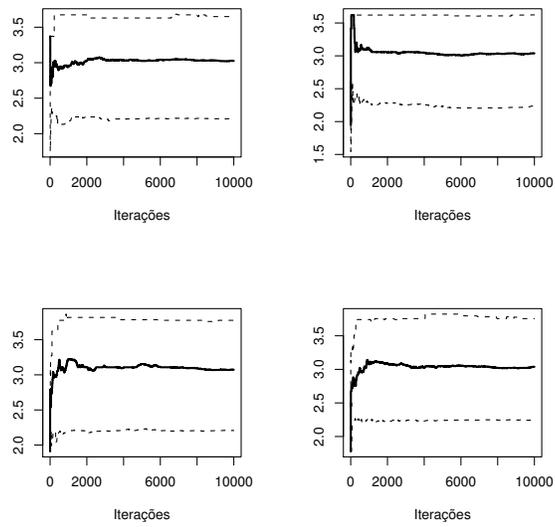


Figura B.9: Diagnóstico para as cadeias 37 à 40 - proposta uniforme.

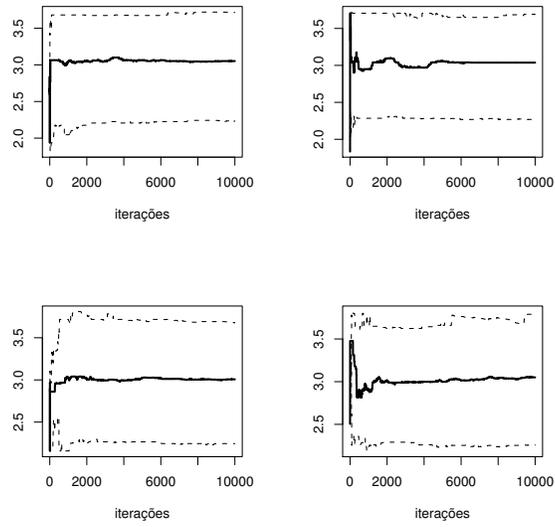


Figura B.10: Diagnóstico para as cadeias 5 à 8 - proposta beta.

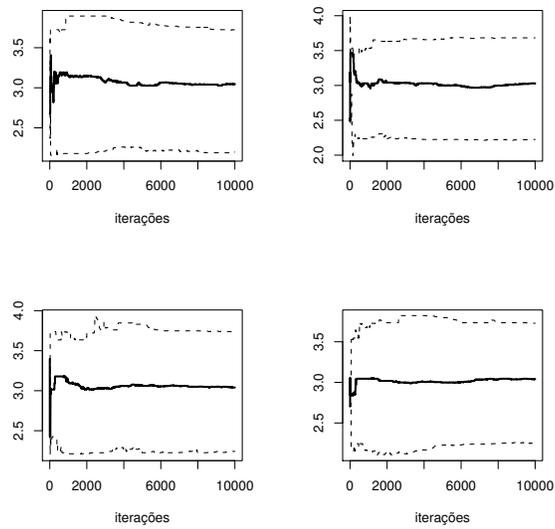


Figura B.11: Diagnóstico para as cadeias 9 à 12 - proposta beta.

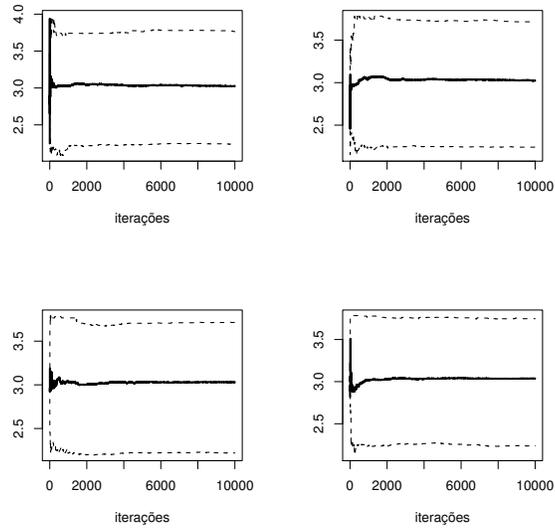


Figura B.12: Diagnóstico para as cadeias 13 à 16 - proposta beta.

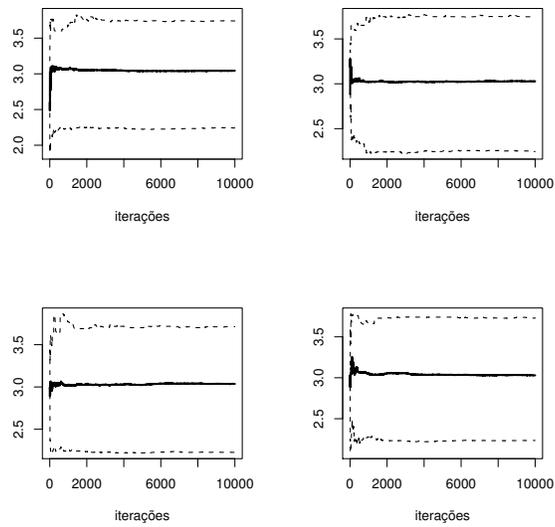


Figura B.13: Diagnóstico para as cadeias 17 à 20 - proposta beta.

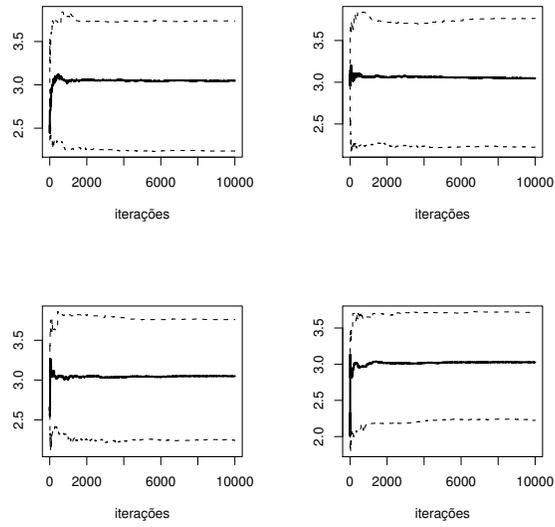


Figura B.14: Diagnóstico para as cadeias 21 à 24 - proposta beta.

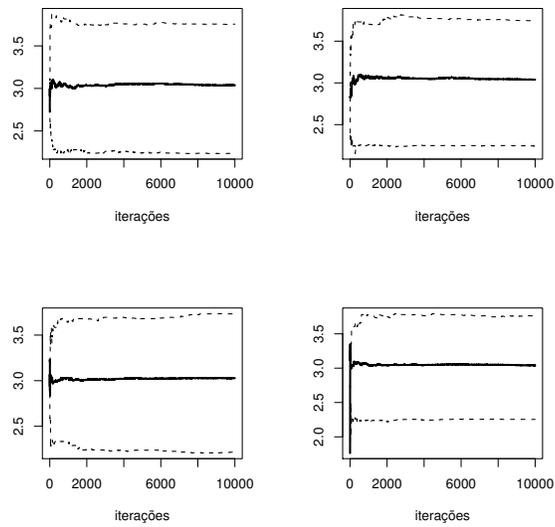


Figura B.15: Diagnóstico para as cadeias 25 à 28 - proposta beta.

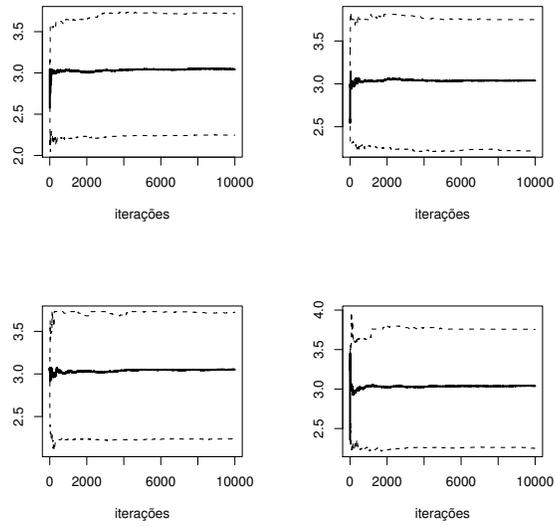


Figura B.16: Diagnóstico para as cadeias 29 à 32 - proposta beta.

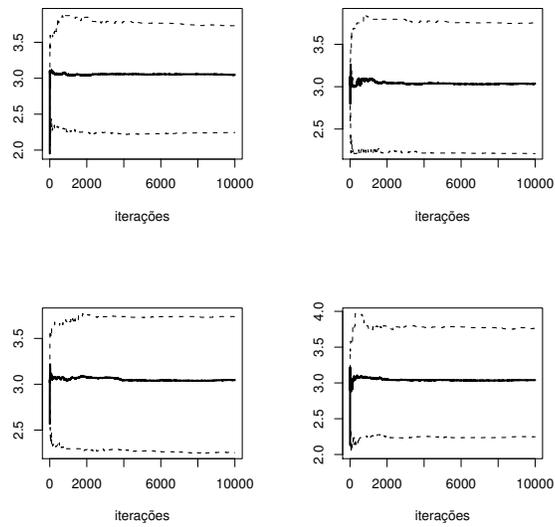


Figura B.17: Diagnóstico para as cadeias 33 à 36 - proposta beta.

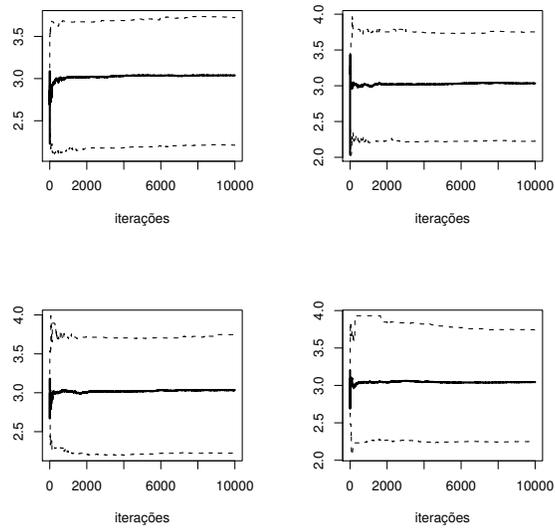


Figura B.18: Diagnóstico para as cadeias 37 à 40 - proposta beta.

Nas Figuras B.19 à B.27 podem ser vistas as funções de autorelação das cadeias geradas através do algoritmo de Metropolis-Hastings com proposta uniforme e nas Figuras B.28 à B.36 são os gráficos das funções de autocorrelação para cadeias com 250 (retirada da dependência).

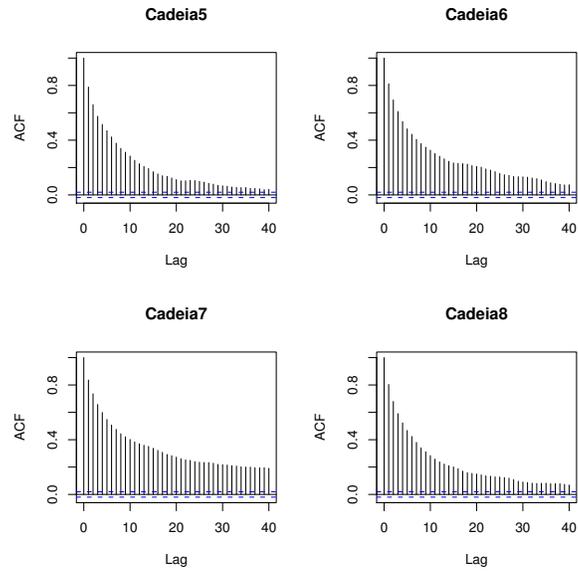


Figura B.19: Acf para as cadeias 5 à 8 - proposta uniforme.

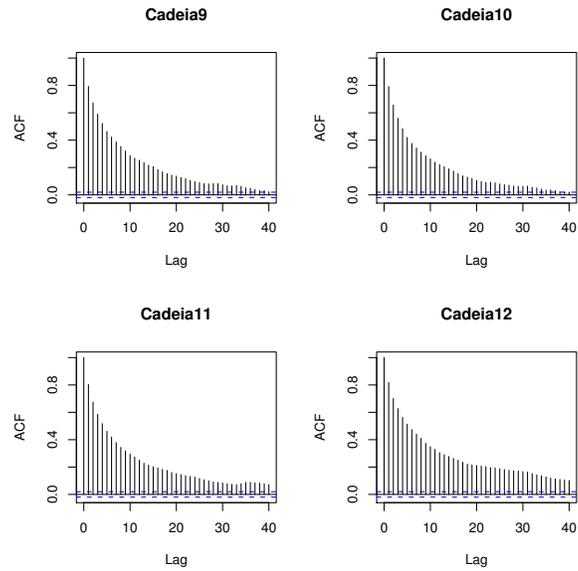


Figura B.20: Acf para as cadeias 9 à 12 - proposta uniforme.

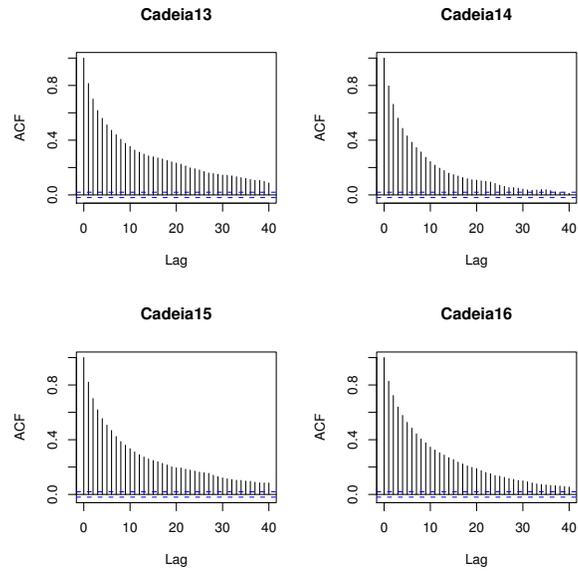


Figura B.21: Acf para as cadeias 13 à 16 - proposta uniforme.

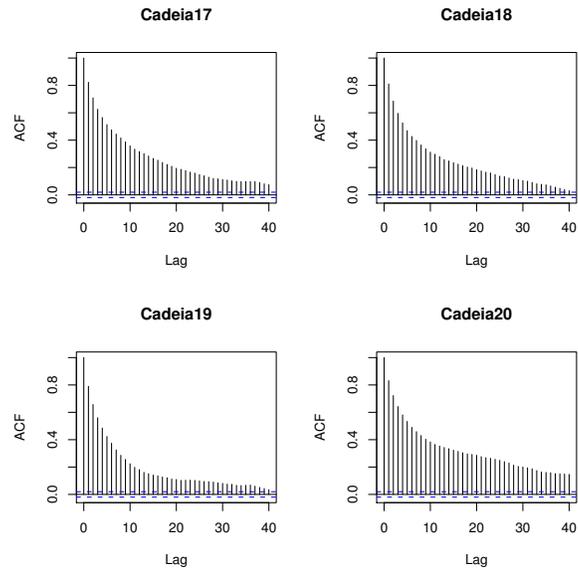


Figura B.22: Acf para as cadeias 17 à 20 - proposta uniforme.

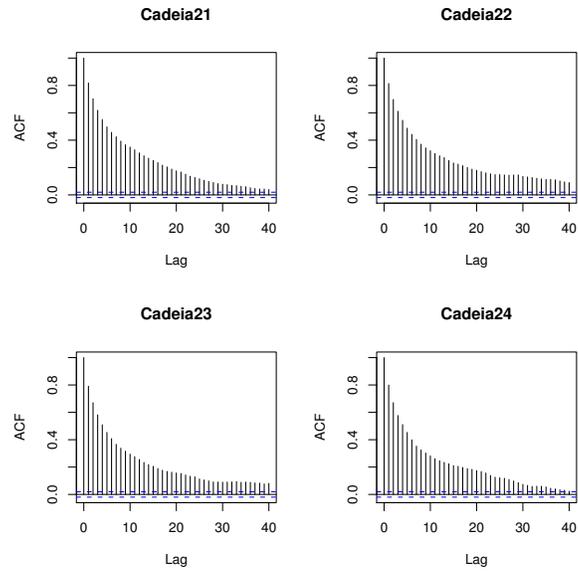


Figura B.23: Acf para as cadeias 21 à 24 - proposta uniforme.

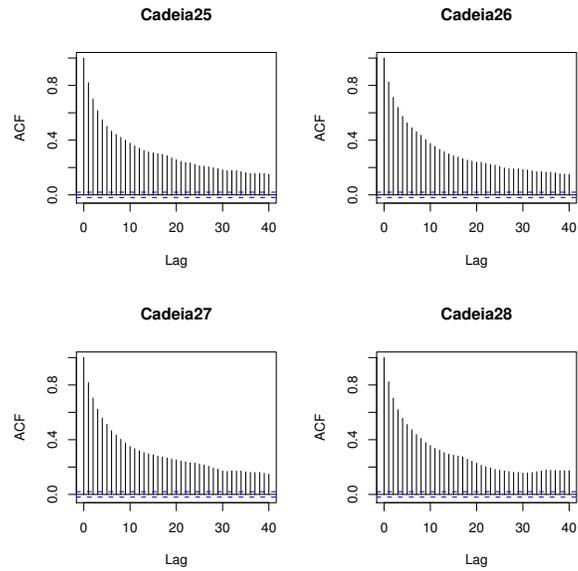


Figura B.24: Acf para as cadeias 25 à 28 - proposta uniforme.

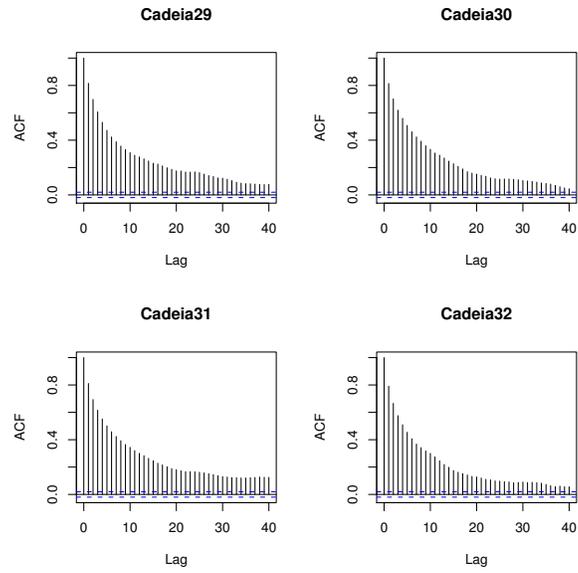


Figura B.25: Acf para as cadeias 29 à 32 - proposta uniforme.

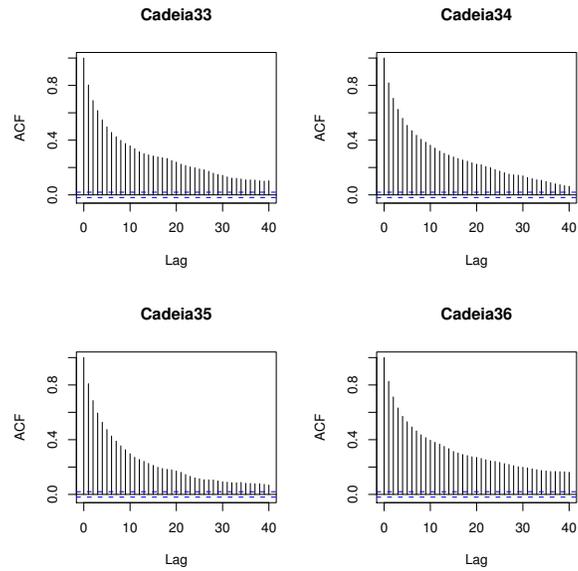


Figura B.26: Acf para as cadeias 33 à 36 - proposta uniforme.

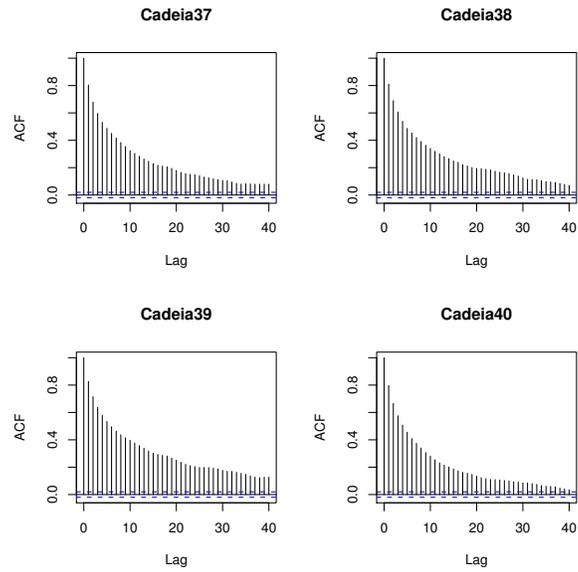


Figura B.27: Acf para as cadeias 37 à 40 - proposta uniforme.

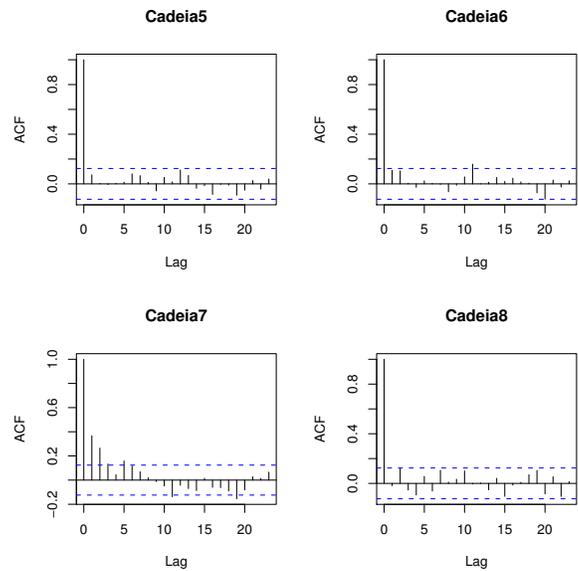


Figura B.28: Acf para as cadeias 5 à 8 com 250 dados- proposta uniforme.

Nas Figuras B.37 à B.45 podem ser vistas as funções de autorelação das cadeias geradas através do algoritmo de Metropolis-Hastings com proposta beta e nas Figuras B.46 à B.54 são os gráficos das funções de autocorrelação para cadeias com 250 (retirada da dependência).

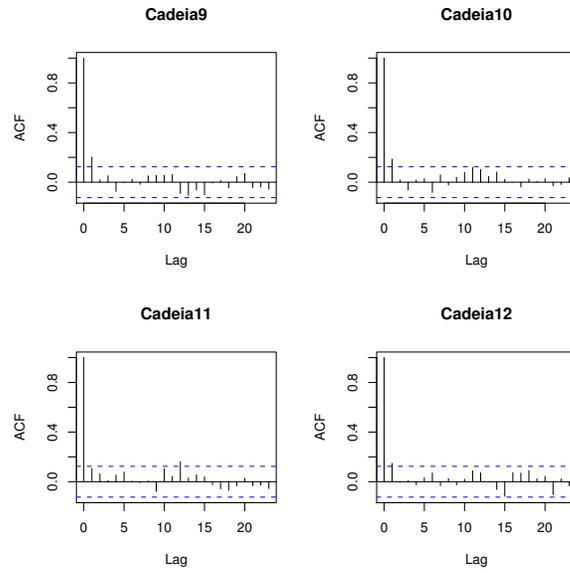


Figura B.29: Acf para as cadeias 9 à 12 com 250 dados- proposta uniforme.

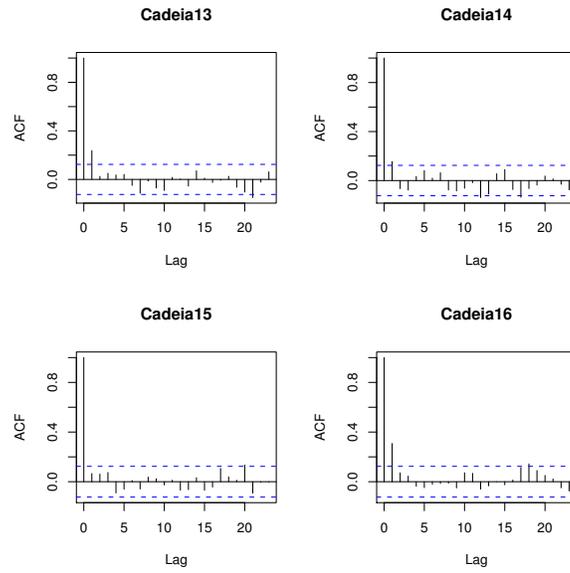


Figura B.30: Acf para as cadeias 13 à 16 com 250 dados- proposta uniforme.

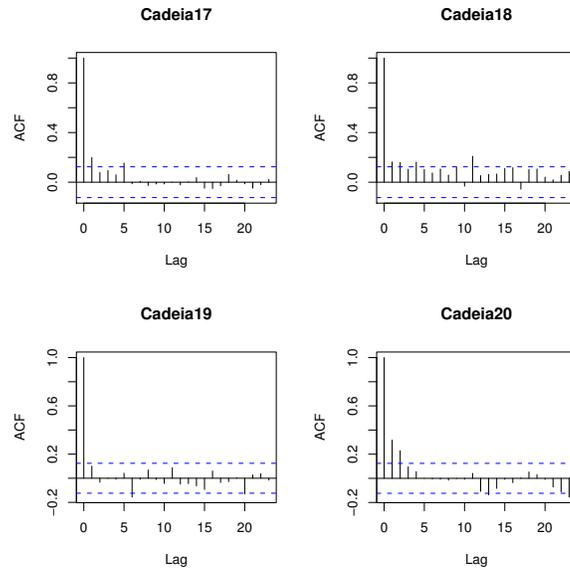


Figura B.31: Acf para as cadeias 17 à 20 com 250 dados- proposta uniforme.

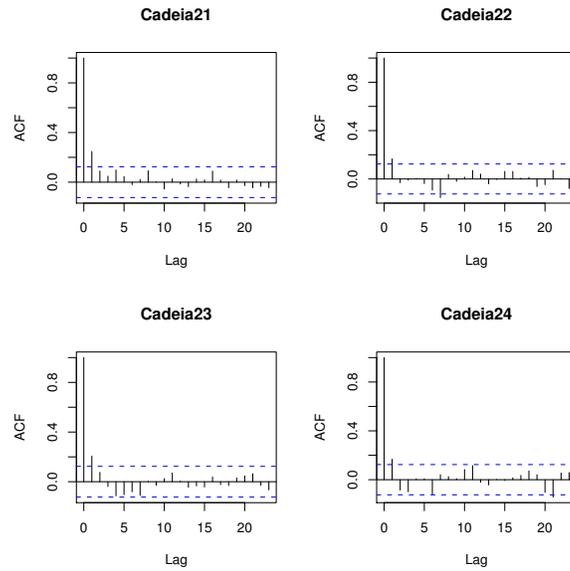


Figura B.32: Acf para as cadeias 21 à 24 com 250 dados- proposta uniforme.

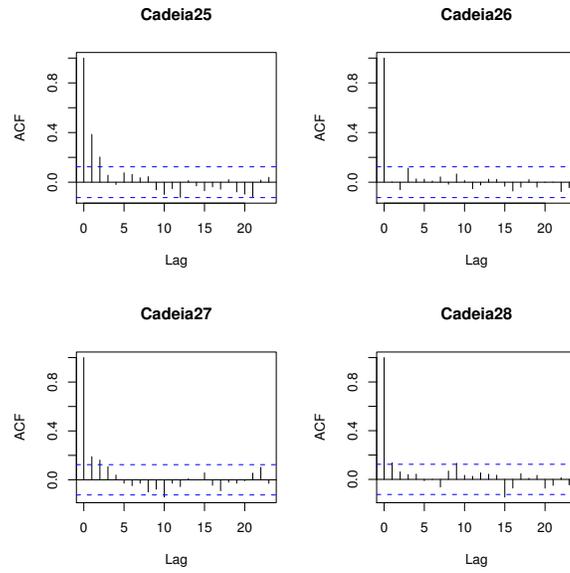


Figura B.33: Acf para as cadeias 25 à 28 com 250 dados- proposta uniforme.

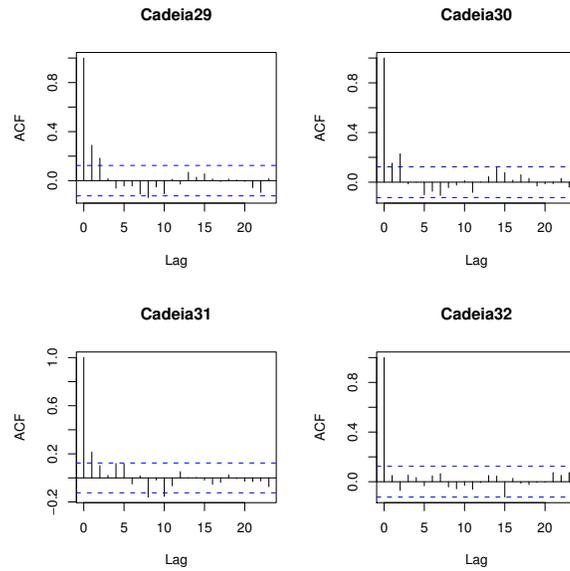


Figura B.34: Acf para as cadeias 29 à 32 com 250 dados- proposta uniforme.

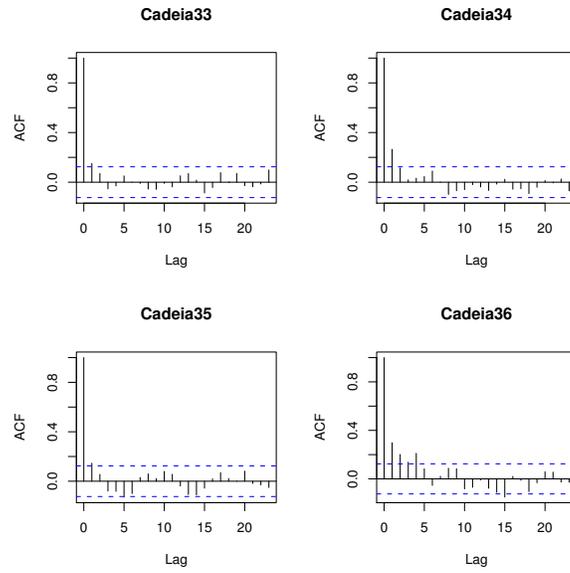


Figura B.35: Acf para as cadeias 33 à 36 com 250 dados- proposta uniforme.

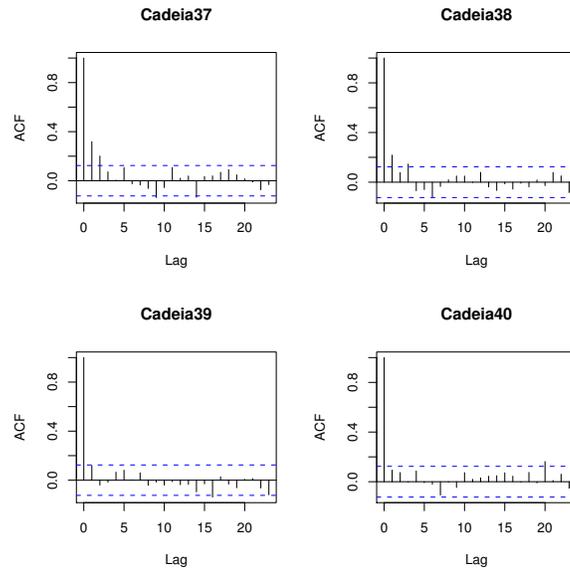


Figura B.36: Acf para as cadeias 37 à 40 com 250 dados- proposta uniforme.

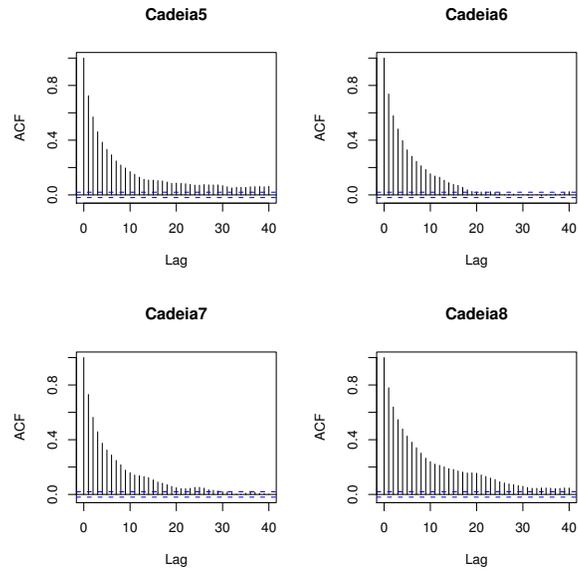


Figura B.37: Acf para as cadeias 5 à 8 - proposta beta.

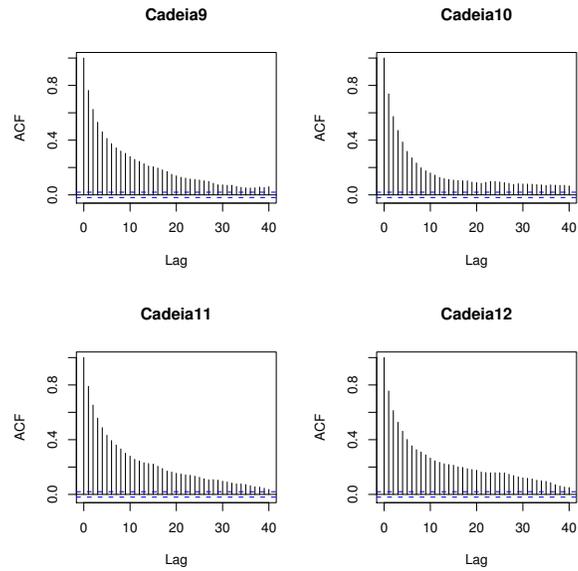


Figura B.38: Acf para as cadeias 9 à 12 - proposta beta.

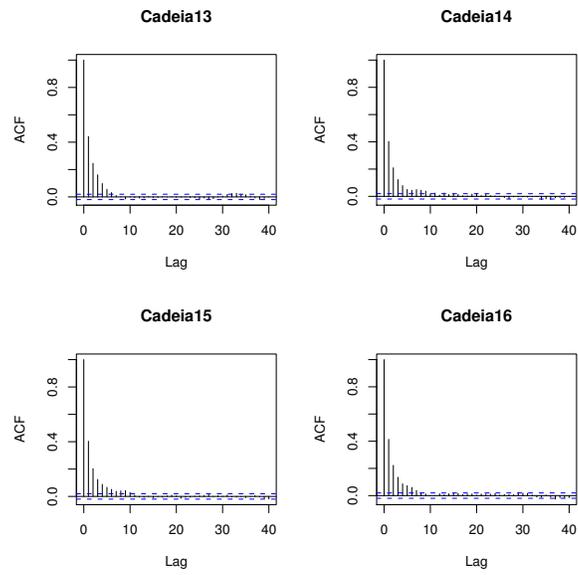


Figura B.39: Acf para as cadeias 13 à 16 - proposta beta.

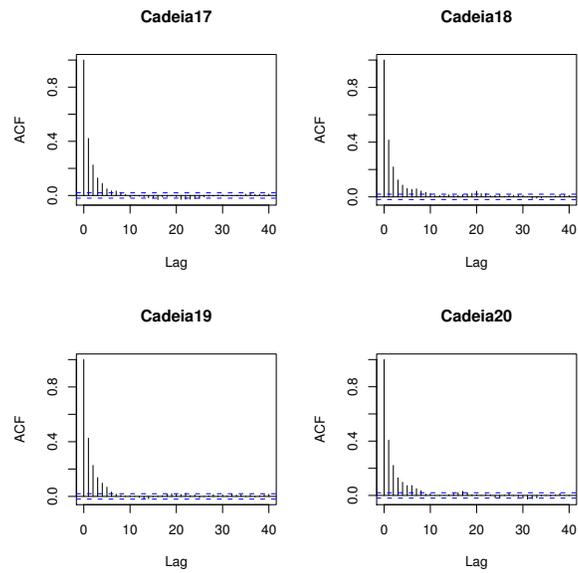


Figura B.40: Acf para as cadeias 17 à 20 - proposta beta.

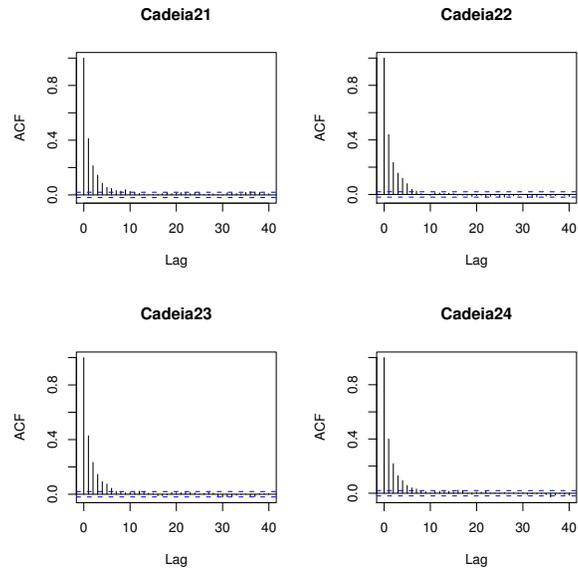


Figura B.41: Acf para as cadeias 21 à 24 - proposta beta.

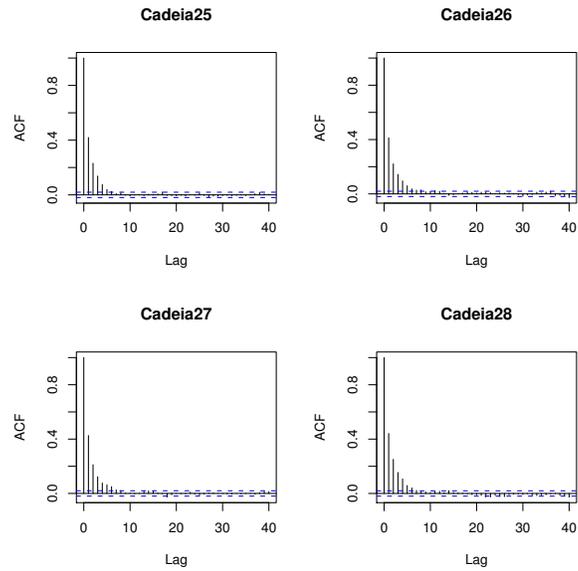


Figura B.42: Acf para as cadeias 25 à 28 - proposta beta.

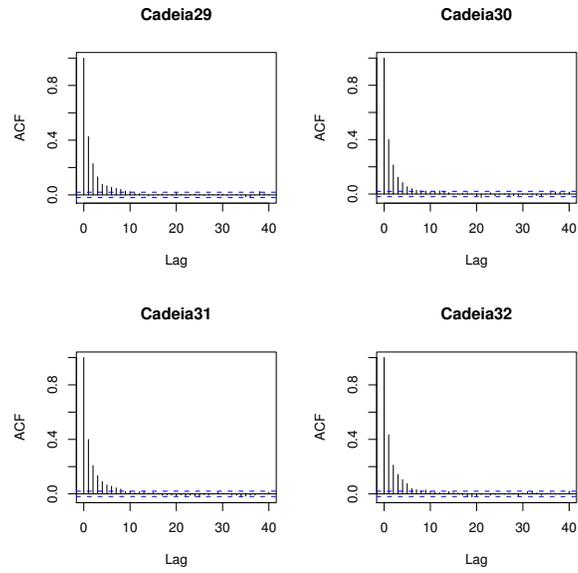


Figura B.43: Acf para as cadeias 29 à 32 - proposta beta.

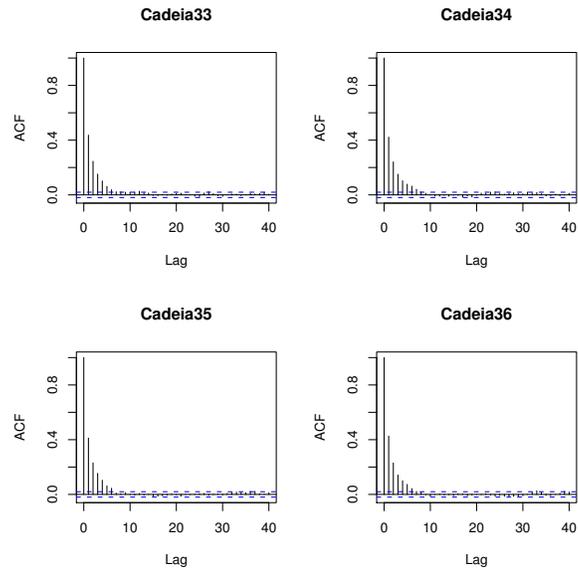


Figura B.44: Acf para as cadeias 33 à 36 - proposta beta.

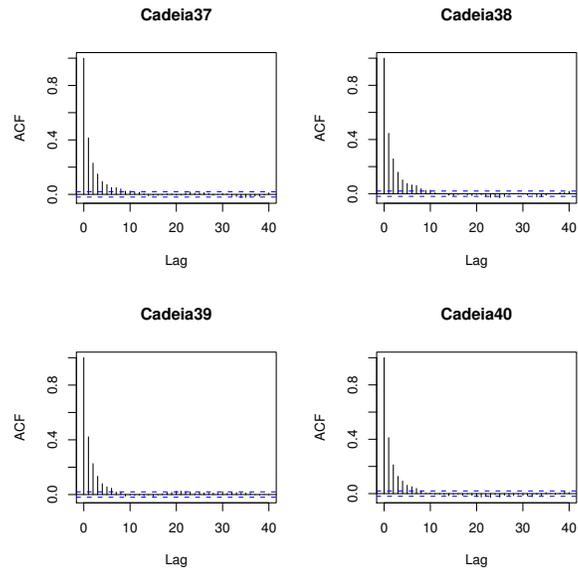


Figura B.45: Acf para as cadeias 37 à 40 - proposta beta.

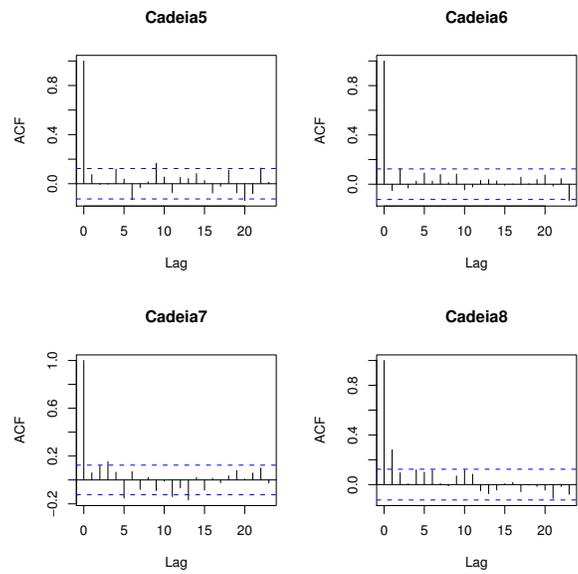


Figura B.46: Acf para as cadeias 5 à 8 com 250 dados - proposta beta.

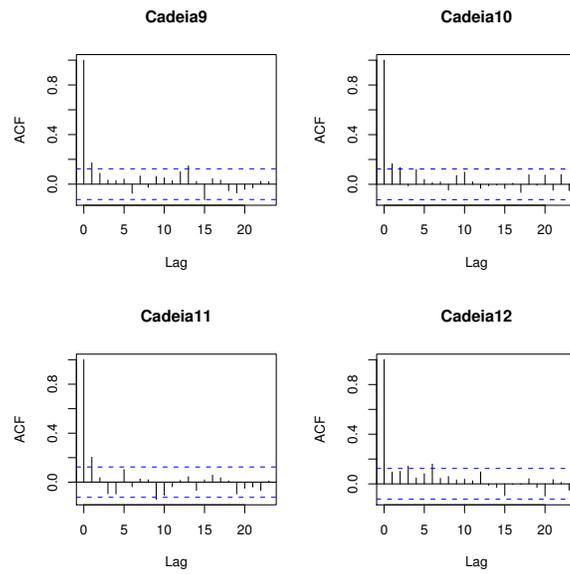


Figura B.47: Acf para as cadeias 9 à 12 com 250 dados- proposta beta.

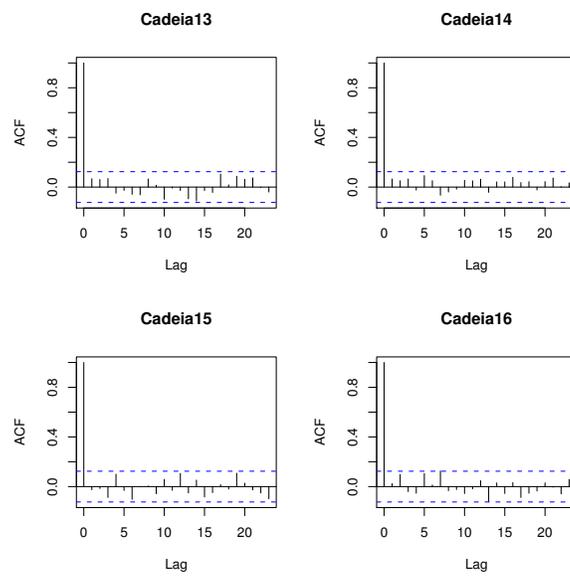


Figura B.48: Acf para as cadeias 13 à 16 com 250 dados - proposta beta.

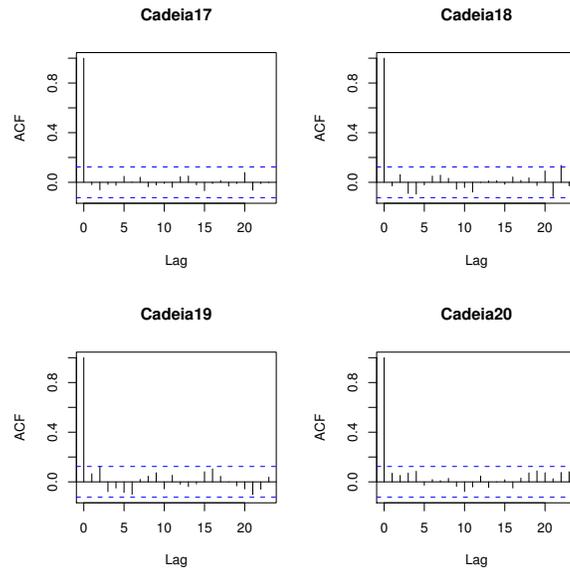


Figura B.49: Acf para as cadeias 17 à 20 com 250 dados - proposta beta.

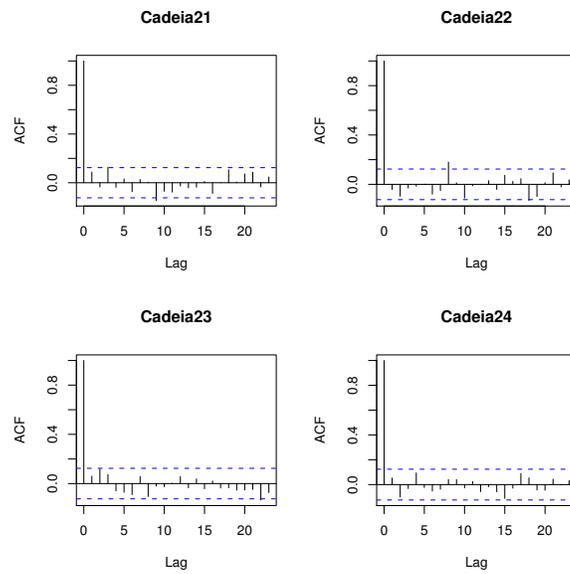


Figura B.50: Acf para as cadeias 21 à 24 com 250 dados - proposta beta.

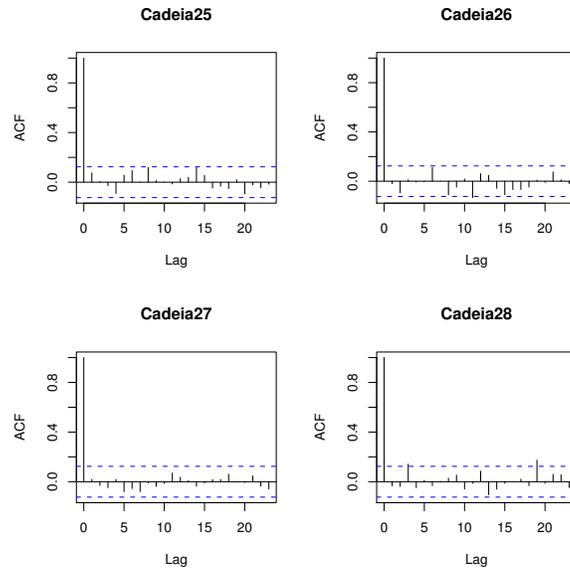


Figura B.51: Acf para as cadeias 25 à 28 com 250 dados - proposta beta.

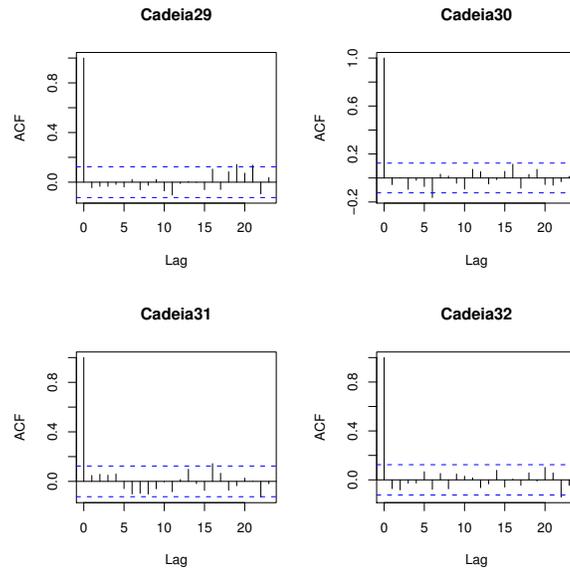


Figura B.52: Acf para as cadeias 29 à 32 com 250 dados - proposta beta.

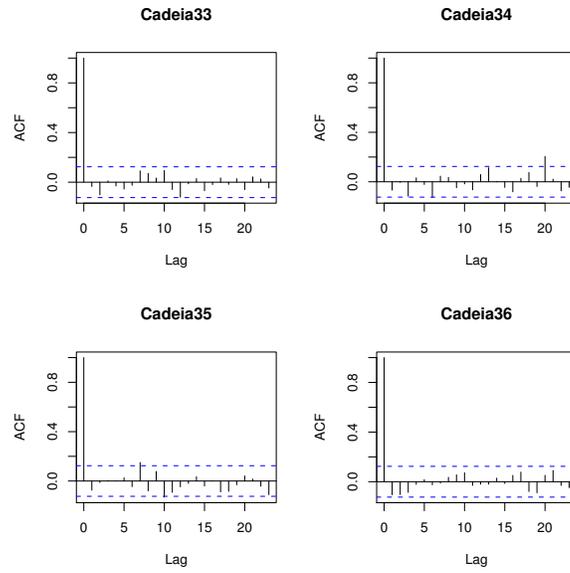


Figura B.53: Acf para as cadeias 33 à 36 com 250 dados - proposta beta.

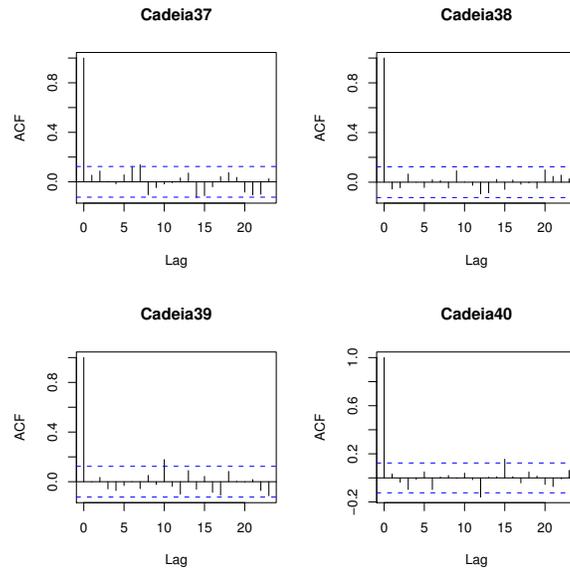


Figura B.54: Acf para as cadeias 37 à 40 com 250 dados - proposta beta.

Apêndice C

Programas

```
##### Geração do processo de nascimento puro através do método da projeção #####
cte = 5
alfa = 2
beta = 1
lambda = 1
K = 5
S = 10

# Gerar v.a.  $X \sim \text{Poisson}(\lambda * |K| * S)$ 
set.seed(1)
X = rpois(1, lambda*K*S )

## Simulando um processo de Poisson com intensidade constante
tau = runif(X, 0, S)
eps = runif(X, 0, K)

#gera C0
C0 = cbind(eps,tau)
PP = C0
geraxy = function(){
m = dim(C0)[1]
y = vector(length=m, mode='numeric')
ystar = vector(length=m, mode='numeric')
xstar = vector(length=m, mode='numeric')
```

```

## Funções
# Remove linhas da matriz de candidatos
remove = function(vector, element) {
  result = matrix(NA, dim(vector)[1] - 1, dim(vector)[2])
  j = 1
  for (i in 1:dim(vector)[1]){
    if (vector[i,1] != element[1] || vector[i, 2] != element[2]) {
      result[j,] = vector[i,]
      j = j + 1
    }
  }
  result
}

# Calcula dentro do mínimo
f = function(element, k, xstar, ystar, y) {
  soma = 0
  for (i in 1:(k-1)) {
    soma = soma + (abs(element[1] - xstar[i]) + cte)alfa
  }
  valor = (element[2] - ystar[k - 1])*exp((-1)*betaalfa*soma) + y[k - 1]
  return(valor)
}

# Calcula f para todos elementos
g = function(C, k, xstar, ystar, y){
  m = dim(C)[1]
  aux = matrix(0, m, 1)
  for (i in 1:m){
    aux[i] = f(C[i, ], k, xstar, ystar, y)
  } return(cbind(min(aux), which.min(aux)))
}

# base do algoritmo recursivo
# para k= 1

```

```

y[1] = min(C0[, 2])
ystar[1] = C0[which.min(C0[, 2]), 2]
xstar[1] = C0[which.min(C0[, 2]), 1]
par = cbind(xstar[1], ystar[1])
C0 = remove(C0, par)

# para 2 <= k <= m
# Gera o novo processo
for(k in 2:m){
  minimo = g(C0, k, xstar, ystar, y)
  y[k] = minimo[1]
  xstar[k] = C0[minimo[2], 1]
  ystar[k] = C0[minimo[2], 2]
  C0 = remove(C0, C0[minimo[2],])
}
xy = cbind(xstar, y)
return(xy)
}
xy = geraxy()
xyold = xy

m = length(xy[,2])
y = sort(xy[,2]) while (y[m] > 5) { set.seed(1)
X = rpois(1, lambda*K*(S+10))
tau2 = runif(X, S, S+10)
eps2 = runif(X, 0, K)
C0 = cbind(c(eps, eps2), c(tau, tau2))
PP = C0
xy = geraxy() y = sort(xy[,2])
m = length(xy[,2])
}
i = 1
for (i in 1:m) {
if (y[i] > 5) break
}
xynew = matrix(0, ncol = 2, nrow=(i-1))

```

```

xynew[,1] = xy[1:(i-1),1]
xynew[,2] = xy[1:(i-1),2]
xy = xynew
m = length(xynew[,2])

```

```

##### Calcula Verossimilhança do Processo #####
##### Ordena os Y #####
ordena = function(vec) {
ordered = matrix(0, dim(vec)[1], dim(vec)[2])
result = sort(vec[, 2], method = "quick", index = TRUE)
for(i in 1:length(result$ix)){
ordered[i,] = vec[result$ix[i],]
}
return(ordered)
}

```

```

##### Função para a primeira parte da verossimilhança #####
parte1 = function(xy, b) {
soma = 0
for (i in 1:(m - 1)) {
for (j in (i+1):m) {
soma = soma + (abs(xy[i,1] - xy[j,1]) + cte)alfa
}
}
return(balfa*soma)
}

```

```

##### Função para a segunda parte da verossimilhança #####
parte2 = function(x, xy, b) {
result = 0
for (l in 2:m){
somay = xy[l,2]- xy[(l-1),2]
somax1 = 0
for (i in 1:(l-1)){
somax1 = somax1 + (abs(x - xy[i,1])+cte)alfa
}
}
}

```

```

valor1 = exp(b(alfa)*somax1)
result = result + somay*valor1
}
somax = 0
for (i in 1:m) {
somax = somax + (abs(x - xy[i,1])+cte)^(alfa)
}
valor = exp(b(alfa)*somax)
result = result + xy[1,2] + (T-xy[m,2])*valor
return(result)
}

```

Calcula a menos log-verossimilhança

```

vero = function(b) {
a = ordena(xy)
N = 5000
x = seq(length=N, 0, 5)
int = (K-0)/(2*N)*(sum(2*parte2(x, a,b)) - parte2(x[1], a, b) - parte2(x[N], a, b)) #regra do trapezio
result = - (parte1(a, b) - int + K*T) ## menos log-verossimilhança
return(result)
}

```

Metropolis-Hastings#####

```

dados = read.table("dados.txt") ### conj. de dados do processo de nascimento puro gerado pelo
metodo da projecao
m = length(dados[,1])
##### Calcula a verossimilhança #####

```

```

vero = function(xy) {
N = 5000
x = seq(length=N, 0, 5)
int = (K-0)/(2*N)*(sum(2*parte2(x, xy)) - parte2(x[1], xy) - parte2(x[N], xy)) ## regra do trapezio
result = exp(parte1 - int + K*T)
return(result)
}

```

```

}

##### ordena xy
ordena = function(vec){ ordered = matrix(0, dim(vec)[1], dim(vec)[2])
result = sort(vec[, 2], method = "quick", index = TRUE)
for(i in 1:length(result$ix)) {
ordered[i,] = vec[result$ix[i],]
}
return(ordered)
}

```

```
##### algoritmo M-H
```

```

n = 10000 ## número de iterações
Y = matrix(NA,(n+1),m)
Y[1,] = runif( m, 0, 5)
xy1 = matrix(0,ncol = 2, nrow = m)
z = matrix(0,ncol=2, nrow=m)
xy1[,1] = dados[,1]
z[,1] = dados[,1]
for(i in 2:(n+1)){
xy1[,2] = Y[i-1,]
ystar = ordena(xy1)
z[,2] = runif(m,0,5)
zstar = ordena(z)
A = vero(zstar)/vero(ystar)
prob = min(1,A)
u = runif(1)
if(u < prob){
y = z[,2]
} else {
y = Y[i-1,]
}
Y[i,] = y
}
##### Monte Carlo #####

```

```

dados = matrix(scan("dados.txt"), ncol=2, byrow=T)
m = length(dados[,1])
beta0 = 1 ##### definir beta0
mc = 250 ### valor utilizado na aproximação Monte Carlo
montecarlo = function(b){
  L = 0
  for(j in 1:mc){
    xy = cbind(dados[,1], mh[j,])
    L = L + vero(xy, b)/vero(xy,beta0)
  }
  dm = log(mc) - log(L) ### menos log-verossimilhança
  return(dm)
}
mh = matrix(scan(paste("MH",s,"_250.txt", sep="´")), ncol=34, byrow=T)
min = optimize(montecarlo, c(0,2))

```

Referências Bibliográficas

- [1] Arthur Dempster, Nan Laird, and Donald Rubin, (1977), “Maximum likelihood from incomplete data via the EM algorithm”, *Journal of the Royal Statistical Society, Series B*, **39(1)**, pp. 1–38
- [2] Beyer, J.E., Keiding, N. and Simonsen, W. (1976) “The exact behaviour of the maximum likelihood estimator in the pure birth process and the pure death process”, *Scandinavian Journal of Statistics*, **3**, pp. 61–72.
- [3] Boots, B., Okabe, A. and Thomas, R., (2003), “Modelling Geographical Systems: Statistical and Computational Applications”, Kluwer Academic Publishers, Dordrecht.
- [4] Cox, D.R. and Isham, V. (1980), *Point Process*, Chapman & Hall/CRC.
- [5] Cressie, N. A. C, (1993), *Statistics for Spatial Data*
- [6] Daley, D. J. and Vere-Jones, D. (1988) *An introduction to the theory of point processes*. Springer-Verlag, New York, NY.
- [7] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) “Maximum likelihood from incomplete data via the EM algorithm (with discussion) ”, *J.R. Statit. Soc. B*, **39**, pp. 1–37.
- [8] Diggle, P. J. (2003) *Statistical Analysis of Spatial Point Patterns*. 2ed. London: Arnold
- [9] Garcia, N.L. (1995a) “Birth and Death Processes as Projections of Higher Dimensional Poisson Processes”, *Advances in Applied Probability*, **27**, No. 4, pp. 911–930.
- [10] Garcia, N. L. (1995b) “Approximations of large population epidemic models”, *Stoch. Proc. and their Appl*, **60**, pp. 147-160.
- [11] Garcia, N. L. e Kurtz, T.G (2008) “Spatial point processes and the projection method” , *Progress in Probability*, **60**, pp. 271-298.
- [12] Garcia, N. L. e Kurtz, T.G (1998) “Maximum Likelihood Estimates for Spatial Pure Birth Process”. Relatório Técnico IMECC 66/98.

- [13] Gelfand, A. E. and Carlin, B. P. (1993) “Maximum-likelihood estimation for constrained- or missing-data models.” *Canad. J. Statist*, **21**, pp. 303-311.
- [14] Geyer, C.J. (1992) “Practical Markov Chain Monte Carlo”, *Statistical Science*, **7**, No. 4, pp. 473–511.
- [15] Geyer, C.J. (1994) “On the Convergence of the Monte Carlo Maximum Likelihood Calculations”, *J. R. Statist. Soc. B*, **56**, pp.261–274.
- [16] Geyer, C. J. (1999). “Likelihood inference for spatial point processes”, *Stochastic Geometry: Likelihood and Computation* (eds. O. E. Barndor–Nielsen, W. S. Kendall and M. N. M. van Lieshout), Chapman & Hall/CRC, Boca Raton, Florida, pp. 79–140.
- [17] Geyer, C. J. and Moller, J. (1994), “Simulation procedures and likelihood inference for spatial point processes”, *Scandinavian Journal of Statistics*, **21**, pp.359–373.
- [18] Guo, S.W. and Thompson, E.A. (1991) “Monte Carlo estimation of variance component models for large complex pedigrees”, *IMA J. Math. Appl. Med. Biol.*, **8**, pp. 171–189.
- [19] Hastings, W., K. (1970) “Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97-109.
- [20] Isham, V. (1981), “An Introduction to Spatial Point Processes and Markov Random Fields”, *International Statistical Review / Revue Internationale de Statistique*, **49**, pp. 21–43
- [21] Keiding, N. (1974) “Estimation in the birth process”, *Biometrika*. **61**, pp. 71–80.
- [22] Kerscher, M. (2000), “Statistical analysis of large-scale structure in the Universe”, *Statistical Physics and Spatial Statistics* (eds. K. R. Mecke and D. Stoyan), Lecture Notes in Physics, Springer, Berlin, 36–71.
- [23] Kipnis, C. and Varadhan, S.R.S. (1986), “Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions”, *Comm. Math. Phys.*, **104**, pp. 1–19.
- [24] Kurtz, T.G. (1989) “Stochastic Processes as Projections of Poisson Random Measures”, Special invited paper at IMS meeting, Washington, D.C.. Unpublished.
- [25] Kong, A., Liu, J. S., and Wong, W. H. (1994) “Sequential imputations and Bayesian missing data problems.” *J. Amer. Statist. Assoc.*, **89**, pp.278-288.

- [26] Lange, K. and Sobel, E. (1991) “A random walk method for computing genetic location scores.” *Am. J. Hum. Genet.*, **49**, pp. 1320-1334.
- [27] Little, R. J. A and Rubin, D. B (2002) *Statistical analysis with Missing Data* 2^a ed., Wiley, Hoboken.
- [28] Moller, J. and Waagepetersen, R. P.(2003) *Statistical Inference and Simulation for Spatial Point Process*, Chapman & Hall/CRC, Boca Raton.
- [29] Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. Teller, E. (1953) “Equation of State Calculations by Fast Computing Machines”, *J. Chem. Phys.*, **21**, pp. 1087–1092.
- [30] Moran, P.A.P. (1951), “Estimation methods for evolutive processes”, *J. R. Stat. Soc. B*, **13**, pp. 141–146.
- [31] Ogata, Y. (1998), “Space-Time point process models for earthquake occurrence”, *Annals of Institute of Statistical Mathematics*, **50**, pp. 379–402
- [32] Ott, J. (1979), “Maximum likelihood estimation by counting methods under polygenic and mixed models in human pedigrees, *Am. J. Hum. Genet.*, **31**, pp.161-175.
- [33] Rathbun, S.L., (1996), “Asymptotic properties of the maximum likelihood estimator for spatio-temporal point processes”, *Jornal of Statistical Planning and Inference*, **51**, pp.55–74.
- [34] Ripley, B. D.(1981) *Spatial Statistics*. Wiley, New York.
- [35] Ripley, B. D. (1988) *Statistical Inference for Spatail Process* Cambridge University Press, Cambridge.
- [36] Ripley, B. D. and Kelly, F. P.,(1977) “ Markov Point Process”, *Jornal of London Mathematical Society*, **15**, pp. 188–192.
- [37] Thompson, E.A. and Guo, S.W. (1991) “Evaluation of likelihood ratios for complex genetic models”, *IMA J. Math. Appl. Med. Biol.*, **8**, pp. 149–169.
- [38] Thompson, E. A. (2003,) Linkage analysis. In *Handbook of Statistical Genetics*. 2^a ed. (D. J. Balding, M. Bishop, and C. Cannings, eds) 893-918. Wiley, Chichester.
- [39] Stoyan, D. and Stoyan H (1998) “Non-homogeneous Gibbs process models for forestry - a case study”, *Biometrical Journal*, **40**, pp.521–531
- [40] Stoyan, D., Kendall, W. S. and Mecke, J. (1995), *Stochastic Geometry and Its Applications*, Wiley, Chichester, 2nd edition.

- [41] Sung, J., Y. and Geyer, J.,C. (2007) “Monte Carlo likelihood inference for missing data models”, *Annals of Statistics* , **3**, pp. 990-1011.
- [42] Wei, G.C.G. and Tanner, M.A. (1990) “ A Monte Carlo implementation of the EM algorithm and poor man’s augmentation”, *J. Am. Stat. Assoc.*, **85**, pp. 699–704.