Márcio Augusto Diniz

Modelos SEIR com taxa de remoção não homogênea

Orientador:
Jorge Alberto Achcar

Coorientador:
Luiz Koodi Hotta

UNIVERSIDADE ESTADUAL DE CAMPINAS

Campinas, São Paulo
2011

Modelos SEIR com taxa de remoção não homogênea

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Márcio Augusto Diniz, aprovada pela Comissão Julgadora.

Campinas 23 de fevereiro de 2011

Prof. Dr.Jorge Alberto Achcar

Orientador

Prof. Dr.Luiz Koodi Hotta

Coorientador

Banca Examinadora:

- Prof. Dr. Jorge Alberto Achcar (FMRP - USP) - Orientador
- Prof. Dr. Aluísio de Souza Pinheiro (IMECC - UNICAMP)
- Prof. Dr. Edson Zangiacomi Martinez (FMRP - USP)

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

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA DO IMECC DA UNICAMP
Bibliotecária: Maria Fabiana Bezerra Müller – CRB8 / 6162

Diniz, Márcio Augusto
D615m Modelos SEIR com taxa de remoção não homogênea/Márcio Augusto Diniz-- Campinas, [S.P. : s.n.], 2011.

Orientador : Jorge Alberto Achcar ; Luiz Koodi Hotta
Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1.Modelos epidemiológicos SEIR. 2.Epidemiologia. 3.Inferência bayesiana. 4.Métodos MCMC. I. Achcar, Jorge Alberto. II.Hotta, Luiz Koodi. III. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Título em inglês: SEIR models with time in-homogeneous removal rate

Palavras-chave em inglês (Keywords): 1.SEIR epidemic models. 2.Epidemiology. 3.Bayesian inference. 4.MCMC methods.

Área de concentração: Métodos Estatísticos

Titulação: Mestre em Estatística

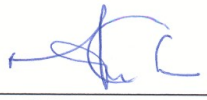
Banca examinadora: Prof. Dr. Jorge Alberto Achcar (FMRP – USP)
Prof. Dr. Aluísio de Souza Pinheiro (IMECC – UNICAMP)
Prof. Dr. Edson Zangiacomi Martinez (FMRP - USP)

Data da defesa: 23/02/2011

Programa de Pós-Graduação: Mestrado em Estatística

Dissertação de Mestrado defendida em 23 de fevereiro de 2011 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.



Prof(a). Dr(a). JORGE ALBERTO ACHCAR

Prof(a). Dr(a). ALUÍSIO DE SOUZA PINHEIRO

Prof(a). Dr(a). EDSON ZANGIACOMI MARTINEZ

"Para tudo há um tempo, para cada coisa há um momento debaixo dos céus:
tempo para nascer,
e tempo para morrer;
tempo para plantar,
e tempo para arrancar
o que foi plantado;
tempo para matar,
e tempo para sarar;
tempo para demolir
e tempo para construir;
tempo para chorar,
e tempo para rir;
tempo para gemer,
e tempo para dançar;
tempo para atirar pedras,
e tempo para ajuntá-las;
tempo para dar abraços,
e tempo para apartar-se.
Tempo para procurar,
e tempo para perder;
tempo para guardar,
e tempo para jogar fora;
tempo para rasgar,
e tempo para costurar;
tempo para calar,
e tempo para falar;
tempo para amar,
e tempo para odiar;
tempo para a guerra,
e tempo para a paz."

Eclesiastes 3, 1 - 8

Agradecimentos

Há inúmeras pessoas que merecem receber os meus agradecimentos, sem elas, não seria possível esta dissertação. Por isso, esta pequena homenagem será extensa.

Gostaria de agradecer a todas as pessoas que se tornaram parte da minha família. Em especial, minha mãe Maria e meu pai Antônio por estarem sempre ao meu lado nas dificuldades, por aceitarem minhas decisões, me apoiarem sem qualquer hesitação e terem me propiciado os princípios para ser este ser humano pelo qual sinto orgulho de ser.

Aos meus irmãos Marco Aurélio, Maurício e minha irmã Márcia por serem meus irmãos mais velhos e exercerem este papel tão bem, pelas jogatinas de video-game e livros emprestados para que eu pudesse espairecer dos estudos. Também, as minhas mães que sempre estão a orar por mim: Ana Maria Pereira, Ana Maria Rosa, Izolina e Narcisa.

Não posso esquecer-me do Colégio Juarez Wanderley Siqueira Britto e seus funcionários que abriram as portas da universidade pública, estas seriam desconhecidas para mim se não fosse a oportunidade de estudo concedida. Agradeço aos meus bons amigos daquela época e que, às vezes, eu tenho o prazer de reencontrá-los: Alessandra, Anderson, Ariane Alves, Camões, Ives, Janaína, Letícia, Maria Cecília, Michelle, Simone, Sônia, Thiago Carneiro e Tatiane. Ao lado de vocês, a jornada da vida tornou-se mais prazerosa.

É imprescindível agradecer aos meus amigos de graduação em estatística, pois ao lado deles, enfrentei os grandes desafios da vida acadêmica e pude me apaixonar por ela: Adriana Gushiken, Ana Carolina, Carlos, Daniel, Fabrício, Felipe, Fernanda Marion, João Ricardo, Jordana, Juliana, Lidiane, Luis, Marcelo, Matheus, Omar, Pedro Flores, Renata, Ricardo, Taísa e Thiago.

E também as amigas do instituto do outro lado da rua que me permitiam esquecer os problemas da estatística e respirar outros ares: Andrea, por mostrar um pouco da verdade sobre o mundo; Fabi, por me dar a oportunidade de ter trabalhado com a demografia; Aldrey e Vanessa, pela alegria constante.

Aos meus companheiros de moradia: Renato, Erick, Gustavo, Denize, Gabrielle, Eveline e, principalmente, Alda Mara e Thiago Juarez, com os quais, aprendi a lidar com as diferenças e superei as dificuldades de viver fora da casa dos meus pais.

Aos meus amigos da pós-graduação, companheiros frente à difíceis provas, churrascos, reclamações, saídas de sábado a noite após longos dias de estudos, boas risadas. Ao lado deles, eu comecei, realmente, a viver bem a vida: Beatriz, por sempre lembrar-me de todas as datas importantes e desesperar-se junto comigo frente as provas; os casais 20, Márcio Valk e Karen, Denise e Airton pelos almoços e passeios; Marcio Rodrigues, Lorena e Christiane pelas boas risadas; Guilherme, pelo companheirismo, conversas filosóficas e dicas de programação; Caroline e Larissa pelos momentos de descontração e torcida; José, pelas dicas sobre MCMC; Sheila e Jonas pelos momentos inusitados; Diego, Karina, Aldo e Pedro por serem os meus veteranos do mestrado e me darem as dicas de sobrevivência.

Aos amigos do Laboratório de Epidemiologia e Fisiologia Matemática: Licia, Roberta, Marcio Sabino, Cintia, com quem compartilhei bolos, limpezas da máquina de café, frio devido ao ar condicionado e ideias sobre epidemiologia.

Aos meus bons amigos Francisco, Denny, Kenya e Érica pela amizade que transpõem o tempo e a distância, por sempre ouvirem minhas conversas sobre as coisas da vida, por compartilhar minhas paixões e por não desistirem da minha amizade por mais que eu me isolasse devido aos estudos. E a Ana Carolina Rodrigues pela amizade cheia de alegria e ser capaz de resgatar-me dos estudos de tempos em tempos.

Ao Max pelos esquemas utilizados na dissertação, por não ligar para as minhas esquisitices, pelas longas conversas e pela grande amizade em tão pouco tempo. A Dinah pela revisão de inglês e as boas risadas proporcionadas.

Ao Fabio pelas constantes conversas em que ao invés de animar-me, me colocava para baixo, fazendo revoltar-me e seguir em frente por pirraça. Ao Thalles e Gladston pelas boas conversas sobre os mais diversos assuntos. Ao Weber por ser um mestre nas coisas boas da vida e ensinar-me um pouco sobre elas.

As minhas terapeutas Lourdes, Maria Lilian e Maria Lidia por caminharem comigo esta longa estrada e permitirem que eu fosse capaz de enxergar-me e acreditar nos meus recursos para superar as dificuldades.

A professora Laura, por ter sido responsável na minha persistência do estudo da estatística no início da graduação. A professora Mariana e Marina por serem exemplos de professores que não se esqueceram que um dia foram estudantes; a professora Nancy por ser um exemplo de dedicação e superação; aos funcionários do Imecc, principalmente, Tânia e Lívia pela grande ajuda em todas as atividades burocráticas da pós-graduação.

Ao meu orientador Luiz Koodi Hotta, que me deu a oportunidade de pesquisa desde a graduação e sempre exigiu que eu melhorasse como pesquisador e também como pessoa com bom humor e rigor quando necessário. A contribuição dada por ele é imensurável em todos os aspectos da minha vida.

Ao meu orientador Jorge Alberto Achcar, que compartilhou seu enorme conhecimento, ideias criativas e compreendeu minhas dificuldades e sempre acreditou na minha

capacidade como pesquisador. Aos professores Aluísio de Souza Pinheiro e Edson Zangiacomi Martinez pela disponibilização do tempo e correções sugeridas.

A FAPESP pelo financiamento do projeto e a Unicamp por toda minha formação como estatístico.

E finalmente, ao meu Pai que está no céu por todas estas pessoas presentes na minha vida, oportunidades concedidas e por permitir enxergar a grandeza de todas as coisas.

Eu agradeço do fundo do meu coração e espero que vocês possam continuar a fazer parte da minha vida sejam através de conversas rápidas, convivência diária, e-mails ou reencontros.

Resumo

A modelagem matemática de epidemias apresenta grande relevância para a área de epidemiologia por possibilitar uma melhor compreensão do desenvolvimento da doença na população e permitir analisar o impacto de medidas de controle e erradicação.

Neste contexto, os modelos compartimentais SEIR (Suscetíveis - Expostos - Infectantes - Removidos) que foram introduzidos por Kermarck e Mckendrick (1927 apud YANG, 2001, Capítulo 1) são extremamente utilizados.

Esta dissertação discute o modelo SEIR encontrado em Lekone e Finkenstädt (2006) que considera a introdução das medidas de intervenção na taxa de contato entre suscetíveis e infectantes, e é aplicado aos dados parcialmente observados da epidemia de febre hemorrágica Ebola, ocorrida no Congo em 1995, através de métodos bayesianos.

Em uma segunda etapa, o modelo é modificado a fim de considerar a introdução das medidas de intervenção também na taxa de remoção. A utilização da taxa de remoção não homogênea permite uma quantificação do impacto das medidas de intervenção mais próxima da realidade.

Além disso, nos dois modelos considerados, uma análise da incerteza gerada pela observação parcial dos dados e uma análise de sensibilidade da escolha das distribuições a priori são realizadas a partir de simulações. E também, é apresentada uma discussão sobre erros de especificação da taxa de remoção.

Por fim, os dois modelos são aplicados aos dados da epidemia de febre hemorrágica Ebola e os resultados são discutidos.

Abstract

Mathematical modeling of epidemics has great relevance to the area of epidemiology by enabling a better understanding of the development of the disease in the population and allow the analysis of the impact of eradication and control measures.

In this context, SEIR (Susceptible-Exposed-Infectious-Removed) compartmental models that were introduced by Kermarck e Mckendrick (1927 apud YANG, 2001, Chapter 1) are extremely used.

This essay discusses the SEIR model found in Lekone e Finkenstädt (2006) that considers the introduction of intervention measures in the contact rate between susceptible and infectious, and is applied to data partially observed the outbreak of Ebola hemorrhagic fever in Congo, held in 1995, by Bayesian methods.

In a second step, the model is modified in order to consider the introduction of intervention measures in the removal rate. The use of time in-homogeneous removal rate allows quantification of the impact of intervention measures closer to reality.

In addition, both models considered, an analysis of uncertainty generated by partial observation and a sensitivity analysis of the choice of a priori distributions are made from simulations. And also, a discussion about errors of removal rate specification.

Finally, the two models are applied to the data of Ebola hemorrhagic fever epidemic and the results are discussed.

Lista de Figuras

Lista de Tabelas

Sumário

1 Introdução

A modelagem matemática de epidemias é de grande importância para os estudos epidemiológicos por possibilitar um melhor entendimento do desenvolvimento da epidemia e a busca por medidas eficientes em sua prevenção e erradicação. (YANG, 2001, Capítulo 1)

Os primórdios da epidemiologia são datados da Grécia Antiga na qual há registros sobre epidemias e especulações de possíveis causas de doenças como, por exemplo, a obra Epidemias de Hipócrates (459 - 377 a.C.).

Todavia, pouco progresso significativo na epidemiologia pôde ser obtido anterior ao surgimento da bacteriologia na metade do século XIX com Louis Pasteur (1822 - 1895) e Robert Koch (1843 - 1910).

Neste período anterior ao surgimento da bacteriologia, há alguns estudos estatísticos e matemáticos realizados, sendo o trabalho de John Graunt (1620 - 1674) e William Petty (1623 - 1687) sobre as tábuas de mortalidade de Londres considerado como o marco inicial da estatística médica e o entendimento em larga escala da relação entre doenças e mortalidade.

Daniel Bernoulli, em 1760, utilizou um método matemático para avaliar a efetividade da técnica de variolação contra a varíola a fim de influenciar as políticas públicas. John Snow mostrou, em 1855, que a cólera era transmitida através da contaminação das fontes de abastecimento de água; William Budd, em 1873, estabeleceu similarmente para o espalhamento da febre tifóide.

A abordagem empírica de tais trabalhos deve-se ao estágio de desenvolvimento das ferramentas matemáticas e à insuficiência de hipóteses precisas sobre os mecanismos de

espalhamento de uma doença, desconhecidos até o surgimento da bacteriologia.

Farr (1840 apud BAILEY, 1975, Capítulo 1), Evans (1875 apud BAILEY, 1975, Capítulo 1) e Brownlee (1906 apud BAILEY, 1975, Capítulo 1) apresentaram trabalhos com uma maior sofisticação matemática ao ajustar dados de mortes de diversas epidemias às curvas normais e tentaram construir predições a partir de tais ajustes, sem obter sucesso significativo.

No final do século 19, a partir do conhecimento gerado pela bacteriologia e a familiaridade com dados epidemiológicos, Hammer (1906 apud BAILEY, 1975, Capítulo 1) estabeleceu o princípio de ação de massas que consiste em afirmar que o desenvolvimento de uma epidemia depende da taxa de contato entre suscetíveis e infectantes.

O princípio de ação de massas é uma hipótese básica para toda a teoria determinística desenvolvida posteriormente e com pequenas modificações, também, para a teoria estocástica. Inicialmente, estabelecido para o tempo discreto, Ross (1908 apud BAILEY, 1975, Capítulo 1) generalizou para o tempo contínuo.

Além disso, Ross (1911 apud BAILEY, 1975, Capítulo 1) formulou os primeiros modelos determinísticos para o desenvolvimento de uma epidemia. Nesta mesma linha, Kermarck e Mckendrick (1927 apud YANG, 2001, Capítulo 1) introduziram o modelo determinístico denominado SEIR (Suscetíveis - Expostos - Infectantes - Removidos) e também estabeleceram o Teorema de Valor Limiar que afirma que a introdução de indivíduos infectantes em uma população de suscetíveis não resultará em uma epidemia se o número de indivíduos suscetíveis estiver abaixo de um limiar. Por outro lado, se o número de suscetíveis estiver acima do limiar, a epidemia será desencadeada até que o número de suscetíveis torne-se menor do que o valor limiar.

Paralelamente, os modelos estocásticos foram desenvolvidos, pois em pequenas epidemias ocorridas em grandes áreas, os elementos aleatórios tornavam-se relevantes e a aproximação dada pelos modelos determinísticos não era suficiente. O primeiro modelo estocástico foi introduzido por Mckendrick (1926 apud BAILEY, 1975, Capítulo 1) e era também baseado no princípio de massas, contudo, numa versão probabilística.

Entretanto, a inexistência de ferramentas suficientes para lidar com a estimação de tal modelo não permitiu o seu desenvolvimento. Um outro modelo estocástico que

tornou-se muito mais popular foi o modelo de Reed e Frost introduzido em 1928, este modelo assume que o período de infecção é extremamente curto e o período de exposição é constante. E, a partir de um caso inicial em um grupo fechado, novos casos surgem em uma série de estágios de geração através da distribuição Binomial. Mais detalhes podem ser encontrados em Abbey (1952 apud BAILEY, 1975, Capítulo 1).

Os modelos epidemiológicos, tanto determinísticos quanto estocásticos, continuaram sendo estudados e no início da década de 40, os avanços na teoria de processos estocásticos permitiram o desenvolvimento do modelo introduzido por Mckendrick (1926 apud BAILEY, 1975, Capítulo 1) e diversas comparações com o modelo de Reed e Frost foram feitas.

Bartlett (1949 apud KYPRAIOS, 2007, Capítulo 2) estudou uma versão estocástica do modelo introduzido por Kermarck e Mckendrick (1927 apud YANG, 2001, Capítulo 1) que se tornou bastante popular e conhecido por Modelo de Epidemia Geral. O modelo SEIR estudado nesta dissertação é uma variação do Modelo de Epidemia Geral.

Mais detalhes podem ser encontrados em Bailey (1975, Capítulo 1) que apresenta um histórico detalhado sobre os trabalhos desenvolvidos e relevantes até o início da década de 1970 na área de epidemiologia.

Além disso, diversas discussões teóricas sobre modelos epidemiológicos e diversas abordagens foram desenvolvidas para lidar com doenças infecciosas via contato direto, vetores, doenças venéreas e sob o ponto de vista temporal e espacial do espalhamento das doenças. Becker (1989 apud DEMIRIS, 2004, Capítulo 1) revisou os métodos estatísticos aplicados a modelagem de epidemias até o final da década de 1980.

Contudo, até o início da década de 1990, o impacto prático de tais desenvolvimentos teóricos ainda não era tão grande, como afirmam Anderson e May (1991, Capítulo 1). A justificativa para este cenário são várias: a falta de comunicação entre as partes envolvidas, a complexidade matemática de modelos mais realísticos, a falta de dados ou a pouca qualidade dos mesmos.

Este cenário está sendo alterado devido a dois fatores segundo O'Neill (2010). O primeiro ponto consiste no engajamento da comunidade estatística com outras áreas a fim de atender questões de saúde pública com diversos exemplos como avaliação de

estratégias de vacinação para HPV nos E.U.A. por Kim e Goldie (2008 apud O'NEILL, 2010) e a estimação de quantidades de interesse epidemiológico do vírus H1N1 por McBryde et al. (2009 apud O'NEILL, 2010). O segundo fator é o acesso e o ganho de grande poder computacional que é utilizado para lidar com grandes bancos de dados e a utilização de métodos estatísticos como MCMC.

O'Neill (2010) faz uma breve revisão do atual desenvolvimento dos métodos estatísticos para a modelagem de epidemias e os desafios a serem enfrentados na área que consistem na inferência para grande populações e modelos complexos, na utilização de dados aumentados e nos critérios de seleção e diagnósticos de modelos. Neste contexto, esta dissertação é proposta como uma contribuição na modificação deste cenário ao considerar um conjunto de dados real, pouco explorado e com a presença de dados faltantes e medidas de intervenção.

Neste capítulo, há a apresentação de conceitos fundamentais da epidemiologia na seção 1.1, do modelo SEIR que será utilizado na modelagem do conjunto de dados na seção 1.2, uma breve explanação sobre a doença Febre Hemorrágica Ebola na seção 1.3, a descrição do conjunto de dados referente a epidemia de Ebola ocorrida na República Democrática do Congo em 1995 na seção 1.4, a apresentação dos objetivos da dissertação na seção 1.5 e, por fim, a apresentação da estrutura da dissertação na seção 1.6.

1.1 Conceitos fundamentais em epidemiologia

A partir de Yang (2001, Capítulo 1), pode-se ressaltar alguns conceitos importantes:

Indivíduo Suscetível: É aquele indivíduo que não teve qualquer indício do agente causador da doença.

Indivíduo Exposto: É aquele indivíduo que apresenta a doença, contudo o indivíduo ainda não é capaz de transmití-la.

Indivíduo Infectante: É aquele indivíduo que apresenta a doença e também é capaz de transmití-la para outros indivíduos.

Indivíduo Removido: É aquele indivíduo que deixou de contribuir para a transmissão da doença seja porque foi isolado, apresentou recuperação ou faleceu.

Doenças de transmissão direta: As doenças de transmissão direta são doenças em que a transmissão ocorre diretamente, através de meio físico, quando há um contato apropriado entre os indivíduos suscetíveis e infectantes.

Período de Exposição: Compreende o período em que o indivíduo é dito ser exposto.

Período de Infecção: Compreende o período em que o indivíduo é dito ser infectante.

Período de Imunidade: Compreende o período em que o indivíduo é dito ser imune.

Endemia: É a ocorrência estável do número de casos da doença dentro de uma comunidade.

Epidemia: É a ocorrência do número de casos da doença em excesso em relação ao número esperado de casos dentro de uma comunidade.

Taxa de contato: É a taxa denotada por β que engloba todas as componentes do padrão de contatos entre indivíduos suscetíveis e individuos infectantes.

Número de reprodutibilidade basal: É o número esperado de casos secundários que um caso primário é capaz de produzir em uma população suscetível.

Número de reprodutibilidade efetivo: É o número esperado de casos secundários que um caso primário é capaz de produzir em uma população suscetível no período t.

1.2 O modelo SEIR

Os modelos SEIR (Suscetíveis - Expostos - Infectantes - Recuperados) foram introduzidos por Kermarck e Mckendrick (1927 apud YANG, 2001, Capítulo 1) e têm sido amplamente utilizados por ajustar-se a diversas epidemias. Tais modelos são construídos ao considerar uma população fechada e na sua divisão em quatro compartimentos:

- Suscetíveis;

- Expostos;
- Infectantes;
- Removidos;

Há também variações do modelo SEIR como, por exemplo, o modelo SIR que desconsidera o compartimento dos indivíduos Expostos ou o modelo SIS que considera que um indivíduo infectante torna-se novamente suscetível.

A estimação de tais modelos é complexa devido, frequentemente, à presença de dados incompletos e, consequentemente, uma função de verossimilhança quase sempre intratável analiticamente.

Há diversas abordagens para a estimação de tais modelos: a aproximação de modelos utilizada por Britton e Becker (2000 apud O'NEILL, 2010), a utilização de dados aumentados e métodos MCMC como em O'Neill e Roberts (1999), martingais em Becker (1989 apud O'NEILL, 2010) e Computação Bayesiana Aproximada (ABC) que surgiu recentemente e não exige a utilização de função de verossimilhança como, por exemplo, em Tony et al. (2009 apud O'NEILL, 2010)).

Esta dissertação utiliza a abordagem de dados aumentados em conjunto com os métodos MCMC. Höhle e Jørgensen (2002) discutem as estimativas para um modelo SIR considerando arranjos espaciais aplicado aos dados de uma clássica epidemia de febre Suína na Holanda em 1997 discutida, inicialmente, por Pluimers et al. (1999 apud HöHLE; JøRGENSEN, 2002).

Stretfaris e Gibson (2004) assumem uma distribuição Weibull para o tempo de remoção em um modelo SEIR, Lekone e Finkenstädt (2006) consideram os dados incompletos e as medidas de intervenção introduzidas na epidemia de febre hemorrágica Ebola considerada anteriormente por Chowell et al. (2004). Por último, Boys e Giles (2007) consideram modelos SEIR com período de exposição constante e período de remoção não homogêneo com um número aleatório de pontos de mudança através do algoritmo *Reversible Jump* e aplicam tal estrutura em dados de varíola e de uma doença respiratória.

1.3 Febre hemorrágica Ebola

A Febre hemorrágica Ebola é uma doença, frequentemente fatal, presente na população humana e não-humana (macacos, gorilas, chimpanzés, antílopes e porco-espinhos) e surge esporadicamente desde seu reconhecimento inicial em 1976. A doença é causada pelo vírus Ebola, nome de um rio na República Democrática do Congo, na África, local em que o vírus foi encontrado pela primeira vez.

Há cinco subtipos de vírus Ebola sendo que quatro deles atingem a população humana: Ebola-Zaire, Ebola-Sudan, Ebola-Ivory Coast e Ebola-Bundibugyo. O quinto, Ebola-Reston, atinge, principalmente a população não-humana primata e pode atingir a população humana, contudo, não gera sintomas sérios ou mortes.

O reservatório natural do vírus é desconhecido, o vírus é encontrado em animais hospedeiros nativos das florestas tropicais da África nos quatro sub-tipos que atingem a população humana e o quinto sub-tipo é encontrado em animais nativos das florestas tropicais do oeste do Pacífico.

Desta forma, acredita-se que a transmissão da doença ocorre quando um ser humano tem contato com um animal infectado pelo vírus. O vírus entre humanos é transmitido via contato físico com sangue ou secreções de pessoas infectadas. O período de exposição do vírus é de 2 a 21 dias com média de 6,3 dias e o período de infecção é estimado ser de 3,5 a 10,7 dias (BREMAN; PIOT; JOHNSON, 1977 apud LEKONE; FINKENSTäDT, 2006) sendo que o surgimento dos sintomas é abrupto e caracterizado por febre, dores de cabeça, dores musculares seguido por diárreia, vômito, conjuntivite, hemorragias internas e externas.

Segundo a Organização Mundial de Saúde, diferentes hipóteses tem sido desenvolvidas para explicar a origem das epidemias de Ebola. Observações laboratoriais identificaram que morcegos infectados experimentalmente não morreram e, desta forma, há especulações sobre o papel de animais mamíferos na manuntenção do vírus nas florestas tropicais e extensos estudos ecológicos estão em desenvolvimento para a identificação da reserva natural do vírus.

Não há um tratamento padrão para a doença, não há cura ou tratamentos efica-

zes. Os pacientes recebem cuidados básicos de suporte vital e devem ser isolados de qualquer contato. Há algumas pessoas que são capazes de apresentar recuperação, todavia, os pesquisadores ainda não foram capazes de identificar o fator responsável pela recuperação destes pacientes.

Não há casos de infecção humana fora do continente africano e o espalhamento da doença é facilitado pela falta de uso de roupas de proteção e material descartável pelo sistema de saúde em países da África.

As medidas de intervenção, em caso de confirmação de casos de Ebola, consistem na utilização de roupas de proteção tais como luvas, máscaras, isolamento dos pacientes suspeitos, acompanhamento de pessoas próximas ao caso suspeito, disponibilização de informações para a população da região afetada e o contato com os corpos daqueles que faleceram também é evitado.

Um mapa da distribuição da epidemias de Febre Hemorrágica Ebola ao longo dos anos é dado na Figura 1.

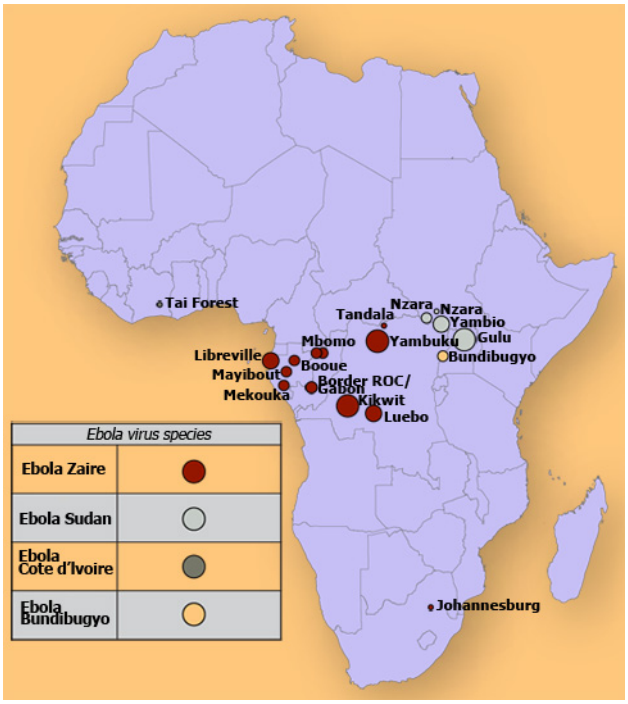


Figura 1: Distribuição geográfica das epidemias de Ebola segundo os quatro sub-tipos fatais para a população humana até o ano de 2009 - Fonte: Centro de Controle de Doenças e Prevenção dos E.U.A.

A Organização Mundial de Saúde fornece informações gerais de caráter histórico e preventivo e as medidas necessárias para lidar com a doença Febre Hemorrágica Ebola. O Centro de Controle de Doenças e Prevenção dos E.U.A. também é uma fonte válida de informações. Mais detalhes podem ser encontrados em WHO (2009), CDDP (2009).

1.4 O conjunto de dados

A epidemia de Ebola no Congo, em 1995, teve início na região de Bandundu, primeiramente em Kikwit, sendo o primeiro caso identificado em 6 de janeiro de 1995 envolvendo um homem de 43 anos de idade que faleceu uma semana depois. Somente no dia 9 de Maio de 1995, o vírus Ebola foi confirmado como agente causador da morte dos casos suspeitos e medidas de intervenção foram introduzidas.

O número total de casos confirmados foi de 316 pessoas, contudo, o período prévio à 1 de Março de 1995 não foi observado e desta forma, 25 indíviduos suscetíveis que se tornaram expostos e 80 indivíduos infectantes que se tornaram removidos não foram registradas em data exata e, por isso, não são considerados no conjunto de dados.

O conjunto de dados consiste em três séries temporais diárias registradas no período de 01 de Março de 1995 à 16 de Julho de 1995. A primeira série, denotada por **B**, é o número de de indivíduos suscetíveis que se tornaram expostos e somente apresenta a informação sobre o primeiro caso; a segunda série, denotada por **C**, é o número de indivíduos expostos que se tornaram infectantes, contabilizando um número total de 291 indivíduos, por fim, a terceira série, denotada por **D**, é o número de indivíduos infectantes que se tornaram removidos em um total de 236 indíviduos. (KHAN et al., 1999). As séries de dados são apresentadas na Figura 2.

1.5 Objetivos

É possível estabelecer os objetivos desta dissertação mais claramente:

- Discutir em detalhes os resultados de Lekone e Finkenstädt (2006) a fim de analisar como a incerteza gerada pelos dados parcialmente observados é incorporada

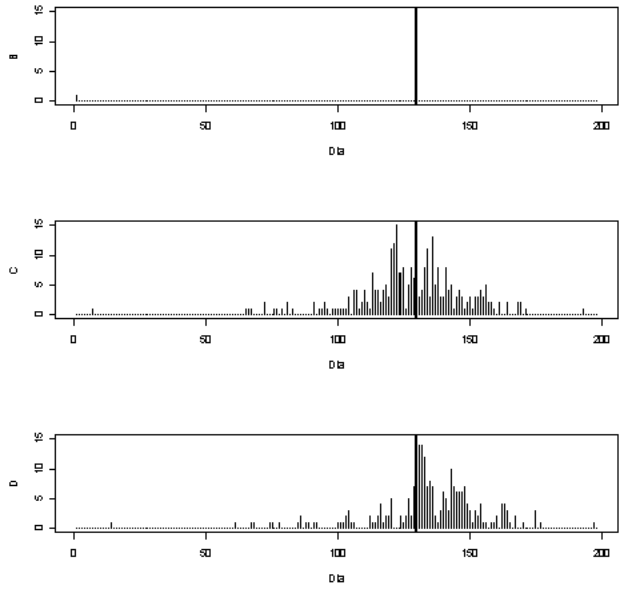


Figura 2: Dados da epidemia de Febre Hemorrágica Ebola ocorrida no Congo em 1995 com medidas de intervenção introduzidas no dia 130 (barra vertical em negrito). A série de dados **B** é o número de de indivíduos suscetíveis que se tornam expostos e somente apresenta a informação sobre primeiro caso; a série de dados **C** é o número de indivíduos expostos que se tornam infectantes, contabilizando um número total de 291 novos casos; a série de dados **D** é o número de indivíduos infectantes que são removidos em um total de 236 remoções

pelo método de estimação segundo a abordagem bayesiana;

- Modificar o modelo de Lekone e Finkenstädt (2006) ao considerar uma taxa de remoção não homogênea no tempo a fim de descrever de maneira mais realística o conjunto de dados da epidemia de Ebola ocorrida no Congo em 1995;
- Analisar o impacto das distribuições a priori nas modelagens propostas;
- Comparar os modelos discutidos.

1.6 Estrutura da dissertação

A dissertação é organizada da seguinte forma: O capítulo 2 consiste na apresentação do ferramental estatístico que consiste em inferência bayesiana, análise de sobrevivência, cadeias de Markov e algoritmos MCMC. O capítulo 3 discute os resultados de Lekone e Finkenstädt (2006) que modelam o conjunto de dados considerado através do modelo SEIR segundo uma abordagem bayesiana e considera a presença de dados faltantes e medidas de intervenção na taxa de contato.

O capítulo 4 apresenta uma modificação do modelo de Lekone e Finkenstädt (2006) ao considerar a taxa de remoção não homogênea no tempo devido às medidas de intervenção, já que um modelo que permita descrever o impacto das medidas de intervenção mais precisamente é essencial para a quantificação das consequências das medidas de intervenção introduzidas em uma epidemia. Por fim, no capítulo 5, os modelos considerados são aplicado ao conjunto de dados reais.

Cabe ressaltar que os gráficos referentes aos métodos MCMC para os Capítulos 3, 4 e 5 estão localizados nos apêndices A, B e C, respectivamente.

2 *Metodologia*

Este capítulo tem por objetivo apresentar uma discussão sobre as ferramentas estatísticas a serem utilizadas na dissertação. Portanto, o texto não irá justificar os resultados teóricos expostos, porém, indicará diversas referências para o leitor interessado em maiores detalhes.

Na seção 2.1, há uma pequena introdução sobre inferência bayesiana; na seção 2.2, sobre análise de sobrevivência; na seção 2.3, sobre cadeias de Markov. Na seção 2.4, há uma extensa discussão sobre algoritmos MCMC dando atenção ao algoritmo Amostrador de Gibbs na seção 2.4.1, o algoritmo Metropolis-Hastings na seção 2.4.2, a utilização conjunta destes dois algoritmos na seção 2.4.3, detalhes da implementação na seção 2.4.4, uma pequena explanação sobre os métodos de diagnóstico de convergência na seção 2.4.5, um pouco sobre a utilização de dados aumentados na seção 2.4.6 e por fim, a seção 2.4.7 estabelece uma direção para a análise dos resultados provenientes dos procedimentos MCMC a serem utilizados nesta dissertação.

2.1 Inferência bayesiana

A inferência bayesiana é uma abordagem alternativa e complementar a inferência clássica. A literatura sobre o assunto é vasta, sendo que bons textos são encontrados em Gamerman e Lopes (2006, Capítulo 2), Paulino, Turkaman e Murteira (2003), Mignon e Gamerman (1999) como textos basilares e Bernardo e Smith (1994) para um texto com uma abordagem mais complexa.

As duas abordagens tomam como base um conjunto de observações x que são descritos por uma função de densidade ou probabilidade $f(x|\boldsymbol{\theta})$, tal que o vetor de parâmetros

$\boldsymbol{\theta}$ define completamente esta distribuição e, portanto, o objetivo é obter informação sobre tal quantidade.

Além da informação dada pelo conjunto de dados, usualmente, há informação disponibilizada pelo pesquisador da área e é neste ponto que a inferência bayesiana difere da inferência clássica. A abordagem bayesiana admite a utilização desta informação proveniente do pesquisador, e a incorpora através de uma distribuição para o vetor de parâmetro $\boldsymbol{\theta}$ denonimada distribuição a priori, denotada por $p(\boldsymbol{\theta})$.

Desta forma, a partir da função de verossimilhança $L(\boldsymbol{\theta}|x)$ e a distribuição a priori $p(\boldsymbol{\theta})$ é possível obter, a partir do Teorema de Bayes, a distribuição a posteriori, isto é,

$$p(\theta|\boldsymbol{\theta}) \quad = \quad \frac{L(\boldsymbol{\theta}|x)p(\boldsymbol{\theta})}{f(x)}, \tag{2.1}$$

em que,

$$f(x) \quad = \quad \int L(\boldsymbol{\theta}|x)p(\boldsymbol{\theta})d\boldsymbol{\theta}. \tag{2.2}$$

Note que as observações x e a densidade $f(x)$ são constantes já que (2.1) é uma densidade em $\boldsymbol{\theta}$. Então, é possível reescrever (2.1) da forma,

$$\pi(\boldsymbol{\theta}) \quad \propto \quad L(\boldsymbol{\theta}|x)p(\boldsymbol{\theta}), \tag{2.3}$$

sendo que, por facilidade de notação, $\pi(\boldsymbol{\theta}) = p(\boldsymbol{\theta}|x)$.

O primeiro elemento na abordagem bayesiana que merece atenção é a escolha da distribuição a priori $p(\boldsymbol{\theta})$ que é uma questão frequentemente discutida, principalmente, quando se buscam distribuições a priori não-informativas. A escolha de distribuições a priori não-informativas pode levar a definição de distribuições impróprias, isto é,

$$\int_{-\infty}^{\infty} p(\boldsymbol{\theta})d\boldsymbol{\theta} \neq 1. \tag{2.4}$$

Na literatura, há algumas definições de distribuições a priori não-informativas comumente aceitas que foram apresentadas por Jeffrey (1961 apud GAMERMAN; LOPES, 2006, Capítulo 2) e Bernardo (1979 apud GAMERMAN; LOPES, 2006, Capítulo 2). De maneira geral, apesar da escolha de densidades a priori impróprias, as densidades a

posteriori podem ser próprias, todavia, há exceções e não é trivial verificar se uma densidade a posteriori é própria dependendo da complexidade do modelo.

Um segundo elemento é a predição de novas observações Y construída a partir da distribuição de $Y|X$. Se Y e X são condicionalmente independentes dado $\boldsymbol{\theta}$, então,

$$\begin{aligned} f(y|x) &= \int f(y, \boldsymbol{\theta}|x) d\boldsymbol{\theta} \\ &= \int f(y|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}. \end{aligned} \tag{2.5}$$

A utilização de distribuições preditivas é extremamente interessante pela possibilidade de comparar os resultados dados pelos métodos estatísticos com os dados reais. Há toda uma discussão sobre distribuições preditivas em Aitchinson e Dunsmore (1975 apud GAMERMAN; LOPES, 2006, Capítulo 2), Geisser (1993 apud GAMERMAN; LOPES, 2006, Capítulo 2) e por fim Gelman, Meng e Stern (1996 apud LEKONE; FINKENSTäDT, 2006, Capítulo 2).

Todavia, o interesse principal ainda está sobre $\boldsymbol{\theta}$ e sua distribuição a posteriori $\pi(\boldsymbol{\theta})$. Para sumarizar a informação contida em $\pi(\boldsymbol{\theta})$, usualmente, são utilizadas estatísticas descritivas como média, mediana e moda como principais medidas de locação e desvio padrão, desvio interquartílico, precisão e curvatura sobre a moda como principais medidas de dispersão.

Tais quantidades são facilmente calculadas se a distribuição a posteriori pertence a uma família de distribuições conhecida. E para garantir que $\pi(\boldsymbol{\theta})$ pertença a uma família de distribuições conhecida, pode-se utilizar o artifício de famílias de distribuições conjugadas.

A definição de famílias de distribuições conjugadas é que uma família de distribuições P é conjugada a um modelo F se para toda distribuição a priori $p \in P$ e para qualquer observação $f \in F$, a distribuição a posterior $\pi \in P$. (GAMERMAN; LOPES, 2006, Capítulo 2)

Tal artifício é extremamente útil no contexto univariado para cálculos analíticos e no contexto multivariado para a aplicação do algoritmo Amostrador de Gibbs que será apresentado. Por outro lado, a utilização de famílias conjugadas limita a escolha de

distribuições a priori e, muitas vezes, a escolha neste conjunto de distribuições pode não ser razoável. Além disso, a existência de distribuições a priori que satisfazem tal definição limitam-se a modelos simples.

A utilização de distribuições a priori que não são conjugadas ao modelo utilizado pode tornar intratável analiticamente a distribuição a posteriori para θ e, por muito tempo, este foi um grande empecilho para o desenvolvimento de aplicações através da inferência bayesiana.

Os métodos MCMC e a evolução dos computadores permitiram e provocaram um grande impacto na utilização de inferência bayesiana por não exigir mais a utilização de famílias conjugadas. Atualmente, estas ferramentas são largamente utilizadas e também são foco de pesquisa.

2.2 Análise de sobrevivência

A teoria de análise de sobrevivência é utilizada para lidar com modelos SEIR por diversos autores como O'Neill e Roberts (1999), Höhle e Jørgensen (2002), Stretfaris e Gibson (2004) e Boys e Giles (2007) que consideram tempo contínuo na modelagem.

Lekone e Finkenstädt (2006) utilizam também a teoria de análise de sobrevivência, mas em menor grau por considerarem uma modelagem em tempo discreto. Desta forma, para um melhor entendimento, alguns conceitos básicos serão apresentados a partir de Colosimo e Giolo (2006, Capítulos 1, 3).

Na teoria de análise de sobrevivência, a variável resposta é o tempo até a ocorrência de um evento de interesse ou falha, este tempo é denominado tempo de falha podendo ser o tempo até que um indivíduo suscetível torne-se exposto ou um indivíduo exposto torne-se infectante ou um indivíduo infectante torne-se removido, isto é, os eventos de interesse são, respectivamente, tornar-se exposto, tornar-se infectante e tornar-se removido.

Uma problemática de dados de sobrevivência é a presença de censura nos dados que consiste na observação parcial do tempo de falha de um indivíduo por outras razões que não seja o evento de interesse. No contexto de modelos SEIR, não há problemas

de censura, porém, existe o problema de falta de informações.

A variável tempo de falha T é não-negativa e, em geral, contínua e para tratá-la, usualmente, são utilizadas a função de sobrevivência e a função taxa de falha ou risco.

A função de sobrevivência é definida como a probabilidade de uma observação não falhar até um certo tempo t, ou seja, uma observação sobreviver ao tempo t. Em termos probabilísticos, é dada por,

$$S(t) = P(T > t). \tag{2.6}$$

Em consequência disso, a função de distribuição acumulada é definida como a probabilidade de uma observação não sobreviver ao tempo t, isto é,

$$F(t) = 1 - S(t). \tag{2.7}$$

E a partir da função de sobrevivência, é possível definir a probabilidade da ocorrência de uma falha no intervalo $(t, t + \Delta t]$ que é dada por, $S(t) - S(t + \Delta t)$.

A taxa de falha neste mesmo intervalo é definida como a probabilidade de ocorrência da falha dado que a falha não ocorreu antes de t, dividida pelo comprimento do intervalo, isto é,

$$\frac{S(t) - S(t + \Delta t)}{\Delta t S(t)}. \tag{2.8}$$

A função taxa de falha ou risco $\lambda(t)$ representa a taxa de falha instantânea no tempo t condicional à sobrevivência até o tempo t. Em termos probabilísticos,

$$\begin{aligned} \lambda(t) &= lim_{\Delta \to 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} \\ &= lim_{\Delta \to 0} \frac{S(t) - S(t + \Delta t)}{\Delta t S(t)}. \end{aligned} \tag{2.9}$$

Observe que a função taxa de falha é sempre positiva, mas não apresenta limite superior. Além disso, ela é mais informativa graficamente do que a função de sobrevivência, dado que diferentes funções de sobrevivência podem ter formas semelhantes, enquanto as respectivas funções de taxa de falha são bastante distintas.

E também pode-se definir a função taxa de falha acumulada,

$$\Lambda(t) = \int_0^t \lambda(u)du. \tag{2.10}$$

Esta função não apresenta uma interpretação direta como a função taxa de falha, porém é útil em manipulações algébricas.

Por fim, algumas relações importantes entre tais funções são dadas abaixo,

$$\begin{aligned}
\lambda(t) &= \frac{f(t)}{S(t)} && (2.11)\\
&= -\frac{d}{dt}(logS(t)), && (2.12)\\
\Lambda(t) &= \int_0^t \lambda(u)du && (2.13)\\
&= -logS(t), && (2.14)\\
S(t) &= exp\{-\Lambda(t)\} && (2.15)\\
&= exp\left\{\int_0^t \lambda(u)du\right\}. && (2.16)
\end{aligned}$$

Nesta dissertação, a utilização de análise de sobrevivência consiste na hipótese em relação ao tempo de permanência dos indivíduos nos compartimentos dos modelos SEIR. Então, T é o tempo de permanência de um indivíduo em qualquer compartimento. E quando T segue distribuição Exponencial, segue que,

$$\begin{aligned}
S(t) &= e^{-\lambda t}, && (2.17)\\
\lambda(t) &= \lambda, && (2.18)\\
\Lambda(t) &= \lambda t. && (2.19)
\end{aligned}$$

Note que a função taxa de falha é constante e, por isso, o risco de um indivíduo mudar de um compartimento para outro é igual para todos os indivíduos sem distinção pelo tempo em que cada indivíduo já permaneceu no compartimento. Além disso, o risco de mudança entre compartimentos é constante por todo o período da epidemia. Estas

duas implicações ao assumir a distribuição Exponencial não são realísticas, todavia, facilitam a modelagem do problema.

Um dos objetivos desta dissertação, como já dito, é modificar a segunda implicação assim como já é feito por Lekone e Finkenstädt (2006) e, mais elaboradamente, por Boys e Giles (2007). Eles consideram a distribuição Exponencial por Partes ($\lambda(t)$) para a transição entre alguns compartimentos.

A distribuição Exponencial por Partes ($\lambda(t)$) é baseada na distribuição Exponencial usual, porém, o parâmetro $\lambda(t)$ é função do tempo,

$$\lambda(t) = \begin{cases} \lambda_1 & \text{se} \quad t_0 \leq t < t_1 \\ \lambda_2 & \text{se} \quad t_1 \leq t < t_2 \\ \vdots & \\ \lambda_k & \text{se} \quad t_{k-1} \leq t < t_k \end{cases} \tag{2.20}$$

para algum conjunto $\{t_0, \ldots, t_k\}$.

Stretfaris e Gibson (2004), Jewell et al. (2009) utilizam outras distribuições no lugar da distribuição Exponencial como, por exemplo, Weibull e Gumbel.

2.3 Cadeias de Markov

A discussão sobre cadeias de Markov é necessária para que seja possível compreender, mesmo intuitivamente, o funcionamento dos métodos computacionais a serem apresentados. Há diversas bibliografias esclarecedoras sobre o assunto: Hoel, Port e Stone (1972), Robert e Casella (2004), Ross (2007) e por último, Meyn e Tweedie (2009) para uma abordagem probabilística e Guttorp (1995 apud GAMERMAN; LOPES, 2006, Capítulo 3) para uma abordagem estatística.

O conceito de cadeias de Markov é atribuído ao matemático russo Andrei Andreivich Markov que estudou, no início do século 20, a alternância entre vogais e consoantes do poema *Onegyn* escrito por Poeshkin. Ele desenvolveu um modelo cujo sucessivos resultados dependiam dos resultadores anteriores somente através do último resultado

antecessor e a partir deste modelo, obteve boas estimativas das frequências de vogais e consoantes. Concomitantemente, o matemático francês Henri Poincaré estudou sequências de variáveis aleatórias que eram cadeias de Markov.

Uma cadeia de Markov é um processo estocástico,

$$X = \{X_t, t \in T\},$$

tal que T representa o tempo da cadeia de Markov podendo ser discreto ou contínuo e X_t pertence ao espaço de estados I que também pode ser discreto ou contínuo.

A teoria sobre cadeias de Markov em tempo discreto para espaço de estados também discreto é bem desenvolvida e pode ser encontrada em qualquer um dos livros citados acima com especial atenção a Hoel, Port e Stone (1972). Para cadeias de Markov em tempo discreto para espaço de estados contínuo, a sugestão é Robert e Casella (2004, Capítulo 4) e por último, Ross (2007, Capítulo 4) é uma boa referência para cadeias de Markov em tempo contínuo para espaço de estados discreto. Meyn e Tweedie (2009) abordam todos estes casos além de descreverem cadeias de Markov em tempo e espaço de estados contínuo, todavia, em um nível muito acima do introdutório.

Este texto segue Robert e Casella (2004, Capítulo 6) em sua discussão dos conceitos essenciais sobre cadeias de Markov para a utilização dos algoritmos MCMC e, por isso, limita-se aos resultados de cadeias de Markov em tempo discreto devido à natureza dos algoritmos de simulação.

Como dito, uma cadeia de Markov é um processo estocástico e este é definido a partir do seu *kernel* de transição dado por $K(.;.)$ que representa uma probabilidade condicional ao último resultado.

Desta forma, um processo estocástico é dito ser uma cadeia de Markov se seu *kernel* de transição $K(.;.)$ satisfazer a propriedade de Markov que consiste na afirmação de que $K(X_0, \ldots, X_{t-1}; X_t) = K(X_{t-1}; X_t)$. Além disso, o *kernel* de transição em z passos é definido por $K^{(z)}(X_{t-1}; X_t) = K(X_{t-z}; X_t)$.

As propriedades limites de uma cadeia de Markov são definidas a partir de seu *kernel* de transição e estas propriedades são de grande interesse para a aplicação dos

métodos MCMC. Neste sentido, a distribuição limite de uma cadeia de Markov é dada por,

$$\pi_\infty = \lim_{z\to\infty} K^{(z)}(X_{t-1}; X_t), \tag{2.21}$$

e sob certas condições, a distribuição limite π_∞ é igual à π que é denominada distribuição estacionária da cadeia de Markov, sendo que uma distribuição π é dita ser uma distribuição estacionária se $X_{t-1} \sim \pi$, então $X_t \sim \pi$.

A simulação da distribuição estacionária é o ponto central dos algoritmos MCMC. Para tanto, estes são construídos a fim de estabelecer cadeias de Markov com *kernels* de transição que satisfaçam as condições conhecidas por irredutibilidade, aperiodicidade e recorrência Harris. Sob estas condições, para z suficientemente grande, a distribuição dos elementos da cadeia de Markov seguem uma distribuição aproximadamente igual à distribuição estacionária, ou seja, a distribuição limite é igual à distribuição estacionária.

É válido notar que a condição de recorrência Harris é difícil de ser verificada e por isso a demonstração da validade dos algoritmos MCMC utilizam a propriedade de reversibilidade. Uma cadeia de Markov é reversível se,

$$\pi(x)K(x;y) = \pi(y)K(y;x) \quad \forall \quad x, y \in I. \tag{2.22}$$

Se a cadeia de Markov apresenta as propriedades de irredutibilidade, aperiodicidade e reversibilidade para π, então π é uma distribuição estacionária e também é a distribuição limite.

Finalmente, pode-se estabelecer o resultado mais interessante para os métodos MCMC cuja demonstração pode ser encontrada em Robert e Casella (2004, Capítulo 6).

Teorema 2.3.1 (Teorema Ergódico) *Se uma cadeia de Markov é irredutível, Harris recorrente e aperiódica, então,*

$$\bar{T}_z = \frac{1}{z}\sum_{t=1}^{z} T(X_t) \to E_\pi(T(X)) \quad \textit{quando} \quad z \to \infty.$$

A partir destes resultados, é possível discutir os algoritmos MCMC.

2.4 Algoritmos MCMC

Os métodos MCMC surgem como solução frente ao problema de lidar com distribuições de interesse analiticamente complexas e/ou difíceis de serem simuladas por métodos usuais. E eles são fortemente relacionados à inferência bayesiana por não exigir, na maioria dos casos, o conhecimento da constante de normalização das densidades a serem simuladas. Todavia, também, podem ser utilizados na inferência clássica. (ROBERT; CASELLA, 2004)

A literatura sobre o assunto é extremamente vasta, mas este texto segue as linhas gerais de (GAMERMAN; LOPES, 2006) pela clareza na exposição das idéias sobre MCMC. A discussão está inserida no contexto bayesiano e portanto, a distribuição de interesse é a distribuição a posteriori.

2.4.1 Amostrador de Gibbs

O Amostrador de Gibbs é o método MCMC mais simples de ser aplicado e teve um grande impacto na inferência bayesiana desde que seu potencial em diversas aplicações foi demonstrado por Gelfand e Smith (1990 apud CASELLA; GEORGE, 1992).

O algoritmo foi introduzido por Geman e Geman (1984 apud GAMERMAN; LOPES, 2006, Capítulo 5) no contexto de processamento de imagens e o esquema de amostragem apresentado explorava a estrutura condicional do problema dada pela distribuição de Gibbs, que dá o nome ao algoritmo. Gelfand e Smith (1990 apud GAMERMAN; LOPES, 2006, Capítulo 5) apontaram que o esquema de amostragem poderia ser utilizado para qualquer outra distribuição condicional além da distribuição de Gibbs.

A bibliografia sobre o Amostrador de Gibbs é extensa e a descrição do algoritmo pode ser encontrada em qualquer livro que trate de inferência bayesiana com uma pequena abordagem computacional ou de métodos computacionais para estatística. As sugestões de leituras para um melhor entendimento do algoritmo são Casella e George (1992) e Gamerman e Lopes (2006, Capítulo 5).

A idéia central do algoritmo é a simulação da densidade de interesse π através da simulação das densidades condicionais completas. Para tanto, seja $\boldsymbol{\theta} = (\boldsymbol{\theta_1}, \ldots, \boldsymbol{\theta_d})$ o vetor de parâmetros de interesse com distribuição conjunta dada por $\pi(\boldsymbol{\theta})$ e a densidade condicional completa de $\boldsymbol{\theta_i}$ é definida como $\pi(\boldsymbol{\theta_i}|\boldsymbol{\theta_1}, \ldots, \boldsymbol{\theta_{i-1}}, \boldsymbol{\theta_{i+1}}, \ldots, \boldsymbol{\theta_d})$ que usualmente é denotada por $\pi(\boldsymbol{\theta_i}|\boldsymbol{\theta_{-i}})$ para $i = 1, \ldots, d$. Então, o algoritmo é dado por,

1. Inicialize o contador da cadeia de Markov $j = 1$ e dê um conjunto de valores iniciais $\boldsymbol{\theta^{(0)}} = (\boldsymbol{\theta_1^{(0)}}, \ldots, \boldsymbol{\theta_d^{(0)}})$;

2. Obtenha um novo valor $\boldsymbol{\theta^{(j)}} = (\boldsymbol{\theta_1^{(j)}}, \ldots, \boldsymbol{\theta_d^{(j)}})$ a partir do valor anterior $\boldsymbol{\theta^{(j-1)}} = (\boldsymbol{\theta_1^{(j-1)}}, \ldots, \boldsymbol{\theta_d^{(j-1)}})$ através da geração dos valores:

   (a) $\boldsymbol{\theta_1^{(j)}} \sim \pi(\boldsymbol{\theta_1}|\boldsymbol{\theta_{-1}})$;

   (b) $\boldsymbol{\theta_2^{(j)}} \sim \pi(\boldsymbol{\theta_2}|\boldsymbol{\theta_{-2}})$;

   $\vdots$

   (c) $\boldsymbol{\theta_d^{(j)}} \sim \pi(\boldsymbol{\theta_d}|\boldsymbol{\theta_{-d}})$;

3. Mude o contador de j para $j + 1$ e retorne ao passo 2 até que a convergência seja alcançada.

Desta forma, uma cadeia de Markov d-variada é formada sendo que $\pi(\theta)$ é a distribuição estacionária e o *kernel* de transição $K(.;.)$ é construído a partir das densidades condicionais completas. E após a convergência ter sido obtida, o elemento gerado apresentará distribuição aproximadamente igual à distribuição de interesse.

Note que $\boldsymbol{\theta_i}$ não é necessariamente um escalar e, em princípio, é necessário saber gerar as densidades condicionais completas de maneira direta. Além disso, o Amostrador de Gibbs pode ser visto como um caso particular de um dos algoritmos a ser apresentado à frente. (GELMAN, 1992 apud CHIB; GREENBERG, 1995)

2.4.2 Metropolis-Hastings

O algoritmo de Metropolis-Hastings foi introduzido por Metropolis et al. (1953 apud CHIB; GREENBERG, 1995) no contexto da Física e generalizado por Hastings (1970 apud CHIB; GREENBERG, 1995).

A idéia central baseia-se no princípio dos algoritmos de rejeição, isto é, como é difícil gerar elementos de uma distribuição de interesse π, considera-se uma densidade proposta que gere candidatos que podem ou não ser aceitos como pertencentes a π. Todavia, como o algoritmo também é baseado em cadeias de Markov, os candidatos gerados podem ou não depender do valor anterior e a aceitação do candidato sempre apresenta tal dependência diferentemente dos algoritmos usuais de rejeição. (CHIB; GREENBERG, 1995)

A estrutura do algoritmo é bastante genérica, o que possibilita acomodar diversas espécies de problemas, ao mesmo tempo, ocasiona a dificuldade de encontrar resultados gerais ou ótimos para a sua implementação e, portanto, o algoritmo não é tão direto de ser aplicado como o Amostrador de Gibbs. Além disso, normalmente, exige um tempo computacional superior.

A bibliografia sobre o algoritmo de Metropolis-Hastings é ainda mais vasta do que sobre o Amostrador de Gibbs e se pode encontrar inúmeras particularizações que tornam o algoritmo mais eficiente.

O conceito de eficiência, neste texto, é no sentido de exigência de tempo computacional para que seja obtida a amostra desejada, o que envolve a capacidade do algoritmo obter uma amostra representativa da distribuição de interesse e próxima de uma amostra aleatóra.

Deste modo, para avaliar a eficiência dos algoritmos, avalia-se a convergência da média amostral para as quantidades de interesse, a autocorrelação entre os elementos da cadeia e por fim, a taxa de aceitação, sendo que tais pontos são discutidos mais à frente com mais detalhes. Roberts e Rosenthal (2001) apresentam uma discussão mais formal sobre eficiência em cadeias de Markov.

Como textos em nível introdutório, pode-se citar Gamerman e Lopes (2006, Capítulo 6) e Ntzoufras (2009, Capítulo 2) enquanto Robert e Casella (2004, Capítulo 6) apresentam uma discussão mais complexa. Para justificar a operação do algoritmo, Chib e Greenberg (1995) é uma ótima referência e por fim, Gilks, Richardson e Spiegelhalter (1996) apresentam uma coletânea de diversas discussões sobre a utilização dos algoritmos MCMC na prática.

Para expor o algoritmo, seja $\boldsymbol{\theta} = (\boldsymbol{\theta_1}, \ldots, \boldsymbol{\theta_d})$ o vetor de parâmetros de interesse com distribuição conjunta dada por $\pi(\boldsymbol{\theta})$ e $q(.;.)$ é a densidade proposta, então

1. Inicialize o contador $j = 1$ e considere um conjunto de valores iniciais $\boldsymbol{\theta^{(0)}}$;
2. Gere um candidato $\boldsymbol{\theta'}$ a partir da densidade $q(\boldsymbol{\theta^{(j-1)}}; \boldsymbol{\theta'})$;
3. Gere $U \sim U(0, 1)$;
4. Calcule $\alpha(\boldsymbol{\theta^{(j-1)}}, \boldsymbol{\theta'})$ dado por

$$\alpha(\boldsymbol{\theta^{(j-1)}}, \boldsymbol{\theta'}) = \min\left\{1, \frac{\pi(\boldsymbol{\theta'})q(\boldsymbol{\theta'}; \boldsymbol{\theta^{(j-1)}})}{\pi(\boldsymbol{\theta^{(j-1)}})q(\boldsymbol{\theta^{(j-1)}}; \boldsymbol{\theta'})}\right\}; \tag{2.23}$$

5. Se $u < \alpha(\boldsymbol{\theta^{(j-1)}}, \boldsymbol{\theta'})$, faça $\boldsymbol{\theta^{(j)}} = \boldsymbol{\theta'}$. Caso contrário, $\boldsymbol{\theta^{(j)}} = \boldsymbol{\theta'}$;
6. Mude o contador de j para $j + 1$ e retorne ao passo 2 até que a convergência seja alcançada.

Novamente, uma cadeia de Markov d-variada é formada sendo que $\pi(\boldsymbol{\theta})$ é a distribuição estacionária e o *kernel* de transição $K(.;.)$ é construído a partir de $q(.;.)$ e $\alpha(.;.)$. E após a convergência ter sido obtida, o elemento gerado apresentará distribuição aproximadamente igual à distribuição de interesse.

Observe que a razão de densidades definida em (2.23) implica que não seja necessário conhecer a constante de normalização da distribuição de interesse, o que torna o algoritmo extremamente atraente para aplicações na inferência bayesiana e aplicável a diversas espécies de problema.

Todavia, como todo algoritmo de rejeição, há a necessidade da escolha da densidade proposta $q(.;.)$ e esta escolha não é trivial. O primeiro critério desta escolha é que a densidade $q(.;.)$ seja fácil de ser gerada e avaliada, o segundo critério é que o suporte da densidade proposta contenha o suporte da distribuição de interesse e por fim, a escolha deve proporcionar uma amostragem eficiente de π.

Chib e Greenberg (1995) apresentam algumas possíveis escolhas e dentre as mais usuais estão aquelas conhecidas por densidades propostas de passeio aleatório e de independência que são discutidas nas próximas seções.

2.4.2.1 Passeio aleatório

A densidade proposta de passeio aleatório consiste em gerar um novo candidato a partir do último valor dado pela cadeia de Markov, isto é, $\boldsymbol{\theta'} = \boldsymbol{\theta^{(j-1)}} + \epsilon_j$ dado que ϵ_j é uma variável aleatória com distribuição independente da cadeia de Markov.

Em geral, ϵ_j são variáveis aleatórias independentes e idênticamente distribuídas por f_ϵ sendo f_ϵ simétrica em torno de 0. Então $q(\boldsymbol{\theta^{(j-1)}}; \boldsymbol{\theta'}) = f_\epsilon(\boldsymbol{\theta'} - \boldsymbol{\theta^{(j-1)}})$ e a razão dada em (2.23) pode ser simplificada para

$$\alpha(\boldsymbol{\theta^{(j-1)}}, \boldsymbol{\theta'}) = \min\left\{1, \frac{\pi(\boldsymbol{\theta'})}{\pi(\boldsymbol{\theta^{(j-1)}})}\right\}. \tag{2.24}$$

A escolha de densidade proposta de passeio aleatório é comum e consiste em explorar localmente a vizinhaça do valor $\boldsymbol{\theta^{(j-1)}}$ na distribuição de interesse (ROBERT; CASELLA, 2004). As distribuições f_ϵ mais usuais são a distribuição Normal segundo Muller (1991 apud GAMERMAN; LOPES, 2006, Capítulo 6) e t-Student segundo Geweke (1992 apud GAMERMAN; LOPES, 2006, Capítulo 6) centradas na origem.

Como a locação é definida pelo último valor dado pela cadeia de Markov, o próximo passo é definir a variância de tais distribuições e é neste ponto que surge uma vasta literatura, pois tal escolha está estritamente ligada à eficiência do algoritmo e a grande utilização de tal densidade proposta.

Tierney (1994 apud GAMERMAN; LOPES, 2006, Capítulo 6) sugere a utilização de cV em que c é um escalar e V é alguma forma de aproximar a variância da distribuição de interesse, por exemplo, a informação de Fisher e, consequentemente, o objetivo é somente calibrar c.

A escolha de c deve ser tal que a variância da densidade proposta não seja muito pequena de tal forma que o algoritmo explore muito vagarosamente o suporte de π e, deste modo, seja obtida uma alta taxa de aceitação, mas uma grande autocorrelação entre os elementos da cadeia de Markov.

Por outro lado, a variância não pode ser muito grande, caso contrário, os candidatos irão pertencer às caudas da distribuição de interesse, consequentemente, a taxa de

aceitação dos candidatos será baixa, o suporte de π não será visitado rapidamente e a autocorrelação entre os elementos da cadeia de Markov será expressiva.

Roberts e Rosenthal (2001), Neal e Roberts (2006) revisam alguns resultados sobre a escolha ótima da constante multiplicativa da variância e apresentam outros resultados sobre a questão. Além disso, há diversas outras discussões sendo que todas são concordantes que a taxa de aceitação deve ser em torno de 20% para $d > 5$ até 50% para $d = 1$, tal que d é a dimensão do vetor $\boldsymbol{\theta}$.

Um resultado bastante interessante é dado por Roberts, Gelman e Gilks (1997 apud ROBERTS; ROSENTHAL, 2001) que afirmam que para uma densidade de interesse que é o produto de densidades de variáveis aleatórias independentes e identicamente distribuídas, a taxa de aceitação ótima é de 23,4% e ela pode ser alcançada ao definir $c = \dfrac{2.38^2}{d}$.

Este resultado, em um primeiro momento, parece ser bastante restrito, mas é diretamente aplicado quando $\boldsymbol{\theta}$ é univariado e ainda é válido para suposições um pouco mais fracas como discutidas por Roberts e Rosenthal (2001) ou Neal e Roberts (2006).

2.4.2.2 Independência

A densidade proposta de independência consiste em gerar um candidato independentemente do último valor dado pela cadeia de Markov, isto é, $q(\boldsymbol{\theta^{(j-1)}}, \boldsymbol{\theta'}) = f(\boldsymbol{\theta'})$. Todavia, esta característica do candidato não invalida a propriedade de Markov da cadeia já que a aceitação dos candidatos ainda depende do valor anterior, como se vê na expressão (2.23).

Uma escolha popular como densidade proposta independente segundo Gamerman e Lopes (2006, Capítulo 6), é a utilização da distribuição a priori do parâmetro do qual se deseja gerar a distribuição de interesse. Desta forma, é possível simplicar a expressão (2.23) para,

$$\alpha(\boldsymbol{\theta^{(j-1)}}, \boldsymbol{\theta'}) = \min\left\{1, \frac{L(\boldsymbol{\theta'}|x)}{L(\boldsymbol{\theta^{(j-1)}}|x)}\right\}. \tag{2.25}$$

sendo que L é a função de verossimilhança.

Esta escolha deve ser feita com cuidado, pois se a distribuição a priori não for concordante com a verossimilhança, a cadeia de Markov será ineficiente, já que o candidato gerado pela priori implicará em um valor pequeno da função de verossimilhança e portanto, dificilmente será aceito.

Uma solução para tal inconveniente é incorporar a informação dada pela verossimilhança na densidade proposta de independência, esta informação permite construir densidades propostas mais adequadas.

Gamerman e Lopes (2006, Capítulo 6) indicam vários exemplos como Bennett, Racine-Poon e Wakefield (1996 apud GAMERMAN; LOPES, 2006, Capítulo 6), Chib e Greenberg (1994). Uma escolha usual é a utilização de uma densidade normal centrada na moda da distribuição de interesse e a variância dada pela informação de Fisher segundo Ntzoufras (2009, Capítulo 2). Lekone e Finkenstädt (2006), Erkanli (1994 apud NTZOUFRAS, 2009, Capítulo 2) são exemplos de tal escolha.

Tierney (1994 apud GAMERMAN; LOPES, 2006, Capítulo 6) sugere a utilização de densidades com caudas mais pesadas do que as caudas da distribuição de interesse e, por isso, a distribuição t-Student com poucos graus de liberdade é uma escolha preferível à distribuição normal.

Robert e Casella (2004, Capítulo 6) comparam com o algoritmo de Aceitação-Rejeição pela natureza da densidade proposta de independência e esta comparação permite justificar a afirmação de Gamerman e Lopes (2006) de que uma regra geral é que π e $q(.;.)$ sejam o mais similares possíveis e, consequentemente, a taxa de aceitação seja alta, diferentemente da densidade proposta de passeio aleatório. Todavia, também é necessário estar atento à autocorrelação decorrente de tal escolha segundo Ntzoufras (2009, Capítulo 2).

Por fim, Robert e Casella (2004, Capítulo 6) afirmam que a utilização da densidade proposta de independência é uma exploração global da densidade de interesse e por isso, pode não ser adequada para grandes dimensões de π.

2.4.3 Metropolis-Hastings em blocos

Na seção anterior, o algoritmo de Metropolis-Hastings foi descrito considerando-se um único bloco do vetor de interesse θ, contudo ele pode ser dividido em blocos menores e, ainda, é possível construir uma cadeia de Markov através do algoritmo Metropolis-Hastings similarmente ao procedimento dado pelo Amostrador de Gibbs. A literatura nomeia este algoritmo como Metropolis-Hastings por Blocos e também por Metropolis-Hastings dentro do Amostrador de Gibbs. (NTZOUFRAS, 2009, Capítulo 2)

A idéia do algoritmo é construir a cadeia de Markov a partir das distribuições condicionais completas dos blocos do vetor de parâmetros θ, sendo que em cada bloco é aplicado o algoritmo de Metropolis-Hastings já que as distribuições condicionais completas são desconhecidas. Então, o algoritmo é dado por,

1. Inicialize o contador da cadeia de Markov em $j = 1$ e dê um conjunto de valores iniciais $\boldsymbol{\theta^{(0)}} = (\boldsymbol{\theta_1^{(0)}}, \ldots, \boldsymbol{\theta_d^{(0)}})$;

2. Obtenha um novo valor $\boldsymbol{\theta^{(j)}} = (\boldsymbol{\theta_1^{(j)}}, \ldots, \boldsymbol{\theta_d^{(j)}})$ a partir do valor anterior $\boldsymbol{\theta^{(j-1)}} = (\boldsymbol{\theta_1^{(j-1)}}, \ldots, \boldsymbol{\theta_d^{(j-1)}})$ através da geração dos valores $\boldsymbol{\theta^{(j)}} \sim \pi(\boldsymbol{\theta_i}|\boldsymbol{\theta_{-i}})$ para $i = 1 \ldots, d$ seguindo os passos:

   (a) Gere um candidato $\boldsymbol{\theta_i}'$ a partir da densidade $q(\boldsymbol{\theta_i^{(j-1)}}; \boldsymbol{\theta_i}')$;

   (b) Gere $U \sim U(0, 1)$;

   (c) Calcule $\alpha(\boldsymbol{\theta_i^{(j-1)}}; \boldsymbol{\theta_i'})$ dado por,

$$\alpha(\boldsymbol{\theta_i^{(j-1)}}; \boldsymbol{\theta_i}) = \min\left\{1, \frac{\pi(\boldsymbol{\theta_i'}|\boldsymbol{\theta_{-i}})q_i(\boldsymbol{\theta_i'}; \boldsymbol{\theta_i^{(j-1)}})}{\pi(\boldsymbol{\theta_i^{(j-1)}}|\boldsymbol{\theta_{-i}})q_i(\boldsymbol{\theta_i^{(j-1)}}; \boldsymbol{\theta_i'})}\right\}; \tag{2.26}$$

   (d) Se $U < \alpha(\boldsymbol{\theta_i^{(j-1)}}; \boldsymbol{\theta_i'})$, faça $\boldsymbol{\theta_i^{(j)}} = \boldsymbol{\theta_i'}$. Caso contrário, $\boldsymbol{\theta_i^{(j)}} = \boldsymbol{\theta_i^{(j-1)}}$;

3. Mude o contador de j para $j + 1$ e retorne ao passo 2 até que a convergência seja alcançada.

Este é o algoritmo mais geral possível e a sua justificativa pode ser encontrada em Gamerman e Lopes (2006, Capítulo 6) ou Chib e Greenberg (1995). A nomenclatura Metropolis-Hastings dentro do Amostrador de Gibbs surge da semelhança com

o algoritmo de Gibbs em relação à amostragem através das distribuições condicionais completas e à possibilidade de que estas distribuições possam ser desconhecidas e conhecidas.

Para as distribuições condicionais completas cuja forma analítica seja conhecida, a melhor escolha da densidade proposta é a própria densidade condicional completa e portanto, a probabilidade de aceitação é $\alpha(\boldsymbol{\theta}_i^{(j-1)}, \boldsymbol{\theta}_i') = 1$, consequentemente, o candidato é gerado diretamente e é sempre aceito.

Para as distribuições condicionais completas com formas desconhecidas, aplica-se o algoritmo de Metropolis-Hastings usual. Além disso, o algoritmo de Metropolis-Hastings pode ser aplicado para gerar candidatos para distribuições condicionais completas conhecidas quando é difícil de gerar um elemento desta distribuição diretamente.

Por fim, é também possível ver o Amostrador de Gibbs como um caso particular deste algoritmo quando todas as distribuições condicionais completas podem ser geradas diretamente.

2.4.4 Implementação

Detalhes sobre a implementação dos algoritmos MCMC descritos nas seções anteriores podem ser encontrados em Ntzoufras (2009), Gamerman e Lopes (2006) e Gelman (2004).

Além disso, Ntzoufras (2009) fornece todo o subsídio necessário para a implementação de algoritmos no programa WinBugs com alguns paralelos no programa R. Por sua vez, Albert (2004) dedica-se a discutir a implementação através do programa R e apresenta o pacote *LearnBayes* desenvolvido pelo próprio autor. Os dois apresentam a ligação existente entre *WinBugs* e R e também os pacotes BOA e CODA para avaliar a convergência das cadeias de Markov para a distribuição de interesse.

Esta dissertação utiliza o programa R e utiliza o pacote BOA para a utilização dos métodos de diagnóstico para a convergência dos algoritmos MCMC como, por exemplo, o método de Gelman e Rubin (1992 apud GAMERMAN; LOPES, 2006, Capítulo 5) que será discutido mais à frente. As seções seguintes seguem o raciocínio de Gamerman e

Lopes (2006, Capítulo 5) complementada por Ntzoufras (2009, Capítulo 2).

A configuração do computador utilizado é dada por um processador Intel 2 Quad 2,83GHz e 3,23 GB de RAM com sistema operacional Windows 32 bits. A programação dos métodos MCMC foi desenvolvida pelo autor e está disponível através do contato via email (marciodiniz@yahoo.com.br) caso o leitor esteja interessado.

2.4.4.1 Formação da amostra

Para a construção de uma amostra da distribuição de interesse, inicialmente, é necessário iterar o algoritmo até que a convergência tenha sido obtida como descrito nos procedimentos apresentados. A avaliação da convergência nas cadeias de Markov para π é uma questão aberta e com diversas discussões que serão brevemente expostas mais à frente.

As iterações necessárias para que a convergência seja obtida são descartadas e elas são denominadas por amostra de aquecimento (*burn-in*). Após obtida a amostra de aquecimento de tamanho z, qualquer elemento da cadeia de Markov amostrado terá aproximadamente a distribuição de interesse e, portanto, a amostra da distribuição de interesse pode ser formada. Há diversas estratégias para obter uma amostra tendo como objetivo a obtenção de uma amostra aleatória de π.

Gamerman e Lopes (2006, Capítulo 5) apresentam algumas possibilidades e comentam sobre a eficiência de cada estratégia. Este texto adota a estratégia discutida por Raftery e Lewis (1992 apud GAMERMAN; LOPES, 2006, Capítulo 5) que consiste em obter uma amostra da distribuição de interesse de tamanho n amostrando um elemento para cada k elementos gerados após obter a amostra de aquecimento de tamanho z. Esta estratégia utiliza os resultados ergódicos e o espaçamento entre os elementos amostrais na cadeia de Markov possibilita que a amostra obtida seja próxima de uma amostra aleatória.

2.4.4.2 Utilização da amostra

A partir da amostra gerada, é possível estimar qualquer quantidade relacionada à distribuição de interesse de $\boldsymbol{\theta}$ ou de maneira mais geral, de $\psi = \psi(\boldsymbol{\theta})$. Tais estimativas são sempre consistentes segundo o Teorema Ergódico dado na seção 2.3.1.

As quantidades usuais para obter informações sobre a distribuição de interesse são a média, o desvio padrão e alguns quantis.

Seja μ_ψ a média de π estimada por,

$$\hat{\mu}_\psi = \frac{1}{n}\sum_{t=1}^{n}\hat{\psi}_t, \tag{2.27}$$

tal que $\psi_t = \psi(\theta^{z+kt})$ para $t = 1, \ldots, n$.

A variância σ^2_ψ é estimada por,

$$\hat{\sigma}^2_\psi = \frac{1}{n}\sum_{t=1}^{n}(\psi_t - \hat{\mu}_\psi)^2. \tag{2.28}$$

Por fim, o quantil α é estimado por,

$$\hat{q}_\psi(\alpha) = \psi_{([\alpha n])}. \tag{2.29}$$

sendo que $\psi_{([s])}$ é uma estatística de ordem s.

2.4.4.3 Reparametrização e blocagem

Para aumentar a eficiência dos algoritmos, há duas estratégias usuais: reparametrização do modelo e blocagem de $\boldsymbol{\theta}$. As duas estratégias buscam aumentar a eficiência do algoritmo ao lidar com a correlação existente entre as componentes do vetor de parâmetros de interesse $\boldsymbol{\theta} = (\boldsymbol{\theta_1}, \ldots, \boldsymbol{\theta_d})$, ou seja, a estrutura de correlação inerente à densidade de interesse.

Usualmente, os candidatos são gerados por densidades propostas que não consideram tal estrutura de correlação por simplicidade na implementação dos algoritmos, isto é, são utilizadas densidades propostas univariadas.

Todavia, a aceitação dos candidatos é governada pela estrutura de correlação existente e se os candidatos são gerados sem que tal estrutura de correlação seja considerada, a eficiência do algoritmo é reduzida, já que candidatos que não são razoáveis, dada a estrutura de correlação, dificilmente serão aceitos. E portanto, quanto mais presente a estrutura de correlação, menor será a eficiência dos algoritmos.

A estratégia de reparametrização do modelo consiste em buscar parametrizações tal que os parâmetros tornem-se independentes ou próximos de independentes a posteriori e, desta forma, o algoritmo possa ser eficiente. Todavia, não há regras gerais para que uma parametrização útil seja encontrada.

Em geral, transformações lineares que resultem em uma matriz de variância assintótica de $\boldsymbol{\theta}$ diagonal são boas escolhas. Há alguns exemplos clássicos como a regressão linear simples ou modelos com efeitos aleatórios que são apresentados por Gamerman e Lopes (2006, Capítulo 5). E tal estratégia é eficiente quando aplicada no Amostrador de Gibbs ou Metropolis-Hastings.

A estratégia de blocagem consiste em construir blocos de parâmetros para as componentes de $\boldsymbol{\theta}$ que apresentem estrutura de correlação e, assim, as densidades propostas são construídas através da incorporação desta estrutura e tornam-se capazes de gerar candidatos mais razoáveis para tais blocos.

Esta é uma estratégia sempre eficiente no Amostrador de Gibbs como mostrado por Liu, Wong e Kong (1994 apud GAMERMAN; LOPES, 2006, Capítulo 5), por outro lado, o mesmo não pode ser dito quando aplicada a Metropolis-Hastings já que as densidades propostas $q_i(.;.)$ tornam-se mais difíceis para serem construídas de maneira similar a $\pi(\boldsymbol{\theta_i}|\boldsymbol{\theta_{-i}})$ conforme o aumento da dimensão do bloco de parâmetros $\boldsymbol{\theta_i}$.

A regra geral é construir blocos sempre que possível desde que as taxas de aceitação não se tornem muito pequenas.

2.4.5 Diagnósticos de convergência

Os algoritmos MCMC são baseados no princípio de que para um z suficientemente grande, a distribuição da cadeia de Markov será próxima à distribuição estacionária

que, por sua vez, é a distribuição de interesse.

A convergência das cadeias de Markov construídas pelos algoritmos MCMC é garantida teoricamente, todavia, os resultados teóricos não são capazes de estipular um valor para z, denominado tamanho da amostra de aquecimento na seção 2.4.4.1.

Para escolhê-lo, é necessário avaliar a convergência da cadeia de Markov para a distribuição estacionária, esta convergência é dividida em dois tipos principais segundo Robert e Casella (2004, Capítulo 8):

- Convergência das médias amostrais: Baseia-se em avaliar se

$$\frac{1}{z}\sum_{t=1}^{z} h(\theta_t) \to E_\pi(h(\theta)) \quad \text{quando} \quad n \to \infty. \tag{2.30}$$

  O propósito desta avaliação é verificar se a cadeia de Markov foi capaz de captar as características da distribuição de interesse. Brooks e Roberts (1998 apud ROBERT; CASELLA, 2004, Capítulo 8) relacionam esta convergência com a velocidade de mistura da cadeia, isto é, o nível de dependência dos valores iniciais e a velocidade na exploração do suporte da distribuição estacionária.

- Convergência para amostras i.i.d: Consiste em avaliar quão distante uma amostra obtida está de uma amostra aleatória.

E para verificar tais pontos, há duas abordagens exploradas: teórica e prática. A abordagem teórica consiste em buscar medidas de distância e estabelecer limites para as distribuições construídas pelas cadeias de Markov e, nesta linha, pode-se citar Meyn e Tweedie (1994 apud GAMERMAN; LOPES, 2006), Polson (1996 apud GAMERMAN; LOPES, 2006) e muitos outros. Contudo, o impacto prático de tais trabalhos ainda é pequeno segundo Cowles e Carlin (1996).

A abordagem prática consiste em avaliar as propriedades das observações da cadeia de Markov. Esta é a abordagem frequentemente utilizada, entretanto, não é possível garantir, de fato, a convergência já que a conclusão é dependente daquelas específicas observações da cadeia de Markov.

No contexto da abordagem prática, há diversos métodos construídos sobre diver-

sas hipóteses sendo os principais dados por Geweke (1992 apud GAMERMAN; LOPES, 2006, Capítulo 5) que considera uma única e longa cadeia de Markov para avaliar a convergência, Gelman e Rubin (1992 apud GAMERMAN; LOPES, 2006, Capítulo 5) que considera múltiplas cadeias de Markov com valores iniciais espalhados pelo espaço paramétrico.

A literatura sugerida sobre métodos de diagnóstico de convergência consiste em Cowles e Carlin (1996), Robert e Casella (2004, Capítulo 8), e também Gamerman e Lopes (2006, Capítulo 5). Os principais métodos estão implementados nos pacotes CODA e BOA do programa R. Ntzoufras (2009) apresenta um guia para a utilização destes pacotes.

Nesta dissertação, o método de Gelman e Rubin (1992 apud GAMERMAN; LOPES, 2006, Capítulo 5) é utilizado. O método consiste em considerar p cadeias paralelas e suas respectivas trajetórias $\{\lambda_i^0, \ldots, \lambda_i^n\}$ para $i = 1, \ldots, p$ tal que os valores iniciais estejam dispersos pelo suporte da distribuição a posteriori.

Então, as variâncias dentro das cadeias e entre as cadeias são dadas por B e W,

$$B = \frac{n}{p-1}\sum_{i=1}^{p}(\bar{\lambda}_i - \bar{\lambda})^2, \tag{2.31}$$

$$W = \frac{1}{p(n-1)}\sum_{i=1}^{p}\sum_{j=1}^{n}(\lambda_i^{(j)} - \bar{\lambda}_i)^2, \tag{2.32}$$

em que, $\bar{\lambda}_i$ é média das observações da i-ésima cadeia para $i = 1, \ldots, p$ e $\bar{\lambda}$ é a média das médias.

Sob convergência, todas as np observações são provenientes da distribuição a posteriori e σ_λ^2, a variância a posteriori de λ pode ser consistentemente estimada por,

$$\hat{\sigma}_\lambda^2 = \left(1 - \frac{1}{n}\right)W + \frac{1}{n}B. \tag{2.33}$$

Se não houve convergência, os valores iniciais ainda exercem influência sobre as cadeias e devido à dispersão dos valores iniciais no suporte da distribuição a posteriori, $\hat{\sigma}_\lambda^2$ irá superestimar σ_λ^2. E também, antes da convergência, W irá subestimar σ_λ^2 porque cada cadeia não terá percorrido adequadamente o suporte da distribuição a posteriori.

Então, é possível definir uma estatística R que é um indicador da convergência da cadeia de Markov para a distribuição estacionária,

$$R = \sqrt{\frac{\hat{\sigma}_{\lambda}^{2}}{W}}. \tag{2.34}$$

Para $n \to \infty$, ambos estimadores convergem para σ_{λ}^{2} pelo Teorema Ergódico e $\hat{R} \to 1$. Note que R é sempre maior do que 1 e segundo Gelman (1996 apud GAMERMAN; LOPES, 2006, Capítulo 5), a convergência é considerada alcançada se $\hat{R}$ é menor do que 1,2.

Um inconveniente deste método é a dependência da escolha dos valores iniciais da cadeia de Markov e da teoria de normalidade presente na formulação dos estimadores para variância. Uma alternativa é a utilização de estimadores não-paramétricos, outra possibilidade é a utilização de reparametrizações para que a distribuição a posteriori torne-se próxima à distribuição normal.

Um ponto ressaltado por todos os autores é que os diagnósticos sobre a convergência de uma cadeia de Markov, a partir da abordagem prática, não são conclusivos e é interessante que diversos métodos sejam utilizados já que cada um dos métodos citados apresentam hipóteses, vantagens, desvantagens e casos que não devem ser utilizados.

Nesta dissertação, além da estatística R de Gelman e Rubin (1992 apud GAMERMAN; LOPES, 2006, Capítulo 5), também são utilizados métodos gráficos que consistem na construção das trajetórias das médias amostrais para as quantidades de interesse considerando múltiplas cadeias de Markov, dos históricos e das autocorrelações para as componentes de θ.

2.4.6 Dados aumentados

A presença de dados faltantes no contexto epidemiológico é uma constante e, por isso, há a necessidade da utilização de dados aumentados. Tanner e Wong (1987 apud GAMERMAN; LOPES, 2006, Capítulo 4) construíram um algoritmo para lidar com parâmetros de não interesse, variáveis latentes e dados adicionais no contexto dos al-

goritmos MCMC. Esta ferramenta é extremamente utilizada e se pode citar diversos autores como, por exemplo, O'Neill e Roberts (1999), Patz e Junker (1999), Paulino, Silva e Achcar (2005), Lekone e Finkenstädt (2006).

Para tanto, seja $\boldsymbol{\theta_i} = (\boldsymbol{\phi}, \boldsymbol{\psi})$ tal que o interesse recaia sobre $\boldsymbol{\phi}$ enquanto $\boldsymbol{\psi}$ seja um conjunto de parâmetros de não interesse, variáveis latentes ou dados adicionais. O algoritmo proposto para atualizar $\pi^{(n)}(\boldsymbol{\phi})$ para $\pi^{(n+1)}(\boldsymbol{\phi})$ é

1. Amostre $\boldsymbol{\phi_1}, \ldots, \boldsymbol{\phi_m}$ a partir de $\pi^{(n)}(\boldsymbol{\phi})$;
2. Amostre $\boldsymbol{\psi_1}, \ldots, \boldsymbol{\psi_m}$ a partir de $\pi(\boldsymbol{\psi})$. Este passo pode ser aproximado pela amostragem $\boldsymbol{\psi_i}$ de $\pi(\boldsymbol{\psi}|\boldsymbol{\phi_i})$ para $i = 1, \ldots, m$;
3. Construa a aproximação Monte Carlo

$$\pi^{(n+1)}(\boldsymbol{\phi}) = \frac{1}{n} \sum_{i=1}^{m} \pi(\boldsymbol{\phi}|\boldsymbol{\psi_i}). \tag{2.35}$$

Tanner e Wong (1987 apud GAMERMAN; LOPES, 2006, Capítulo 4) mostram que o algoritmo terá uma única distribuição estacionária $\pi(\boldsymbol{\phi})$ para qualquer m escolhido. Em especial, utiliza-se $m = 1$ e o último passo não precisa necessariamente ser realizado.

3 Modelo SEIR com taxa de remoção homogênea

Este capítulo tem por objetivo apresentar e discutir o modelo SEIR estocástico utilizado por Lekone e Finkenstädt (2006) para o estudo da epidemia de febre hemorrágica Ébola ocorrida na República Democrática do Congo em 1995. O modelo utiliza toda informação disponível, ou seja, lida com dados incompletos e também com a introdução de medidas de intervenção ao longo da epidemia. Para tanto, procedimentos bayesianos são considerados.

Previamente a Lekone e Finkenstädt (2006), Chowell et al. (2004) modela esta epidemia somente a partir da série diária incompleta do número de indivíduos expostos que se tornam infectantes e também considera a introdução de medidas de intervenção a partir de uma abordagem determinística. A estimação de parâmetros é realizada via Mínimos Quadrados.

Apesar das duas abordagens distintas, estocástica e determinística, tais modelos têm pontos comuns e dentre eles, a suposição de taxas de exposição e remoção homogêneas no tempo, isto é, tais taxas são modeladas a partir de uma distribuição *Exponencial*. Esta suposição é usual na abordagem estocástica pela facilidade algébrica na modelagem de epidemias, todavia, ela não é sempre razoável, principalmente, quando há medidas de intervenção que reduzem o período de remoção. Neste capítulo, a suposição de tempo exponencial para o período de remoção é mantida assim como em Lekone e Finkenstädt (2006) e será questionada mais à frente.

O capítulo está organizado da seguinte forma: Na seção 3.1, o modelo é apresentado; Nas seções 3.2 e 3.3, discute-se os procedimentos bayesianos para o modelo estudado, e

por fim, a seção 3.4 é constituída de simulações.

3.1 O modelo

O modelo estocástico SEIR estudado considera tempo discretizado e o número de indivíduos como uma quantidade discreta sendo baseado na Figura 3.

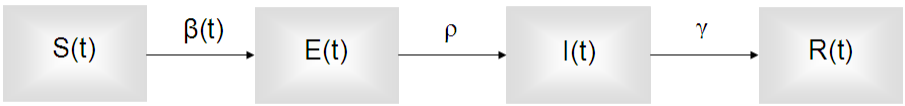


Figura 3: Modelo SEIR com taxa de remoção homogênea

O intervalo de tempo é definido por $[t, t+h)$ e denominado período t, onde h representa o tamanho do intervalo, no caso, $h = 1$ dia. Considere,

- $B(t)$ é o número de indivíduos suscetíveis que se tornam expostos no período t;
- $C(t)$ é o número de indivíduos expostos que se tornam infectantes no período t;
- $D(t)$ é o número de indivíduos infectantes que se tornam removidos no período t;

O período que indica o término da epidemia é denotado por τ^* e desta forma, define-se as séries temporais desde o início até o final da epidemia, $\mathbf{B} = \{B(t)\}_{t=1}^{\tau^*}$, $\mathbf{C} = \{C(t)\}_{t=1}^{\tau^*}$ e $\mathbf{D} = \{D(t)\}_{t=1}^{\tau^*}$.

Sejam $S(t)$, $E(t)$, $I(t)$ e $R(t)$, respectivamente, o número de indivíduos suscetíveis, expostos, infectantes e removidos no período t. Tais valores são atualizados no início do intervalo de tempo $[t, t+h)$.

Então, dadas as condições iniciais ($S(1) = s_0$, $E(1) = e_0$, $I(1) = i_0$, $R(1) = r_0$) e o

tamanho da população N constante, o modelo é dado por,

$$\begin{aligned} S(t+h) &= S(t) - B(t) \\ E(t+h) &= E(t) + B(t) - C(t) \\ I(t+h) &= I(t) + C(t) - D(t) \\ N &= S(t) + E(t) + I(t) + R(t), \end{aligned} \tag{3.1}$$

em que

$$\begin{aligned} B(t) &\dot{\sim} \; Binomial(S(t),\, p_B(t)) \\ C(t) &\dot{\sim} \; Binomial(E(t),\, p_C) \\ D(t) &\dot{\sim} \; Binomial(I(t),\, p_D), \end{aligned} \tag{3.2}$$

tal que

$$\begin{aligned} p_B(t) &= 1 - exp\left(-\frac{\beta(t)}{N} h I(t)\right) \\ p_C &= 1 - exp(-\rho h) \\ p_D &= 1 - exp(-\gamma h). \end{aligned} \tag{3.3}$$

Os parâmetros $\beta(t)$, $\frac{1}{\rho}$, $\frac{1}{\gamma}$ representam, respectivamente, a taxa de contato no período t e os períodos médios de exposição e remoção. Note que ρ, γ são as taxas médias de exposição e remoção, respectivamente, enquanto, a taxa de infecção no período t é dada por $\frac{\beta(t)}{N} I(t)$.

Observe que a transição dos indivíduos entre os compartimentos consiste em um movimento estocástico. Em cada intervalo de tempo $[t, t+h)$, cada indivíduo pode permanecer no compartimento atual ou mover-se para o próximo compartimento, ou seja, o movimento de um indivíduo pode ser considerado um ensaio de *Bernoulli*.

As distribuições *Binomiais* resultam da soma de todos os ensaios de *Bernoulli* independentes e identicamente distribuídos para os indivíduos de cada compartimento desde que o número de indivíduos no compartimento permaneça constante no intervalo de tempo considerado.

É necessário fazer algumas considerações sobre tais suposições: O número de indivíduos em um compartimento é aproximadamente constante se o intervalo de tempo é pequeno suficientemente. Além disso, as suposições de que os indivíduos são independêntes e identicamente distribuídos é uma grande simplificação.

Sob a suposição de que o período que um indivíduo permanece em um compartimento é exponencialmente distribuído com uma taxa específica do compartimento λ, a probabilidade do indivíduo permanecer no compartimento por um período de extensão h é $exp\{-\lambda h\}$, consequentemente, a probabilidade do indivíduo mudar de compartimento no intervalo $[t, t+h)$ é $1 - exp\{-\lambda h\}$.

Observe que λ não é constante na situação real, contudo, pode ser considerado constante em um pequeno intervalo de tempo, desde que não ocorra qualquer mudança de compartimento entre os indivíduos, por isso, a distribuição aproximada na equação (3.2).

A partir de tais considerações, as equações (3.2) e (3.3) são estabelecidas. Por fim, para considerar os efeitos da intervenção, define-se,

$$\beta(t) = \begin{cases} \beta & \text{se t} < \tau \\ \beta e^{-q(t-\tau)} & \text{se t} \geq \tau, \end{cases} \tag{3.4}$$

sendo τ é o período t em que as medidas de intervenção foram introduzidas, β é a taxa de contato inicial e q é a taxa de decaimento de $\beta(t)$ a partir do período τ. Note que ao considerar medidas de intervenção, a taxa de contato deixa de ser homogênea no tempo.

É importante ressaltar que τ também é denominado ponto de mudança e para a equação (3.4) é conhecido.

3.2 Estimação dos parâmetros

3.2.1 A função de verossimilhança

O modelo dado pelo conjunto de equações (3.1) - (3.4) é definido a partir do conjunto de parâmetros $\boldsymbol{\theta} = (\beta, \rho, \gamma, q)$ que pode ser estimado a partir do conhecimento das condições iniciais, tamanho da população e as séries diárias $\{\mathbf{B}, \mathbf{C}, \mathbf{D}\}$.

Além disso, note que $\{S(t),\ E(t),\ I(t)\}$ são completamente determinados a partir do sistema de equações (3.1) e como $\{\mathbf{B},\ \mathbf{C},\ \mathbf{D}\}$ são condicionalmente independentes, a função verossimilhança dos dados é definida por,

$$L(\mathbf{B}, \mathbf{C}, \mathbf{D}|\boldsymbol{\theta}) = \prod_{t=1}^{\tau^*-1} g_1(B(t)|\boldsymbol{\theta}, I_t) g_2(C(t)|\boldsymbol{\theta}, I_t) g_3(D(t)|\boldsymbol{\theta}, I_t), \tag{3.5}$$

onde g_1, g_2 e g_3 são as funções de probabilidades estabelecidas em (3.2) - (3.4) condicionadas em $\boldsymbol{\theta}$ e toda informação disponível no período t denotada por I_t. Então,

$$\begin{aligned}
L(\mathbf{B}, \mathbf{C}, \mathbf{D}|\boldsymbol{\theta}) \;=\; & \prod_{t=1}^{\tau-1} \binom{S(t)}{B(t)} \left(1 - \exp\left(-\frac{\beta}{N} I(t)\right)\right)^{B(t)} \exp\left(-\frac{\beta}{N} I(t)(S(t) - B(t))\right) \times \\
& \times \prod_{t=\tau}^{\tau^*-1} \binom{S(t)}{B(t)} \left(1 - \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right)^{B(t)} \times \\
& \times \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)(S(t) - B(t))\right) \times \\
& \times \prod_{t=1}^{\tau^*-1} \binom{E(t)}{C(t)} (1 - e^{-\rho})^{C(t)} e^{-\rho(E(t)-C(t))} \times \\
& \times \prod_{t=1}^{\tau^*-1} \binom{I(t)}{D(t)} (1 - e^{-\gamma})^{D(t)} e^{-\gamma(I(t)-D(t))}.
\end{aligned} \tag{3.6}$$

Como $S(1) = N$ e $S(\tau^* - 1) = N - m$ sendo que o tamanho da população N, em geral, é muito grande e o tamanho da epidemia m é pequeno além do fato de que a probabilidade $P(t)$, dada por (3.3), é pequena para quase todo t, pode-se aproximar a densidade de $\mathbf{B}$ pela distribuição *Poisson*. A função de verossimilhança aproximada segue,

$$
\begin{aligned}
L(\mathbf{B},\mathbf{C},\mathbf{D}|\boldsymbol{\theta}) &= \prod_{t=1}^{\tau-1} \frac{1}{B(t)!} \exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right) \times \\
&\times \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right]^{B(t)} \times \\
&\times \prod_{t=\tau}^{\tau^*-1} \frac{1}{B(t)!} \exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right) \times \\
&\times \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right]^{B(t)} \times \\
&\times \prod_{t=1}^{\tau^*-1} \binom{E(t)}{C(t)} (1-e^{-\rho})^{C(t)} e^{-\rho(E(t)-C(t))} \times \\
&\times \prod_{t=1}^{\tau^*-1} \binom{I(t)}{D(t)} (1-e^{-\gamma})^{D(t)} e^{-\gamma(I(t)-D(t))}. \qquad (3.7)
\end{aligned}
$$

Todavia, lembre-se que a série diária $\mathbf{B}$ é praticamente desconhecida com exceção do primeiro caso e as séries diárias de dados $\mathbf{C}$ e $\mathbf{D}$ são parcialmente conhecidas, por isso, é necessário considerar dados aumentados. Os dados aumentados considerados por Lekone e Finkenstädt (2006) são exatamente as séries $\mathbf{B}$, $\mathbf{C}$ e $\mathbf{D}$.

Desta forma, é necessário introduzir a seguinte notação: $\mathbf{B}$, $\mathbf{C}$ e $\mathbf{D}$ são as séries diárias de dados aumentados; $\mathbf{B}_p$, $\mathbf{C}_p$ e $\mathbf{D}_p$ são as séries diárias de dados parcialmente conhecidas.

Portanto, a função de verossimilhança para os dados aumentados é ainda dada pela expressão (3.7). O próximo passo, segundo a abordagem bayesiana, é estabelecer distribuições a priori para $\boldsymbol{\theta}$.

3.2.2 As distribuições a priori

Para a escolha das distribuições a priori, é necessário considerar alguns pontos. O primeiro ponto a ser considerado são as restrições sobre o espaço paramétrico: $0 < \beta < 1$, $q > 0$, $0 < \gamma < 1$ e $0 < \rho < 1$.

Tais restrições para γ e ρ surgem da utilização de dados diários que impossibilita a

estimação de $\frac{1}{\gamma}$, $\frac{1}{\rho}$ menores do que 1. A restrição sobre β é dada pela definição de taxa de contato e para q consiste no princípio de que medidas de intervenção implicam em uma redução da taxa de contato entre suscetíveis e infectantes.

O segundo ponto é o conhecimento apresentado na literatura sobre o período de exposição com duração de 2 a 21 dias com média 6,1 segundo Breman, Piot e Johnson (1977 apud LEKONE; FINKENSTäDT, 2006), e o período de remoção de 3,5 a 10,7 dias segundo Birmingham e Cooney (2002 apud CHOWELL et al., 2004). Por último, assume-se independência entre os parâmetros por facilidade algébrica.

Desta forma, a partir da positividade dos parâmetros, uma primeira possibilidade é a escolha de distribuções $Gama(a_\lambda, b_\lambda)$. A escolha de a_λ e b_λ para $\lambda \in \{\rho, \gamma\}$, pode ser definida de forma que a média da distribuição a priori seja próxima da taxa média de exposição dada na literatura e a taxa média de remoção, sendo que esta última pode ser calculada a partir da média do intervalo em que é limitado o período de remoção na literatura.

Para β e q, a escolha de a_λ e b_λ é arbitrária desde que a massa das distribuições a priori limite-se ao intervalo (0,1) já que por definição $\beta \in (0,1)$ e, no caso do parâmetro q, é a partir da consideração de que $q = 1$ significa uma redução de 70% da taxa de contato em um único dia da introdução das medidas de intervenção, um valor muito alto para qualquer epidemia.

Por fim, para uma análise de sensibilidade da distribuição a priori no processo de estimação, define-se mais de um conjunto de parâmetros de maneira similar a Lekone e Finkenstädt (2006): $((a_\beta, b_\beta), (a_\rho, b_\rho), (a_\gamma, b_\gamma), (a_q, b_q)) \in \{((2,10),(2,12),(2,14),(2,10)),$ $((20,100),(20,120),(20,140),(2,10)),((20,40),(20,40),(20,40),(20,40))\}$.

O primeiro vetor de parâmetros é denominado distribuição a priori não informativa, o segundo vetor de parâmetros é denominado distribuição a priori informativa e, por fim, o terceiro vetor de parâmetros é denominado distribuição a priori não centrada. Note que a distribuição a priori para q no segundo conjunto de parâmetros continua não informativa, pois é difícil estabelecer o efeito das medidas de intervenção.

Uma segunda possibilidade de distribuições a priori seria a utilização de distri-

buições uniformes tal que as restrições sobre o espaço paramétrico fossem consideradas. Todavia, esta escolha não permite que a distribuição a posteriori de q seja maximizada numericamente quando os dados são incompletos.

3.2.3 A distribuição a posteriori

A distribuição a posteriori para o conjunto de parâmetros $\boldsymbol{\theta}$ e os dados aumentados $\mathbf{B}$, $\mathbf{C}$ e $\mathbf{D}$ a partir de $\mathbf{B}_p$, $\mathbf{C}_p$ e $\mathbf{D}_p$ segue,

$$\pi(\boldsymbol{\theta}, \mathbf{B}, \mathbf{C}, \mathbf{D}|\mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) \propto L(\mathbf{B}, \mathbf{C}, \mathbf{D}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}). \tag{3.8}$$

Com a suposição de distribuições a prioris como distribuições *Gama*, tem-se,

$$\begin{aligned}
\pi(\boldsymbol{\theta}, \mathbf{B}, \mathbf{C}, \mathbf{D}|\mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) \quad \propto \quad & \prod_{t=1}^{\tau-1} \frac{1}{B(t)!} \exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right) \times \\
\times \quad & \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right]^{B(t)} \times \\
\times \quad & \prod_{t=\tau}^{\tau^*-1} \frac{1}{B(t)!} \exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right) \times \\
\times \quad & \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right]^{B(t)} \times \\
\times \quad & \prod_{t=1}^{\tau^*-1} \binom{E(t)}{C(t)} (1-e^{-\rho})^{C(t)} e^{-\rho(E(t)-C(t))} \times \\
\times \quad & \prod_{t=1}^{\tau^*-1} \binom{I(t)}{D(t)} (1-e^{-\gamma})^{D(t)} e^{-\gamma(I(t)-D(t))} \times \\
\times \quad & \beta^{a_\beta-1} e^{-b_\beta\beta} \rho^{a_\rho-1} e^{-b_\rho\rho} \gamma^{a_\gamma-1} e^{-b_\gamma\gamma} q^{a_q-1} e^{-b_q q} \times \\
\times \quad & I(\beta)_{(0,\infty)} I(\rho)_{(0,\infty)} I(\gamma)_{(0,\infty)} I(q)_{(0,\infty)}.
\end{aligned} \tag{3.9}$$

A estimação dos parâmetros e a simulação dos dados aumentados são realizadas a partir de métodos MCMC.

3.3 Métodos MCMC

O algoritmo Metropolis-Hasting por Blocos é a ferramenta adequada para a estimação do vetor de parâmetros $\boldsymbol{\theta}$ e a simulação dos dados aumentados a partir de $\pi(\boldsymbol{\theta}, \mathbf{B}, \mathbf{C}, \mathbf{D}|\mathbf{C}_p, \mathbf{D}_p)$. O algoritmo geral segue,

1. Inicialize as séries diárias de dados **B**, **C** e **D**.
2. Reconstrua as trajetórias $\{S(t),\, E(t),\, I(t)\}$ a partir das equações em (3.1);
3. Inicialize o vetor $\boldsymbol{\theta}$;
4. Atualize **B** a partir de $\pi(\mathbf{B}|\mathbf{C}, \mathbf{D}, \boldsymbol{\theta})$ e reconstrua as trajetórias $\{S(t),\, E(t)\}$;
5. Atualize **C** a partir de $\pi(\mathbf{C}|\mathbf{B}, \mathbf{D}, \boldsymbol{\theta})$ e reconstrua as trajetórias $\{E(t),\, I(t)\}$;
6. Atualize **D** a partir de $\pi(\mathbf{D}|\mathbf{B}, \mathbf{C}, \boldsymbol{\theta})$ e reconstrua as trajetórias $\{I(t)\}$;
7. Atualize $\boldsymbol{\theta}$ a partir de $\pi(\boldsymbol{\theta}|\mathbf{B}, \mathbf{C}, \mathbf{D})$;
8. Repita os passos 4 - 7 até que se obtenha a amostra de tamanho desejado após a convergência ter sido alcançada.

3.3.1 Inicialização de B, C e D

As inicializações das séries diárias de dados **B**, **C** e **D** podem ser feitas de diversas formas desde que não modifiquem os dados conhecidos $\mathbf{B}_p$, $\mathbf{C}_p$ e $\mathbf{D}_p$ e satisfaçam as seguintes condições conjuntamente,

1. $\sum_{t=1}^{\tau^*-1} B(t) = m$;
2. $\sum_{t=1}^{\tau^*-1} C(t) = m$;
3. $\sum_{t=1}^{\tau^*-1} D(t) = m$;
4. $E(t) - C(t) \geq 0 \ \forall t$;

5. $I(t) - D(t) \geq 0 \ \forall t$;

6. $E(t) + I(t) > 0$ para $t \in \{2, \ldots, \tau^* - 1\}$;

7. $E(\tau^*) + I(\tau^*) = 0$;

Lekone e Finkenstädt (2006) consideram a condição 5 como $E(t) \geq 0 \ \forall t$, todavia, não é suficiente que $E(t)$ satisfaça tal condição pois é possível, por exemplo, a partir de um $\mathbf{B}$ aumentado que tal $B(t) = 1$, $E(t) = 0$ e a partir dos dados $C(t) = 1$, então,

$$E(t+1) = E(t) + B(t) - C(t) = 0 \tag{3.10}$$

e a condição dada em Lekone e Finkenstädt (2006)) é satisfeita. Entretanto, é incoerente que os indivíduos expostos que se tornam infectantes no período t seja igual a 1, enquanto o número de indivíduos expostos no mesmo período seja 0.

Além disso, as condições 4 - 7 e o conjunto de dados considerado implicam em,

1. $B(t)$ pode ser modificado $\forall t \in \{8, \ldots, \tau^* - 3\}$

2. $C(t)$ pode ser modificado $\forall t \in \{9, \ldots, \tau^* - 2\}$

3. $D(t)$ pode ser modificado $\forall t \in \{8, \ldots, \tau^* - 1\}$

já que o primeiro caso torna-se infectante somente no dia 8. A utilização destas condições permite reduzir as possibilidades na construção de candidatos na atualização para $\mathbf{B}, \mathbf{C}$ e $\mathbf{D}$.

Dadas as condições acima, a inicialização das séries diárias pode ser feita e, novamente, este texto não segue Lekone e Finkenstädt (2006): A série diária de dados $\mathbf{B}$ é inicializada dividindo os indivíduos em 10 parcelas iguais e os posicionando ao longo dos 20 dias alternadamente a partir do dia 8 sendo que o resto da divisão em 10 parcelas é somado aos indivíduos já posicionados no dia 8.

Este procedimento tem como justificativa aumentar a velocidade do espalhamento dos indivíduos na série diária de dados $\mathbf{B}$ para se aproximar da série de diária de dados desconhecida. Lekone e Finkenstädt (2006) concentram todos os indivíduos no dia 8.

A inicialização da série diária de dados **C** e **D** não é apresentada em Lekone e Finkenstädt (2006), desta forma, o procedimento utilizado foi o sorteio aleatório das posições dos indivíduos faltantes entres todos os dias possíveis para a série diária de dados **C** e entre os dias $\tau/2$ até $\tau^* - 1$ para a série diária de dados **D**. A determinação deste último intervalo de tempo tem como razão aumentar a probabilidade de encontrar uma série diária de dados inicial **D** que satisfaça as condições 1 - 7 que foram apresentadas acima.

A reconstrução das trajetórias $\{S(t),\ E(t),\ I(t)\}$ é feita a partir das equações em (3.1) e das condições iniciais: $s_0 = 5363500$, $e_0 = i_0 = r_0 = 0$.

3.3.2 Inicialização de $\boldsymbol{\theta}$

Os valores iniciais para $\boldsymbol{\theta}$ podem ser encontrados a partir da distribuição a priori para β, q e para ρ e γ podem ser utilizados as informações dada na literatura sobre a doença. Neste texto, $\rho_0 = \dfrac{1}{6,1}$, $\gamma_0 = \dfrac{1}{7,2}$ e $\beta_0 = 0,5$ por ser a metade do intervalo em que o parâmetro β está definido e $q_0 = 0,5$ foi escolhido arbitrariamente com o cuidado de que a redução na taxa de contato não fosse muito grande, no caso, a redução é de 40%.

E como é necessário considerar pelo menos duas cadeias de Markov para verificar a convergência para a distribuição estacionária como apresentado na subseção 2.4.4.4 do Capítulo 2, o segundo conjunto de valores iniciais é dado por $\beta_0 = 0,1$, $\rho_0 = 0,45$, $\gamma_0 = 0,25$ e $q_0 = 0,1$.

3.3.3 Atualização de B, C e D

A atualização das séries diárias de dados **B**, **C** e **D** é realizada sob as mesmas condições consideradas para encontrar as séries diárias de dados iniciais. Para tanto, é necessário calcular as distribuições marginais a posteriori de tais séries diárias de dados,

$$
\begin{aligned}
\pi(\mathbf{B}|\mathbf{C},\mathbf{D},\boldsymbol{\theta}) \;\propto\; & \prod_{t=1}^{\tau-1}\frac{1}{B(t)!}\exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right)\times \\
& \times\; \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right]^{B(t)}\times \\
& \times\; \prod_{t=\tau}^{\tau^*-1}\frac{1}{B(t)!}\exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right)\times \\
& \times\; \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right]^{B(t)}\times \\
& \times\; \prod_{t=1}^{\tau^*-1}\binom{E(t)}{C(t)}e^{-\rho E(t)}, \qquad (3.11)\\
\pi(\mathbf{C}|\mathbf{B},\mathbf{D},\boldsymbol{\theta}) \;\propto\; & \prod_{t=1}^{\tau-1}\exp\left(S(t)\exp\left(-\frac{\beta}{N}I(t)\right)\right)\times \\
& \times\; \left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)^{B(t)}\times \\
& \times\; \prod_{t=\tau}^{\tau^*-1}\exp\left(S(t)\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\times \\
& \times\; \left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)^{B(t)}\times \\
& \times\; \prod_{t=1}^{\tau^*-1}\binom{E(t)}{C(t)}(1-e^{-\rho})^{C(t)}e^{-\rho(E(t)-C(t))}\times \\
& \times\; \prod_{t=1}^{\tau^*-1}\binom{I(t)}{D(t)}e^{-\gamma I(t)}, \qquad (3.12)\\
\pi(\mathbf{D}|\mathbf{B},\mathbf{C},\boldsymbol{\theta}) \;\propto\; & \prod_{t=1}^{\tau-1}\exp\left(S(t)\exp\left(-\frac{\beta}{N}I(t)\right)\right)\times \\
& \times\; \left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)^{B(t)}\times \\
& \times\; \prod_{t=\tau}^{\tau^*-1}\exp\left(S(t)\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\times \\
& \times\; \left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)^{B(t)}\times \\
& \times\; \prod_{t=1}^{\tau^*-1}\binom{I(t)}{D(t)}(1-e^{-\gamma})^{D(t)}e^{-\gamma(I(t)-D(t))}. \qquad (3.13)
\end{aligned}
$$

Então, denote $\mathbf{B}'$ a nova série diária de dados candidata ao lugar de $\mathbf{B}$, o procedimento para a atualização de $\mathbf{B}$ segue,

1. Faça $\mathbf{B}' = \mathbf{B}$;
2. Verifique quais são as posições não nulas da série diária de dados $\mathbf{B}'$;
3. Sorteie uma posição dentre as posições não nulas de $B'(t)$ para $8 \leq t \leq \tau^* - 3$ e faça $B'(t) = B'(t) - 1$;
4. Sorteie uma posição da série diária de dados $B'(t)$ para $8 \leq t \leq \tau^* - 3$ e faça $B'(t) = B'(t) + 1$;
5. Repita $nmodB$ vezes, os passos 2 - 4;
6. Reconstrua as trajetórias $\{S'(t),\, E'(t)\}$ a partir de (3.1);
7. Verifique se a série diária de dados $\mathbf{B}'$ satisfaz as condições 1, 4, 6 e 7 listadas na subseção 3.3.1;
8. Aceite $\mathbf{B}'$ com a seguinte probabilidade:

$$\alpha(\mathbf{B}, \mathbf{B}') \;=\; min\{1, R_B\} \tag{3.14}$$

em que

$$\begin{aligned} R_B &= \frac{\pi(\mathbf{B}'|\mathbf{C}, \mathbf{D}, \boldsymbol{\theta})}{\pi(\mathbf{B}|\mathbf{C}, \mathbf{D}, \boldsymbol{\theta})} \\ &= \exp(log(\pi(\mathbf{B}'|\mathbf{C}, \mathbf{D}, \boldsymbol{\theta})) - log(\pi(\mathbf{B}|\mathbf{C}, \mathbf{D}, \boldsymbol{\theta}))). \end{aligned} \tag{3.15}$$

Denote $\mathbf{C}'$ a nova série diária de dados candidata ao lugar de $\mathbf{C}$. O procedimento para a atualização de $\mathbf{C}$ segue,

1. Faça $\mathbf{C}' = \mathbf{C}$;
2. Seja $\mathbf{C}_a$ a série diária de dados que complementa a série diária de dados conhecida $\mathbf{C}_p$, faça $\mathbf{C}_a = \mathbf{C}'$ - $\mathbf{C}_p$;
3. Verifique quais são as posições não nulas da série diária de dados $\mathbf{C}_a$;

4. Sorteie uma posição dentre as posições não nulas de $\mathbf{C}_a$ e faça $C_a(t) = C_a(t) - 1$;

5. Sorteie uma posição da série diária de dados $C_a(t)$ para $9 \leq t \leq \tau^* - 2$ e faça $C_a(t) = C_a(t) + 1$;

6. Repita $nmodC$ vezes, os passos 3 - 5;

7. Calcule $\mathbf{C}' = \mathbf{C}_a + \mathbf{C}_p$;

8. Reconstrua as trajetórias $\{E'(t)\ I'(t)\}$ a partir de (3.1);

9. Verifique se a série diária de dados $\mathbf{C}'$ satisfaz as condições 2, 4, 5, 6 e 7 listadas na subseção 3.3.1;

10. Aceite $\mathbf{C}'$ com a seguinte probabilidade:

$$\alpha(\mathbf{C}, \mathbf{C}') = min\{1, R_C\} \tag{3.16}$$

em que

$$\begin{aligned} R_C &= \frac{\pi(\mathbf{C}'|\mathbf{B}, \mathbf{D}, \boldsymbol{\theta})}{\pi(\mathbf{C}|\mathbf{B}, \mathbf{D}, \boldsymbol{\theta})} \\ &= \exp(log(\pi(\mathbf{C}'|\mathbf{B}, \mathbf{D}, \boldsymbol{\theta})) - log(\pi(\mathbf{C}|\mathbf{B}, \mathbf{D}, \boldsymbol{\theta}))). \end{aligned} \tag{3.17}$$

Denote $\mathbf{D}'$ a nova série diária de dados candidata ao lugar de $\mathbf{D}$. E por fim, o procedimento para a atualização de $\mathbf{D}$ segue,

1. Faça $\mathbf{D}' = \mathbf{D}$;

2. Seja $\mathbf{D}_a$ a série diária de dados que complementa a série diária de dados conhecida $\mathbf{D}_p$, faça $\mathbf{D}_a = \mathbf{D}'$ - $\mathbf{D}_p$;

3. Verifique quais são as posições não nulas da série diária de dados $\mathbf{D}_a$;

4. Sorteie uma posição dentre as posições não nulas de $\mathbf{D}_a$ e faça $D_a(t) = D_a(t) - 1$;

5. Sorteie uma posição da série diária de dados $D_a(t)$ para $8 \leq t \leq \tau^* - 1$ e faça $D_a(t) = D_a(t) + 1$;

6. Repita $nmodD$ vezes, os passos 3 - 5;

7. Calcule $\mathbf{D}' = \mathbf{D}_a + \mathbf{D}_p$;

8. Reconstrua as trajetórias $\{I'(t)\}$ a partir de (3.1);

9. Verifique se a série diária de dados $\mathbf{D}'$ satisfaz as condições 3, 4, 5, 6 e 7 listadas na subseção 3.3.1;

10. Aceite $\mathbf{D}'$ com a seguinte probabilidade:

$$\alpha(\mathbf{D}, \mathbf{D}') = min\{1, R_D\} \tag{3.18}$$

em que

$$\begin{aligned} R_D &= \frac{\pi(\mathbf{D}'|\mathbf{B}, \mathbf{C}, \boldsymbol{\theta})}{\pi(\mathbf{D}|\mathbf{B}, \mathbf{C}, \boldsymbol{\theta})} \\ &= \exp(log(\pi(\mathbf{D}'|\mathbf{B}, \mathbf{C}, \boldsymbol{\theta})) - log(\pi(\mathbf{D}|\mathbf{B}, \mathbf{C}, \boldsymbol{\theta}))). \end{aligned} \tag{3.19}$$

Note que o cálculo das razões nas probabilidades de aceitação é realizado considerando o logaritmo das distribuições marginais a posteriori para evitar problemas numéricos. Abaixo tais expressões,

$$\begin{aligned} log(\pi(\mathbf{B}|\mathbf{C}, \mathbf{D}, \boldsymbol{\theta})) \propto & \sum_{t=1}^{\tau-1} -S(t)\left(1 - \exp\left(-\frac{\beta}{N}I(t)\right)\right) + \\ & + B(t)log(S(t)) + B(t)log\left(1 - \exp\left(-\frac{\beta}{N}I(t)\right)\right) - \\ & - log(B(t)!) + \sum_{t=\tau}^{\tau^*-1} -S(t)\left(1 - \exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right) + \\ & + B(t)log(S(t)) + B(t)log\left(1 - \exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right) - \\ & - log(B(t)!) + \sum_{t=1}^{\tau^*-1} log\binom{E(t)}{C(t)} - \rho E(t), \end{aligned} \tag{3.20}$$

$$
\begin{aligned}
log(\pi(\mathbf{C}|\mathbf{B},\mathbf{D},\boldsymbol{\theta})) \;\propto\; & \sum_{t=1}^{\tau-1} S(t)\exp\left(-\frac{\beta}{N}I(t)\right) + \\
+ \;& B(t)log\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right) + \\
+ \;& \sum_{t=\tau}^{\tau^*-1} S(t)\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right) + \\
+ \;& B(t)log\left(\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right) + \\
+ \;& \sum_{t=1}^{\tau^*-1} log\binom{E(t)}{C(t)} + C(t)(1-e^{-\rho}) - \rho(E(t)-C(t)) + \\
+ \;& \sum_{t=1}^{\tau^*-1} log\binom{I(t)}{D(t)} - \gamma I(t), \qquad (3.21)
\end{aligned}
$$

$$
\begin{aligned}
log(\pi(\mathbf{D}|\mathbf{B},\mathbf{C},\boldsymbol{\theta})) \;\propto\; & \sum_{t=1}^{\tau-1} S(t)\exp\left(-\frac{\beta}{N}I(t)\right) + \\
+ \;& B(t)log\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right) + \\
+ \;& \sum_{t=\tau}^{\tau^*-1} S(t)\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right) + \\
+ \;& B(t)log\left(\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right) + \\
+ \;& \sum_{t=1}^{\tau^*-1} log\binom{I(t)}{D(t)} + D(t)log(1-e^{-\gamma}) - \gamma(I(t)-D(t)) \qquad (3.22)
\end{aligned}
$$

Além disso $nmodB$, $nmodC$ e $nmodD$ são determinados a partir da quantidade de dados faltantes, respectivamente, nas séries diárias de dados **B**, **C** e **D**. Tais valores influenciam na taxa de aceitação dos procedimentos de atualização e também nas misturas das cadeias de Markov.

Lekone e Finkenstädt (2006) sugerem que $nmodB$ seja igual a 10% do tamanho m da epidemia e nada falam sobre $nmodC$ e $nmodD$, todavia, esta escolha para $nmodB$resulta em baixíssimas taxas de aceitação para a série diária de dados **B** e o mesmo pode ser dito para **C** e **D**.

Desta forma, é necessário considerar diversos valores a fim de encontrar valores razoáveis para as taxas de aceitação assim como para os níveis de mistura das cadeias de Markov e este procedimento será realizado na seção 4.

3.3.4 Atualização de $\boldsymbol{\theta}$

A atualização de $\boldsymbol{\theta}$ pode ser feita através da utilização, mais uma vez, do Metropolis-Hastings por Blocos e de diversas formas pela escolha da dimensão dos blocos de parâmetros e das densidades propostas. Independente de tais escolhas, é necessário conhecer as distribuições marginais a posteriori de β, ρ, γ e q.

$$\begin{aligned}
\pi(\beta|\boldsymbol{\theta}_\beta, \mathbf{B}, \mathbf{C}, \mathbf{D}) &\propto \beta^{a_\beta - 1} e^{-b_\beta \beta} \prod_{t=1}^{\tau-1} \exp\left(S(t) \exp\left(-\frac{\beta}{N} I(t)\right)\right) \times \\
&\times \left(1 - \exp\left(-\frac{\beta}{N} I(t)\right)\right)^{B(t)} \times \\
&\times \prod_{t=\tau}^{\tau^*-1} \exp\left(S(t) \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right) \times \\
&\times \left(1 - \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right)^{B(t)} I(\beta)_{(0,\infty)}, \qquad (3.23) \\
\pi(\rho|\boldsymbol{\theta}_\rho, \mathbf{B}, \mathbf{C}, \mathbf{D}) &\propto \rho^{a_\rho - 1} e^{-b_\rho \rho} \prod_{t=1}^{\tau^*-1} (1 - e^{-\rho})^{C(t)} e^{-\rho(E(t) - C(t))} I(\rho)_{(0,\infty)}, \qquad (3.24) \\
\pi(\gamma|\boldsymbol{\theta}_\gamma, \mathbf{B}, \mathbf{C}, \mathbf{D}) &\propto \gamma^{a_\gamma - 1} e^{-b_\gamma \gamma} \prod_{t=1}^{\tau^*-1} (1 - e^{-\gamma})^{D(t)} e^{-\gamma(I(t) - D(t))} I(\gamma)_{(0,\infty)}, \qquad (3.25) \\
\pi(q|\boldsymbol{\theta}_q, \mathbf{B}, \mathbf{C}, \mathbf{D}) &\propto q^{a_q - 1} e^{-b_q q} \prod_{t=\tau}^{\tau^*-1} \exp\left(S(t) \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right) \times \\
&\times \left(1 - \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right)^{B(t)} I(q)_{(0,\infty)}. \qquad (3.26)
\end{aligned}$$

3.3.4.1 Densidades propostas univariadas

A primeira possibilidade é a utilização de densidades propostas univariadas. Então, o procedimento para atualização de λ para $\lambda \in \{\beta, \rho, \gamma, q\}$ baseado numa densidade

proposta normal univariada de passeio aleatório segue,

1. Calcule o ponto de máximo μ_λ de $log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))$;
2. Calcule a segunda derivada de $log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))$ avaliada em μ_λ;
3. Faça $\sigma^2_\lambda = -\dfrac{\partial^2 log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))}{\partial \lambda^2}$;
4. Faça $\sigma^2_\lambda = c_\lambda \sigma^2_\lambda$;
5. Seja λ' candidato ao lugar de λ, gere $\lambda' \sim q(\lambda, \lambda') \equiv N(\lambda, \sigma^2_\lambda)$;
6. Aceite λ' com a seguinte probabilidade:

$$\alpha(\lambda', \lambda) = min\{1, R_\lambda\} \tag{3.27}$$

em que

$$\begin{aligned} R_\lambda &= \frac{\pi(\lambda'|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})}{\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})} \\ &= \exp(log(\pi(\lambda'|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})) - log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))). \end{aligned} \tag{3.28}$$

Se uma densidade proposta normal univariada de independência for escolhida, os passos 5 e 6 do procedimento anterior são alterados para,

5. Seja λ' candidato ao lugar de λ, gere $\lambda' \sim q(\lambda') \equiv N(\mu_\lambda, \sigma^2_\lambda)$;
6. Aceite λ' com a seguinte probabilidade:

$$\alpha(\lambda', \lambda) = min\{1, R_\lambda\} \tag{3.29}$$

em que

$$\begin{aligned} R_\lambda &= \frac{\pi(\lambda'|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}) q(\lambda)}{\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}) q(\lambda')} \\ &= \exp(log(\pi(\lambda'|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})) - log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})) + \\ &+ log(q(\lambda)) - log(q(\lambda'))). \end{aligned} \tag{3.30}$$

Note que $log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))$ para $\lambda \in \{\beta, \rho, \gamma, q\}$ é necessário tanto no passo 1 como no passo 6 para evitar problemas numéricos. Deste modo, tem-se que,

$$
\begin{aligned}
log(\pi(\beta|\boldsymbol{\theta}_\beta, \mathbf{B}, \mathbf{C}, \mathbf{D})) &\propto (a_\beta - 1)log(\beta) - b_\beta\beta + \sum_{t=1}^{\tau-1} -S(t)\left(1 - \exp\left(-\frac{\beta}{N}I(t)\right)\right) + \\
&+ B(t)log\left(1 - \exp\left(-\frac{\beta}{N}I(t)\right)\right) + \\
&+ \sum_{t=\tau}^{\tau^*-1} -S(t)\left(1 - \exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right) + \\
&+ B(t)log\left(1 - \exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right), \qquad (3.31) \\
log(\pi(\rho|\boldsymbol{\theta}_\rho, \mathbf{B}, \mathbf{C}, \mathbf{D})) &\propto (a_\rho - 1)log(\rho) - b_\rho\rho + \sum_{t=1}^{\tau^*-1} C(t)log(1 - e^{-\rho}) - \\
&- \rho(E(t) - C(t)), \qquad (3.32) \\
log(\pi(\gamma|\boldsymbol{\theta}_\gamma, \mathbf{B}, \mathbf{C}, \mathbf{D})) &\propto (a_\gamma - 1)log(\gamma) - b_\gamma\gamma + \sum_{t=1}^{\tau^*-1} D(t)log(1 - e^{-\gamma}) - \\
&- \gamma(I(t) - D(t)), \qquad (3.33) \\
log(\pi(q|\boldsymbol{\theta}_q, \mathbf{B}, \mathbf{C}, \mathbf{D})) &\propto (a_q - 1)log(q) - b_q q + \\
&+ \sum_{t=\tau}^{\tau^*-1} -S(t)\left(1 - \exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right) + \\
&+ B(t)log\left(1 - \exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right). \qquad (3.34)
\end{aligned}
$$

E a segunda derivada de $log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))$ para $\lambda = \{\beta,\ \rho,\ \gamma,\ q\}$ é dada por,

$$
\begin{aligned}
\frac{\partial^2 log(\pi(\beta|\boldsymbol{\theta}_\beta, \mathbf{B}, \mathbf{C}, \mathbf{D}))}{\partial\beta^2} &\propto -\frac{a_\beta - 1}{\beta^2} + \sum_{t=1}^{\tau-1} \frac{S(t)\exp\left(-\frac{\beta}{N}I(t)\right)I(t)^2}{N^2} - \\
&- \frac{B(t)I(t)^2\exp\left(-\frac{\beta}{N}I(t)\right)}{N^2(1 - \exp\left(-\frac{\beta}{N}I(t)\right))^2} + \\
&+ \sum_{t=\tau}^{\tau^*-1} \frac{S(t)\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)e^{-2q(t-\tau)}I(t)^2}{N^2} - \\
&- \frac{B(t)I(t)^2\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)e^{-2q(t-\tau)}}{N^2(1 - \exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right))^2}, \qquad (3.35)
\end{aligned}
$$

$$\frac{\partial^2 log(\pi(\rho|\boldsymbol{\theta}_\rho,\mathbf{B},\mathbf{C},\mathbf{D}))}{\partial\rho^2} \propto -\frac{a_\rho-1}{\rho}+\sum_{t=1}^{\tau-1}-C(t)\frac{e^{-\rho}}{(1-e^{-\rho})^2}, \tag{3.36}$$

$$\frac{\partial^2 log(\pi(\gamma|\boldsymbol{\theta}_\gamma,\mathbf{B},\mathbf{C},\mathbf{D}))}{\partial\gamma^2} \propto -\frac{a_\gamma-1}{\gamma^2}+\sum_{t=1}^{\tau^*-1}-D(t)\frac{e^{-\gamma}}{(1-e^{-\gamma})^2}, \tag{3.37}$$

$$\begin{aligned}\frac{\partial^2 log(\pi(q|\boldsymbol{\theta}_q,\mathbf{B},\mathbf{C},\mathbf{D}))}{\partial q^2} &\propto -\frac{(a_q-1)}{q^2}+\\ &+ \sum_{t=\tau}^{\tau^*-1}\frac{S(t)\beta I(t)(t-\tau)^2\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)e^{-q(t-\tau)}}{N}\times\\ &\times \left(\frac{\beta e^{-q(t-\tau)}I(t)}{N}-1\right)+\\ &+ \frac{B(t)\beta I(t)(t-\tau)^2\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)e^{-q(t-\tau)}}{N(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right))}-\\ &- \frac{B(t)\beta^2 I(t)^2(t-\tau)^2\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)e^{-2q(t-\tau)}}{N^2(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right))^2}\end{aligned} \tag{3.38}$$

A utilização de tais procedimentos univariados ignora completamente a possibilidade de existência de correlação significativa entre as componentes do vetor $\boldsymbol{\theta}$ que pode reduzir a eficiência do algoritmo como discutido no Capítulo 2. Nesta dissertação, a solução escolhida para lidar com a correlação entre os parâmetros é a utilização de blocos que englobem os parâmetros correlacionados sendo que os blocos são construídos de tamanho máximo igual a 2.

3.3.4.2 Densidades propostas bivariadas

O procedimento para a atualização de $\Lambda = (\lambda_1, \lambda_2)$, isto é, a utilização de densidades propostas bivariadas é similar ao caso univariado. E neste texto, os blocos a serem construídos têm a segunda derivada mista da densidade posteriori conjunta igual a zero, consequentemente, a construção da densidade proposta bivariada é feita a partir das expressões univariadas com a inserção da covariância amostral entre λ_1 e λ_2.

O procedimento para uma densidade proposta normal bivariada de passeio aleatório é dado por,

1. Calcule o ponto de máximo μ_λ de $log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))$ para $\lambda \in \{\lambda_1, \lambda_2\}$;

2. Faça $\mu_\Lambda = (\mu_{\lambda_1}, \mu_{\lambda_2})$;

3. Calcule a segunda derivada de $log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))$ avaliada em μ_λ para $\lambda \in \{\lambda_1, \lambda_2\}$;

4. Faça $\sigma^2_\lambda = -\dfrac{\partial^2 log(\pi(\lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))}{\partial \lambda^2}$ para $\lambda \in \{\lambda_1, \lambda_2\}$;

5. Faça $\sigma^2_\lambda = c_\lambda \sigma^2_\lambda$ para $\lambda \in \{\lambda_1, \lambda_2\}$;

6. Faça

$$\Sigma_\Lambda = \begin{bmatrix} \sigma^2_{\lambda_1} & \sigma_{\lambda_1\lambda_2} \\ \sigma_{\lambda_2\lambda_1} & \sigma^2_{\lambda_2} \end{bmatrix}$$

7. Seja Λ' candidato ao lugar de Λ, gere $\Lambda \sim q(\Lambda, \Lambda') \equiv N(\Lambda, \Sigma_\Lambda)$;

8. Aceite Λ' com a seguinte probabilidade:

$$\alpha(\Lambda, \Lambda') = min\{1, R_\Lambda\} \tag{3.39}$$

em que

$$\begin{aligned} R_\Lambda &= \frac{\pi(\Lambda'|\boldsymbol{\theta}_\Lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})}{\pi(\Lambda|\boldsymbol{\theta}_\Lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})} \\ &= \exp(log(\pi(\Lambda'|\boldsymbol{\theta}_\Lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})) - log(\pi(\Lambda|\boldsymbol{\theta}_\Lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))). \end{aligned} \tag{3.40}$$

As expressões (3.31) a (3.34) são utilizadas para o cálculo da moda e as expressões (3.35) a (3.38) para a construção da matriz de variâncias e covariâncias. As covariâncias são definidas da seguinte forma,

$$\sigma_{\lambda_1\lambda_2} = \hat{\rho}_{\lambda_1\lambda_2} \sqrt{\sigma^2_{\lambda_1}} \sqrt{\sigma^2_{\lambda_2}}, \tag{3.41}$$

$$\sigma_{\lambda_2\lambda_1} = \sigma_{\lambda_1\lambda_2}. \tag{3.42}$$

Se uma densidade proposta normal bivariada de independência é escolhida, os passos 7 e 8 são modificados para,

7. Seja Λ' candidato ao lugar de Λ, gere $\Lambda \sim q(\Lambda, \Lambda') \equiv N(\mu_\Lambda, \Sigma_\Lambda)$;

8. Aceite Λ' com a seguinte probabilidade:

$$\alpha(\Lambda', \Lambda) = min\{1, R_\Lambda\} \tag{3.43}$$

em que

$$\begin{aligned} R_\Lambda &= \frac{\pi(\Lambda'|\boldsymbol{\theta}_\Lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})q(\Lambda)}{\pi(\Lambda|\boldsymbol{\theta}_\lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})q(\Lambda')} \\ &= \exp(log(\pi(\Lambda'|\boldsymbol{\theta}_\Lambda, \mathbf{B}, \mathbf{C}, \mathbf{D})) - log(\pi(\Lambda|\boldsymbol{\theta}_\Lambda, \mathbf{B}, \mathbf{C}, \mathbf{D}))) + \\ &+ log(q(\Lambda)) - log(q(\Lambda'))\}. \end{aligned} \tag{3.44}$$

3.3.4.3 Constantes multiplicativas da variância

Por último, as constantes multiplicativas das variâncias são escolhidas segundo a discussão das subseções 2.4.2 e 2.4.3 do Capítulo 2. Então, para as densidades propostas normais univariadas de passeio aleatório, define-se $c_\lambda = 2,38$ e para densidades propostas normais bivariadas, uma primeira tentativa é $c_\lambda = \frac{2,38}{2} = 1,19$. Caso esta tentativa não seja adequada e para as distribuições propostas normais de independência, utiliza-se a grade $\mathrm{C} = \{0,75; \ldots; 2,25\}$ para a escolha da constante.

3.4 Simulações

Para realizar as simulações desta seção, uma epidemia foi gerada segundo o modelo (3.1) - (3.4) a partir dos seguintes parâmetros escolhidos concordantes com a literatura: $\beta = 0,2$; $\rho = \frac{1}{6,1}$, $\gamma = \frac{1}{7,2}$ e $q = 0,2$. A epidemia simulada teve tamanho 72 com duração de 176 dias e medidas de intervenção introduzidas no dia 130 na tentativa de criar um cenário similar aos dados reais.

Além disso, o tamanho m da epidemia é uma quantidade aleatória e, portanto, possui uma distribuição de probabilidade. A precisão das estimativas dos parâmetros está intrisicamente ligada à probabilidade de se observar uma epidemia de tamanho m como é discutido em Manfredini (2009) para o modelo SIR. Por isso, a escolha da epidemia simulada também seguiu o critério de que o tamanho da epidemia simulada fosse típica, isto é, que não esteja localizada nas caudas da distribuição do tamanho m

da epidemia.

Para os parâmetros definidos nesta simulação, a distribuição do tamanho m da epidemia obtida via simulações segue na Figura 4.

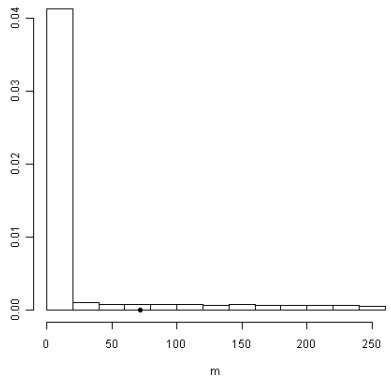


Figura 4: Distribuição do tamanho m da epidemia do modelo (3.1) - (3.4) até o quantil 0,75 a partir de 100000 simulações; o ponto representa a epidemia de tamanho igual a 72 indivíduos

A epidemia simulada é o quantil 0,655 e é representada pelo ponto na distribuição dada pela Figura 4, que foi construída a partir de 100000 simulações. Por isso, a epidemia escolhida pode ser considerada uma epidemia típica a partir dos parâmetros definidos, ou seja, com probabilidade significativa de ser observada.

3.4.1 Dados completos

O primeiro cenário considerado para as simulações é a presença de todos os dados, ou seja, as séries de dados **B**, **C** e **D** são totalmente conhecidas.

A primeira tentativa de implementação do algoritmo é através da utilização de densidades propostas normais univariadas de passeio aleatório para todos os parâmetros, pois a constante multiplicativa da variância é bem definida neste caso.

Note que a distribuição a priori considerada para as tentativas de implementação do algoritmo é a distribuição a priori não informativa e espera-se que a escolha mais eficiente para a distribuição a priori não informativa também seja a mesma ou, ao menos, similar para as outras distribuições a priori.

E, a partir de tal escolha, foi possível obter bons resultados, possivelmente, porque a estrutura amostral de correlação entre os parâmetros não foi significativa e deste modo, não há necessidade de tentar outras densidades propostas.

3.4.1.1 Análise de sensibilidade

As três distribuições a priori são consideradas. O tamanho da amostra de aquecimento é determinado a partir dos gráficos de médias amostrais e históricos. Após a definição do tamanho da amostra de aquecimento, a estatística R de Gelman e Rubin (1992 apud GAMERMAN; LOPES, 2006, Capítulo 5) é calculada para cada cadeia de Markov a fim de verificar se a convergência para a distribuição a posteriori foi alcançada, isto é, se R$<$ 2.

O passo seguinte consiste na escolha do espaçamento amostral, isto é, o número de observações da cadeia de Markov que devem ser gerados para que uma observação seja incluída na amostra da distribuição a posteriori.

Por fim, os gráficos das autocorrelações das amostras geradas para cada parâmetro e das distribuições a posteriori são apresentados assim como as estatísticas descritivas.

Para a distribuição a priori não informativa, os gráficos para a definição das condições da simulação são dados nas Figuras 53, 54 e 55. Para a distribuição informativa, as Figuras 56, 57 e 58 são consideradas. Por fim, para a distribuição a priori não centrada, as Figuras 59, 60 e 61. E os detalhes de cada simulação são apresentadas na Tabela 1.

Os resultados das simulações seguem nas Figuras 5, 6, 7, 8, 9 e 10, respectivamente, para as distribuições a priori não informativa, informativa e não centrada. As estatísticas descritivas estão na Tabela 2.

Tabela 1: Detalhes das simulações considerando dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

Priori	TAA	EA	T (segundos)
Não informativa	1000	5	56,21
Informativa	1000	5	52,31
Não centrada	1000	5	48,26
Parâmetros	**Estatística R**		**Taxas de aceitação (%)**
Priori não informativa			
β	1,001		45,24
ρ	1,000		44,84
γ	1,001		44,26
q	1,000		43,94
Priori informativa			
β	1,000		43,80
ρ	1,000		40,34
γ	1,001		41,16
q	1,001		45,24
Priori não centrada			
β	1,000		45,68
ρ	1,004		41,54
γ	1,002		41,16
q	1,005		44,86

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, T = Tempo

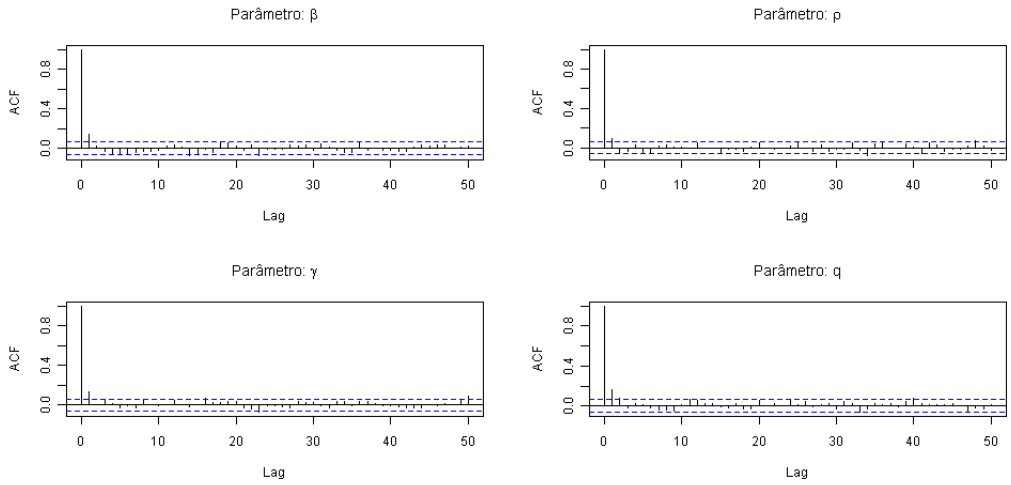


Figura 5: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

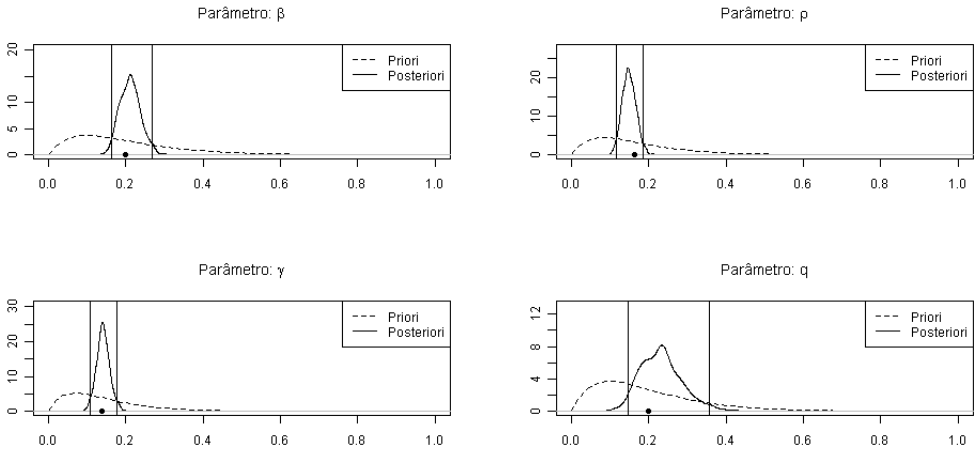


Figura 6: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

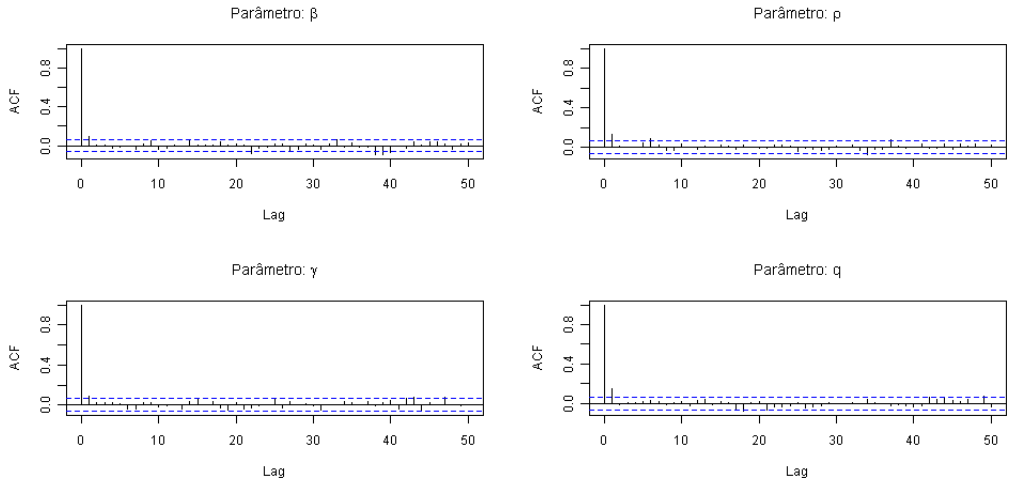


Figura 7: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

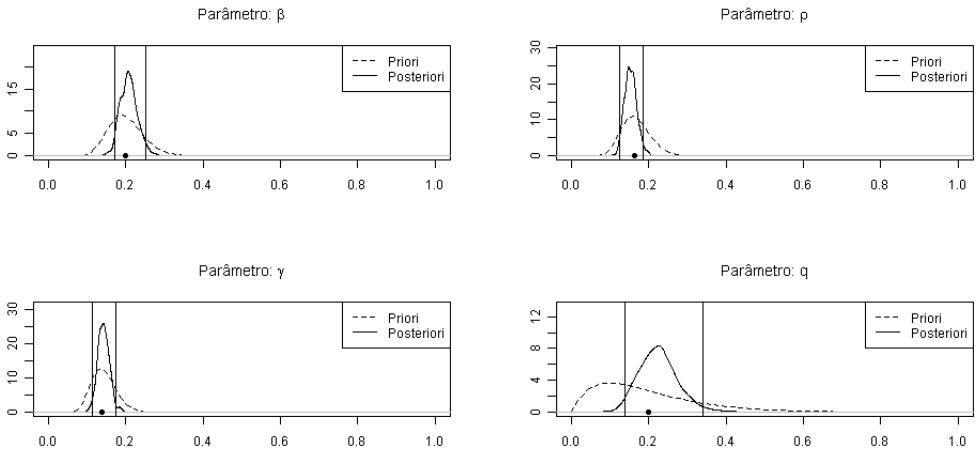


Figura 8: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

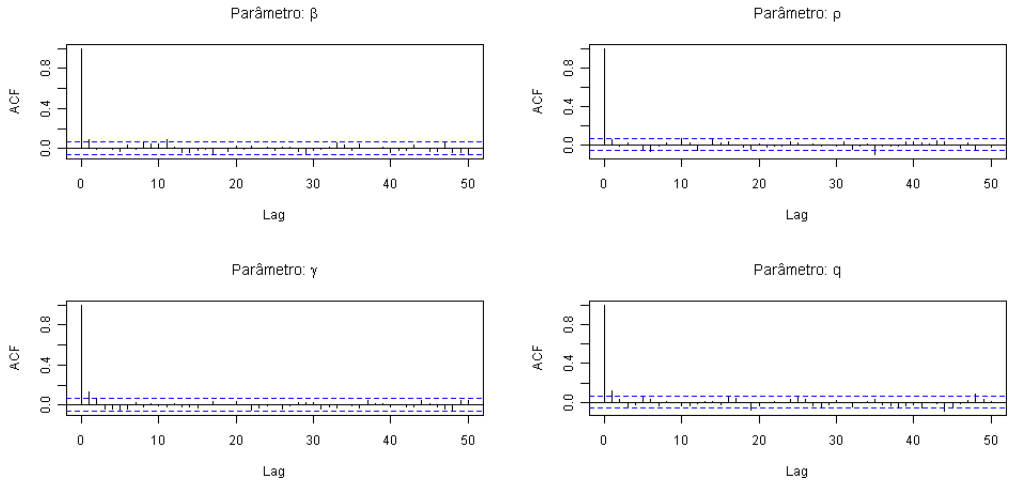


Figura 9: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não centrada com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

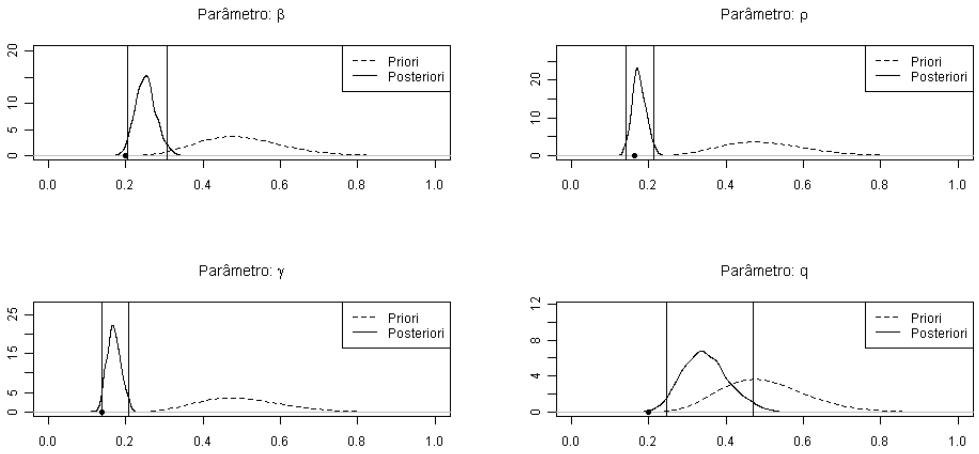


Figura 10: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não centrada com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

Tabela 2: Estatísticas descritivas das distribuições a posteriori considerando dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

β				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,211	0,026	(0,163 ; 0,267)	0,2
Informativa	0,208	0,022	(0,171 ; 0,253)	0,2
Não-Centrada	0,252	0,026	(0,204 ; 0,308)	0,2
ρ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,149	0,018	(0,115 ; 0,186)	0,164
Informativa	0,153	0,016	(0,124 ; 0,186)	0,164
Não-Centrada	0,175	0,018	(0,141 ; 0,213)	0,164
γ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,141	0,017	(0,108 ; 0,177)	0,139
Informativa	0,143	0,015	(0,114 ; 0,174)	0,139
Não-Centrada	0,169	0,018	(0,137 ; 0,208)	0,139
q				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,232	0,054	(0,146 ; 0,357)	0,2
Informativa	0,232	0,052	(0,139 ; 0,340)	0,2
Não-Centrada	0,347	0,059	(0,245 ; 0,470)	0,2

3.4.1.2 Conclusões

Os resultados apresentados junto às estatísticas descritivas da Tabela 2 permitem afirmar que o método de estimação apresenta boas estimativas considerando tanto distribuições a priori informativas quanto não informativas, como se pode ver pelas Figuras 6 e 8 desde que a distribuição a priori apresente uma probabilidade significativa numa vizinhança dos valores verdadeiros.

Se a distribuição a priori não atribui densidade significante ao redor dos valores verdadeiros, como no caso da utilização de uma distribuição a priori não centrada, o método é funcional, todavia não apresenta tão bons resultados. É importante ressaltar que o método de estimação ainda é razoável pois é capaz de aproximar as distribuições a posteriori do valor verdadeiro como se pode ver pela Figura 10.

Nos três casos, as distribuições a posteriori para β, ρ e γ apresentam pequenos

desvios padrão e o mesmo pode ser dito sobre o parâmetro q sob a consideração de que a informação dada pelos dados sobre q limitam ao período posterior a introdução das medidas de intervenção, ou seja, a quantidade de informação sobre q disponível é menor em relação aos outros parâmetros.

3.4.2 B Incompleto

O segundo cenário a ser considerado é na ausência quase total de informações sobre a série diária de dados $\mathbf{B}$, isto é, a ausência do registro das datas de quando os indivíduos deixaram de ser suscetíveis para se tornarem expostos.

Espera-se que o do vetor de dados $\mathbf{B}$ aumentado influenciará os parâmetros β, q por modificar a trajetória $S(t)$ e fortemente o parâmetro ρ por modificar a trajetória $E(t)$. O impacto em $E(t)$ é mais visível do que em $S(t)$ devido a magnitude do número de indivíduos registrados em $S(t)$ ser muito maior do que em relação a $B(t)$, diferentemente do que ocorre entre $E(t)$ e $B(t)$.

Além das constantes de variância a serem definidas independentemente das densidades propostas escolhidas, é necessário também definir $nmodB$ que consiste no número de modificações a serem feitas no candidato $\mathbf{B}'$ para $\mathbf{B}$ no procedimento de imputação dos dados. Para este conjunto de dados simulado, a sugestão de Lekone e Finkenstädt (2006) é $nmodB = 7$.

A primeira tentativa de implementação do algoritmo é através de densidades propostas normais univariadas de passeio aleatório, como no caso anterior, com $c_\lambda = 2,38$ para $\lambda = \beta, \rho, \gamma, q$ que é a escolha ótima segundo a literatura apresentada no Capítulo 2. Resta definir uma boa escolha para $nmodB$.

Então, o impacto da escolha de $nmodB$ deve ser observado nos resultados. Inicialmente, a taxa de aceitação do vetor $\mathbf{B}$ junto ao tempo computacional são apresentados na Tabela 3.

Os gráficos de autocorrelações para as amostras geradas dos parãmetros são dados nas Figuras 11, 12 e 13 considerando, respectivamente, $nmodB = 7, 5, 2$.

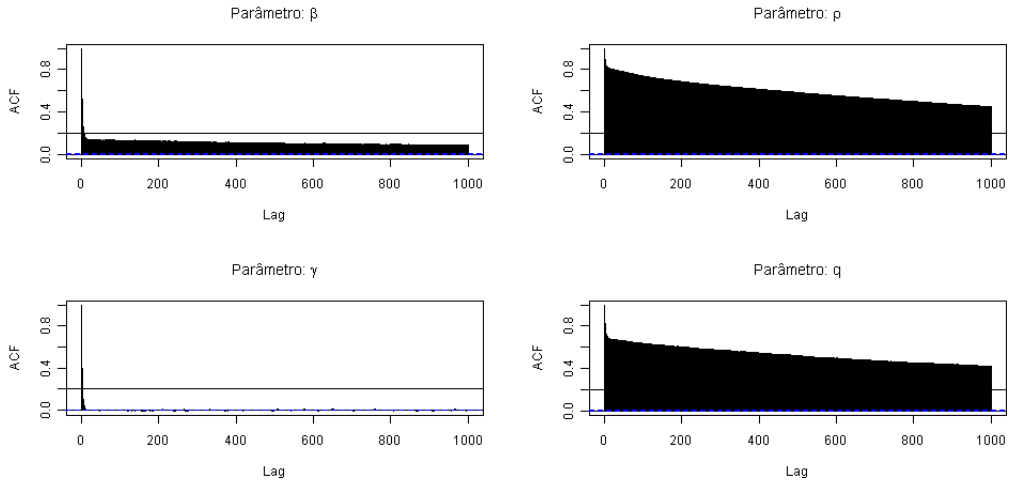


Figura 11: Autocorrelações sem espaçamento com $nmodB = 7$

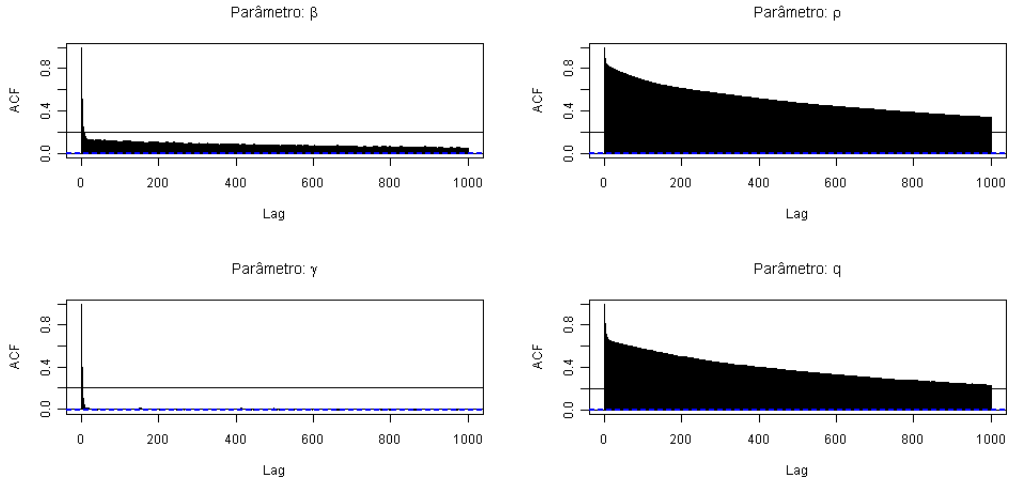


Figura 12: Autocorrelações sem espaçamento com $nmodB = 5$

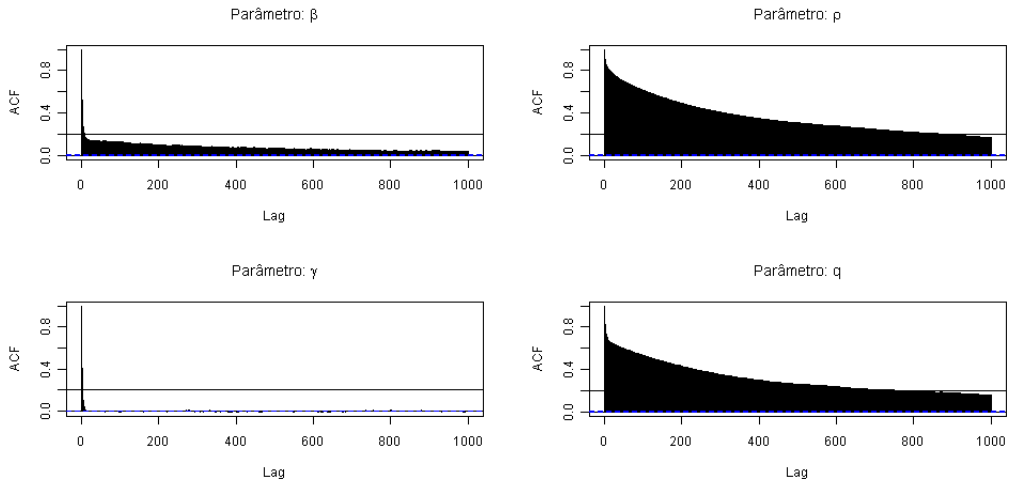


Figura 13: Autocorrelações sem espaçamento com $nmodB = 2$

Tabela 3: Taxa de aceitação e tempo computacional para diferentes valores de $nmodB$

$nmodB$	Taxa de aceitação	Tempo computacional (segundos)
7	$\approx 1\%$	7680,51
5	$\approx 3\%$	7181,54
2	$\approx 13\%$	6304,25

A escolha $nmodB = 1$ não é considerada porque a mistura das cadeias de Markov tornariam-se muito lentas segundo Lekone e Finkenstädt (2006), todavia, eles nada dizem sobre a taxa de aceitação da série de dados **B**. A partir de tais resultados, a escolha mais razoável é $nmodB = 2$ e a escolha de Lekone e Finkenstädt (2006) é descartada.

Dado $nmodB = 2$, o próximo passo é a busca por densidades propostas que tornem o algoritmo eficiente supondo que o impacto da escolha de $nmodB$ seja constante ou próximo de constante independentemente das escolhas das densidades propostas. Esta suposição é confirmada através das diversas tentativas de densidades propostas consideradas mais à frente.

Para a primeira tentativa, há dois pontos notáveis dentre os resultados: o lento decaimento da autocorrelação das amostras geradas dos parâmetros ρ e q que torna-se menor do que a reta 0,2 somente em torno do espaçamento 700 e a correlação negativa entre estes parâmetros na magnitude de 0,5 que surge pela introdução do vetor **B**

aumentado. Além disso, a correlação entre β e q torna-se maior, contudo, não tão significativa em relação a correlação negativa.

Então, a segunda tentativa de implementação é a utilização de densidades propostas univariadas de independência para ρ e q segundo a grade C para as constantes multiplicativas da variância e mantendo as densidades propostas univariadas de passeio aleatório para β e γ. Os resultados permanecem similares à primeira tentativa.

Desta forma, parece que a utilização de densidades propostas univariadas não é uma escolha adequada frente a estrutura de correlação observada. Para lidar com a autocorrelação dos parâmetros ρ e q, pode-se considerar uma densidade proposta bivariada de passeio aleatório para tais parâmetros e a escolha de densidades propostas para β e γ é mantida.

No caso bivariado de $\Lambda = (\rho, q)$, é necessário definir c_λ, a constante multiplicativa de variância, e também $Cov(\lambda_1, \lambda_2)$. A escolha de passeio aleatório é otimizada, segundo a literatura apresentada no Capítulo 2, por $c_\lambda = 2,38/2$ se a distribuição condicional completa de (ρ, q) é similar a distribuição normal, caso contrário, é necessário a utilização da grade C. Além disso, $Cov(\rho, q)$ pode ser definida a partir da equação (3.42), portanto, $Cov(\rho, q) = -0,5Var(\rho)Var(q)$.

Os resultados obtidos são um pouco melhores, as autocorrelações de ρ e q tornam menores que 0,2 em torno do espaçamento 550 e 600 sendo que o melhor resultado é quando $c_\lambda = 1,75$.

Ainda resta a possibilidade de uma escolha de densidade proposta bivariada de independência para (ρ, q). As mesmas considerações do caso anterior são válidas sobre c_λ e $Cov(\lambda_1, \lambda_2)$, porém, não é necessário testar $c_\lambda = 1,19$. Sob tais condições, os melhores resultados são obtidos para a escolha de $c_\lambda = 1,75$.

E na comparação entre as duas densidades propostas bivariadas, a escolha mais adequada é a utilização de uma densidade proposta bivariada de independência.

3.4.2.1 Análise de sensibilidade

As três distribuições a priori são consideradas. E o procedimento anterior é repetido.

Para a distribuição a priori não informativa, os gráficos para a definição das condições da simulação são dados na Figuras 62, 63 e 64. Para a distribuição informativa, as Figuras 65, 66 e 67 são consideradas. Por fim, para a distribuição a priori não centrada, as Figuras 68, 69 e 70. E os detalhes de cada simulação são apresentadas na Tabela 4.

Os resultados das simulações seguem nas Figuras 14, 15, 16, 17, 18 e 19, respectivamente, para as distribuições a priori não informativa, informativa e não centrada. As estatísticas descritivas estão na Tabela 5.

Tabela 4: Detalhes das simulações considerando **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

Priori	TAA	EA	T (horas)
Não informativa	100000	500	3,13
Informativa	100000	400	2,73
Não centrada	100000	600	3,72
Parâmetros	**Estatística R**		**Taxas de aceitação (%)**
Priori não informativa			
B	-		14,95
β	1,000		44,47
ρ	1,000		44,16*
γ	1,000		51,11
q	1,000		44,16*
Priori informativa			
B	-		13,77
β	1,000		44,57
ρ	1,002		40,83*
γ	1,000		47,36
q	1,001		40,83*
Priori não centrada			
B	-		10,24
β	1,000		44,54
ρ	1,000		40,89*
γ	1,000		47,66
q	1,000		40,89*

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, T = Tempo
*A taxa de aceitação é igual nos dois casos, pois é a taxa de aceitação conjunta.

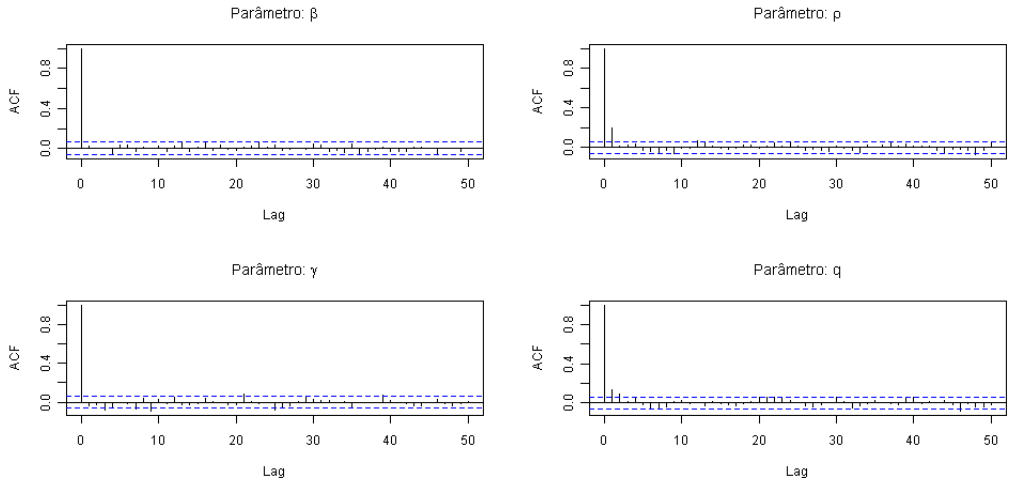


Figura 14: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

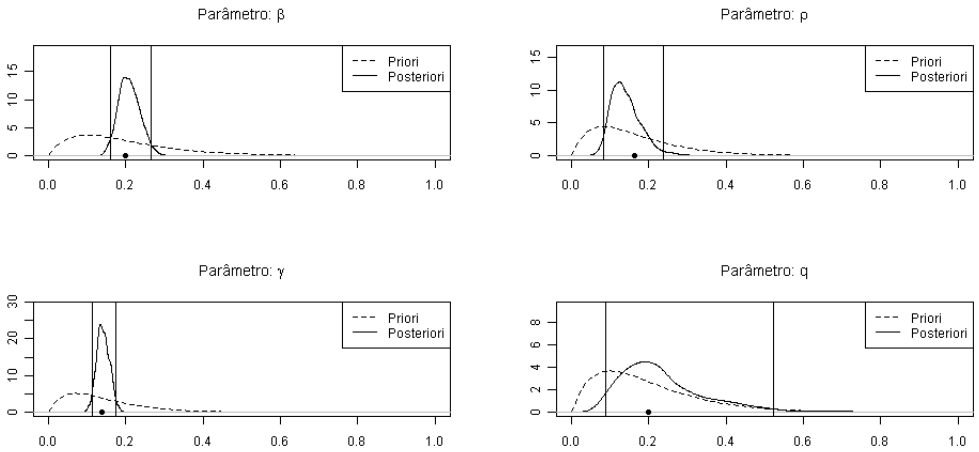


Figura 15: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

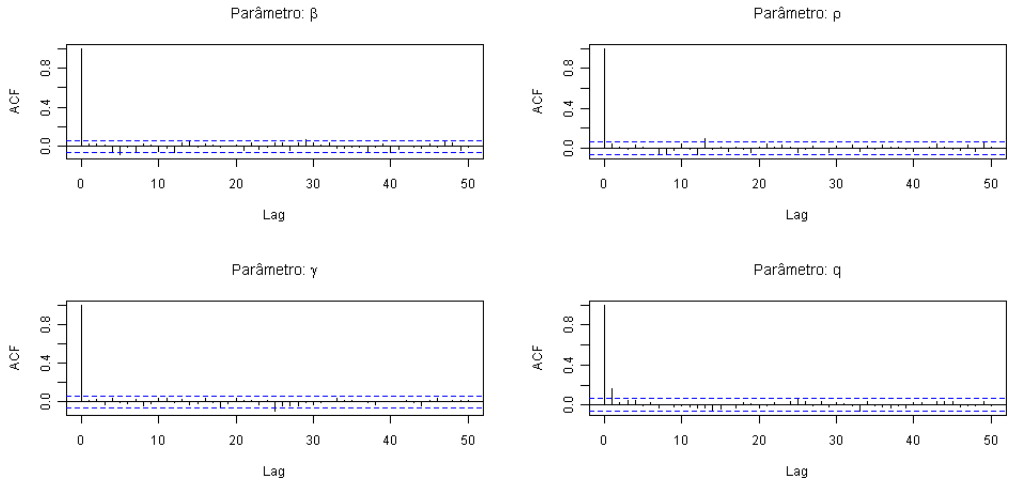


Figura 16: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

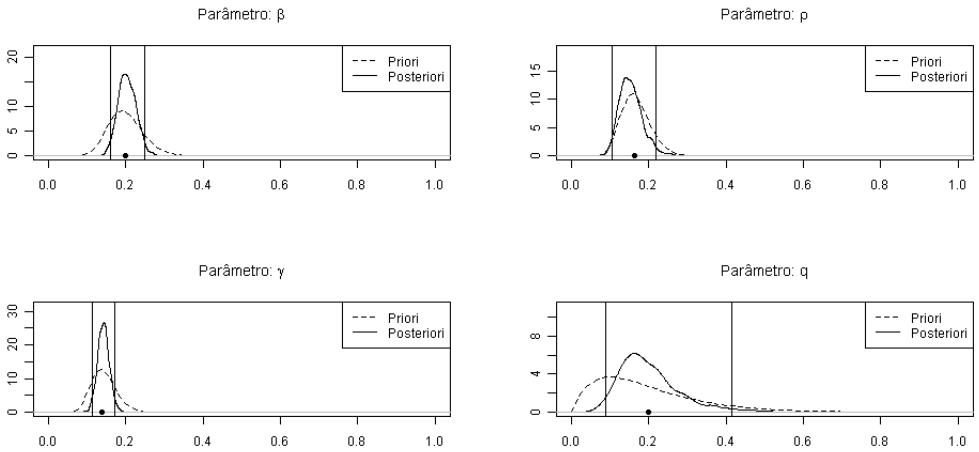


Figura 17: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

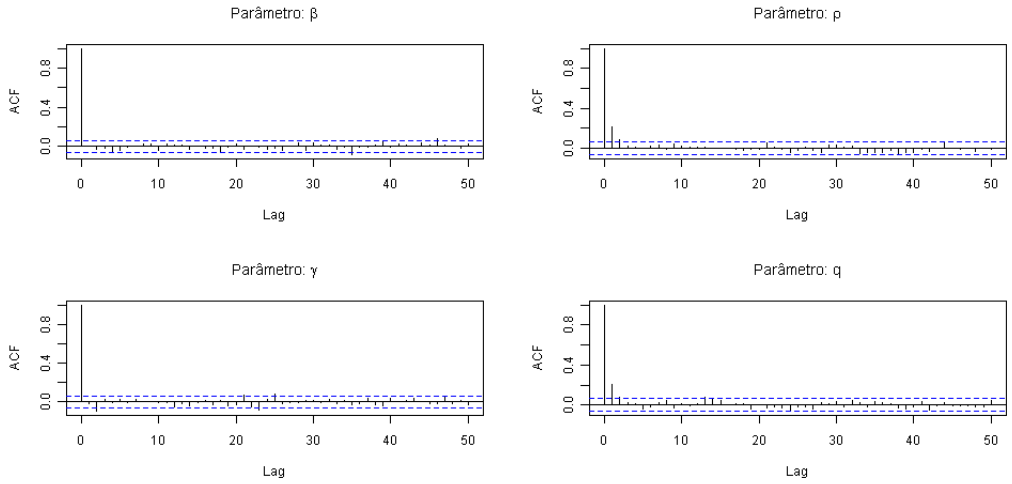


Figura 18: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não centrada com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

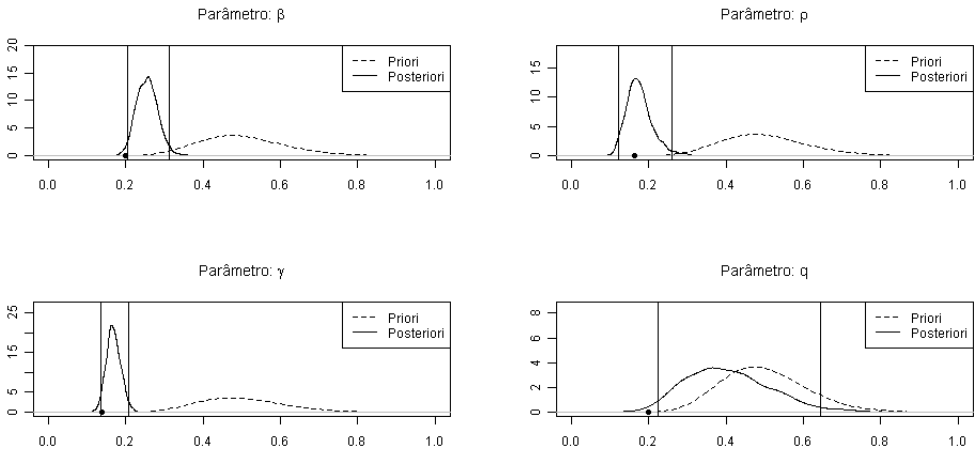


Figura 19: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não centrada com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

Tabela 5: Estatísticas descritivas das distribuições a posteriori considerando **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

β				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,209	0,027	(0,159 ; 0,264)	0,2
Informativa	0,203	0,023	(0,160 ; 0,248)	0,2
Não-Centrada	0,255	0,027	(0,205 ; 0,311)	0,2
ρ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,142	0,040	(0,083 ; 0,239)	0,164
Informativa	0,153	0,029	(0,105 ; 0,218)	0,164
Não-Centrada	0,176	0,034	(0,121 ; 0,259)	0,164
γ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,141	0,016	(0,112 ; 0,173)	0,139
Informativa	0,142	0,015	(0,113 ; 0,173)	0,139
Não-Centrada	0,168	0,019	(0,134 ; 0,206)	0,139
q				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,239	0,117	(0,088 ; 0,523)	0,2
Informativa	0,204	0,083	(0,090 ; 0,414)	0,2
Não-Centrada	0,407	0,111	(0,224 ; 0,644)	0,2

3.4.2.2 Conclusões

As conclusões para este segundo cenário são similares às conclusões do primeiro cenário em relação às estimativas pontuais. Cabe ressaltar o aumento do desvio padrão das estimativas de ρ e q que pode ser visto nas comparações entre as Tabelas 2 e 5, principalmente, para as distribuições a priori não informativa e não centrada. O aumento dos desvios padrão é devido à incorporação da incerteza sobre a série diária de dados **B**, como era esperado em relação aos parâmetros ρ e q.

Também, é interessante verificar que as distribuições a posteriori para q são muito similares às suas respectivas distribuições a priori. A única diferença notável, ainda pequena, é na utilização da distribuição a priori informativa observada como se pode ver nas Figuras 15, 17 e 19.

Outro ponto importante é o tempo computacional que se tornou notavelmente maior

na introdução do vetor aumentado **B**. Este tempo computacional deve-se à busca por candidatos razoáveis para **B**.

3.4.3 B, C e D Incompletos

O último cenário considerado é com dados faltantes em todas as séries de dados seguindo a proporção de 99,6%, 8% e 25% para, respectivamente, **B**, **C** e **D**. A utilização de **B** aumentado, como na seção anterior, irá influenciar a estimação de β, q e ρ, por sua vez, **C** aumentado influencia as estimativas de todos os parâmetros e, por fim, **D** influencia as estimativas de β, q e γ.

As constantes de variância, novamente, devem ser definidas assim como as constantes $nmodB$, $nmodC$ e $nmodD$. Por facilidade, $nmodB$ é mantido igual a 2 e $nmodC$, $nmodD$ são escolhidos também iguais a 2.

A primeira tentantiva de escolha das densidades propostas é a utilização de densidades univariadas de passeio aleatório. Mais uma vez, é observado o lento decaimento da correlação para os parâmetros ρ e q até a reta 0,2 e também, a correlação negativa entre estes parâmetros na magnitude de 0,5 além da correlação positiva entre β e γ na mesma magnitude.

Uma segunda tentativa é a utilização de densidades propostas univariadas de independência para ρ e q considerando a grade C para as constantes de variância. Há melhores resultados em relação à primeira tentativa sendo que a cadeia de Markov mais eficiente foi obtida com $c_\rho = c_q = 1,5$. A terceira possibilidade consiste na utilização de uma densidade proposta bivariada de independência de maneira similar à apresentada na seção anterior para (ρ, q), ou seja, considerando as mesmas constantes de variância. Os resultados são piores do que aqueles encontrados na segunda tentativa.

A quarta possibilidade foi a utilização de uma densidade proposta bivariada de independência para (ρ, q) considerando a constante de variância definida na seção anterior ($c_\rho = c_q = 1,75$) e a utilização de uma densidade proposta bivariada de passeio aleatório para (β, γ). Contudo, os resultados não foram satisfatórios.

Por fim, considerou-se densidades propostas bivariadas de independência para os

dois pares construídos na quarta tentativa, considerando ($c_\rho = c_q = 1,5$) devido aos resultados univariados e a grade C para (β, γ). Dentre todas possibilidades, os melhores resultados foram obtidos para $c_\rho = c_q = 1,75$ e $c_\beta = c_\gamma = 1,25$.

Entretanto, os resultados desta última tentativa foram bastante similares aqueles obtidos da segunda tentativa com a desvantagem de exigir um tempo computacional um pouco maior. Portanto, por simplicidade e tempo computacional, a configuração de densidades propostas utilizadas para este caso consiste na utilização de densidades propostas univariadas de passeio aleatório para β, γ e de independência com $c_\rho = c_q = 1,5$ para ρ e q.

3.4.3.1 Análise de sensibilidade

Para a distribuição a priori não informativa, os gráficos para a definição das condições da simulação são dados na Figuras 71, 72 e 73. Para a distribuição informativa, as Figuras 74, 75 e 76 são consideradas. Por último, para a distribuição a priori não centrada, as Figuras 77, 78 e 79. E os detalhes de cada simulação são apresentadas na Tabela 6.

Os resultados das simulações seguem nas Figuras 20, 21, 22, 23, 24 e 25, respectivamente, para as distribuições a priori não informativa, informativa e não centrada. As estatísticas descritivas estão na Tabela 7.

Tabela 6: Detalhes das simulações considerando **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

Priori	TAA	EA	T (horas)
Não informativa	100000	500	7,08
Informativa	100000	250	4,80
Não centrada	100000	800	15,92

Parâmetros	Estatística R	Taxas de aceitação (%)
Priori não informativa		
B	-	16,80
C	-	24,34
D	-	22,59
β	1,000	44,58
ρ	1,000	74,42
γ	1,000	44,23
q	1,000	75,19
Priori informativa		
B	-	16,04
C	-	26,51
D	-	21,80
β	1,000	44,49
ρ	1,002	68,15
γ	1,000	40,84
q	1,001	75,24
Priori não centrada		
B	-	9,54
C	-	28,20
D	-	9,73
β	1,012	44,53
ρ	1,001	68,10
γ	1,007	40,82
q	1,002	74,96

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, T = Tempo

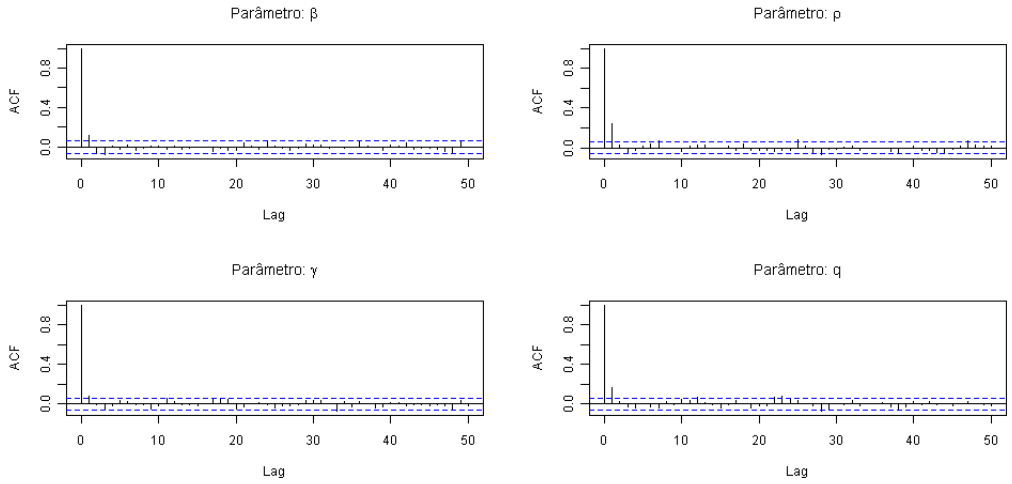


Figura 20: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

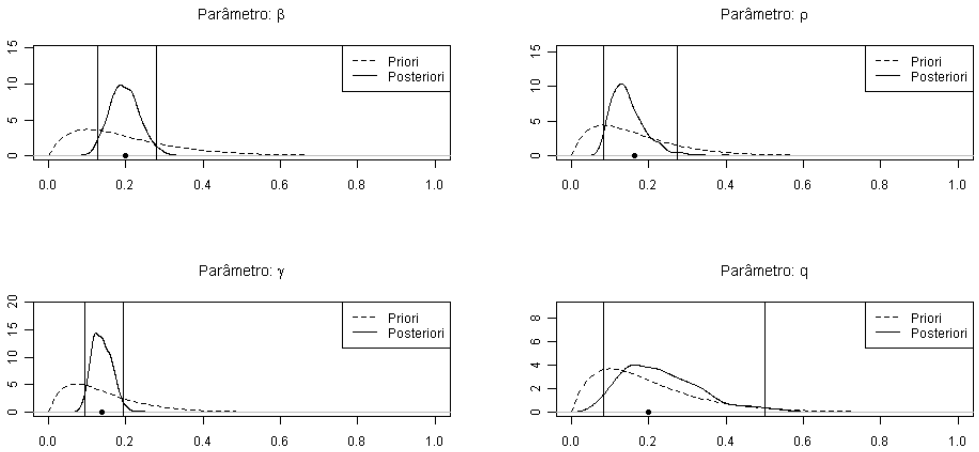


Figura 21: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

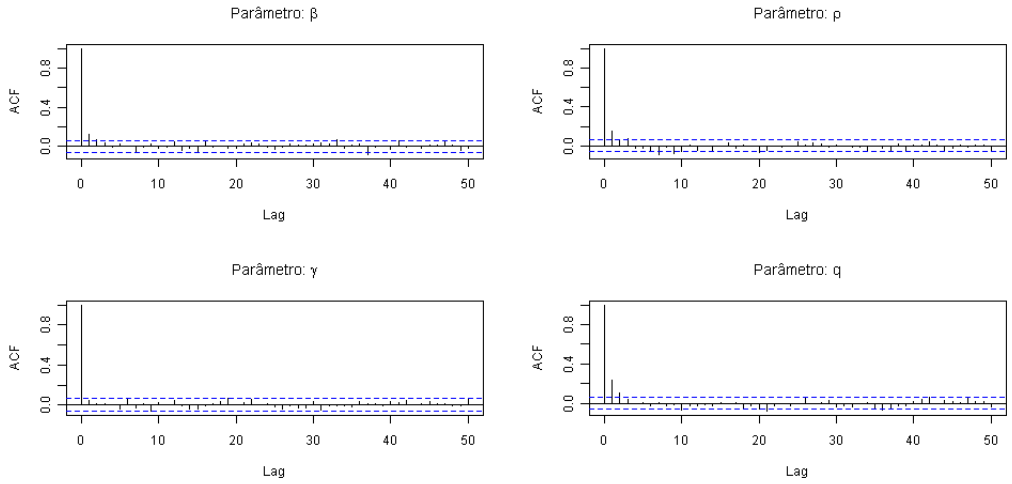


Figura 22: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

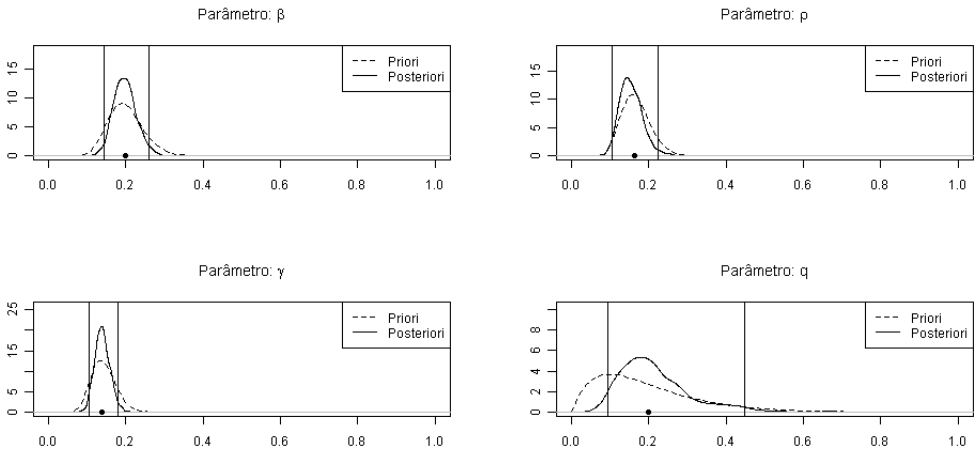


Figura 23: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

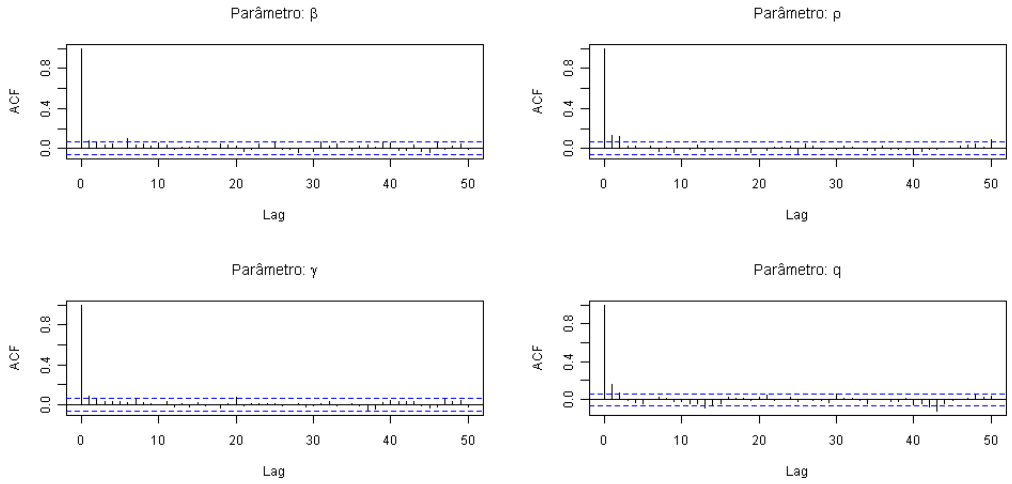


Figura 24: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não centrada com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

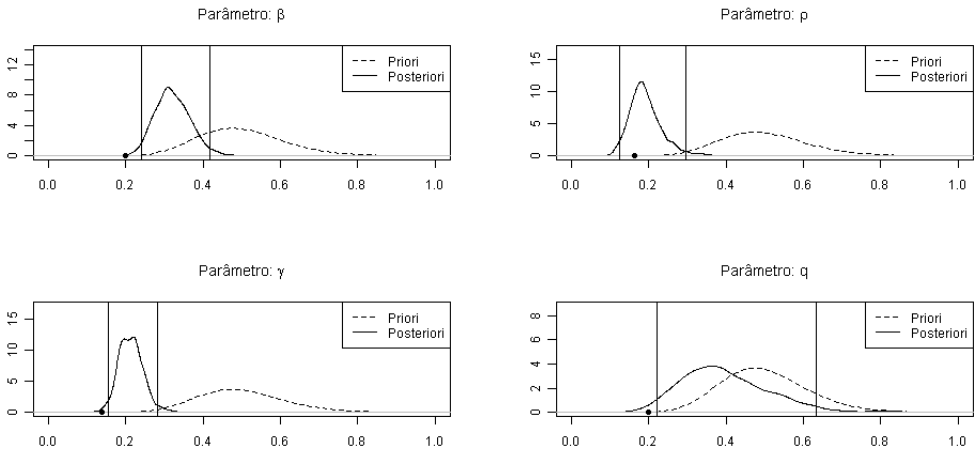


Figura 25: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

Tabela 7: Estatísticas descritivas das distribuições a posteriori considerando **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

β				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,197	0,039	(0,127 ; 0,279)	0,2
Informativa	0,198	0,029	(0,144 ; 0,259)	0,2
Não-Centrada	0,321	0,045	(0,241 ; 0,416)	0,2
ρ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,148	0,048	(0,083 ; 0,274)	0,164
Informativa	0,154	0,029	(0,105 ; 0,223)	0,164
Não-Centrada	0,192	0,043	(0,123 ; 0,296)	0,164
γ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,139	0,026	(0,093 ; 0,192)	0,139
Informativa	0,139	0,019	(0,104 ; 0,180)	0,139
Não-Centrada	0,213	0,031	(0,155 ; 0,282)	0,139
q				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,236	0,108	(0,082 ; 0,501)	0,2
Informativa	0,217	0,094	(0,093 ; 0,448)	0,2
Não-Centrada	0,397	0,111	(0,220 ; 0,633)	0,2

3.4.3.2 Conclusões

Os resultados sobre o método de estimação ainda são válidos para **B**, **C** e **D** incompletos como se vê na Tabela 7. Como nos casos anteriores, a informação a priori tem um impacto notável que pode ser visto pelos resultados considerando uma distribuição a priori informativa. Os resultados ainda são bons para uma distribuição a priori não informativa, todavia, não são tão razoáveis com uma distribuição a priori não centrada que atribui probabilidade próxima de zero na vizinhança dos valores verdadeiros dos parâmetros.

A incorporação da incerteza gerada por **C** e **D** é observada ao comparar as Tabelas 5 e 7. Note que todos os desvios padrão tornam-se notavelmente maiores com exceção do parâmetro q. Além disso, observe que as taxas de aceitação de **B**, **C** e **D** são quase constantes para as escolhas das distribuições a priori com exceção da escolha da

distribuição a priori não centrada em que há uma redução pequena para **B** e drástica para **D**.

Por fim, o tempo computacional torna-se ainda maior e também é diretamente proporcional a informação a priori.

4 *Modelo SEIR com taxa de remoção não-homogênea*

Este capítulo tem por objetivo apresentar a modificação do modelo discutido no Capítulo 3. O modelo SEIR estocástico discutido para a epidemia de febre hemorrágica Ebola, ocorrida na República Democrática do Congo em 1995, lida com dados incompletos e com a introdução de medidas de intervenção ao longo da epidemia. Além disso, Lekone e Finkenstädt (2006) consideram os períodos de incubação e remoção homogêneos no tempo, ou seja, tais períodos são modelados segundo uma distribuição *Exponencial*.

O artigo de Boys e Giles (2007) considera a distribuição *Exponencial por Partes* para o tempo de remoção e o artigo de Stretfaris e Gibson (2004) também discute modelos SEIR considerando tempos de transição não Exponenciais tais como Weibull ou Gumbel.

Neste sentido, este capítulo considera a distribuição *Exponencial por Partes* para a taxa de remoção e tal objetivo é alcançado ao considerar modelos com k pontos de mudança, isto é, o número de pontos de mudança é conhecido, mas suas respectivas posições são desconhecidas. A metodologia é desenvolvida para o caso geral, contudo, somente o caso $k = 1$ é considerado.

O capítulo apresenta a seguinte estrutura: Na seção 4.1, a modificação do modelo é apresentada; Nas seções 4.2 e 4.3, discute-se os procedimentos bayesianos para o modelo estudado, e por fim, a seção 4.4 é constituída de simulações.

4.1 O modelo

O modelo estocástico SEIR estudado neste capítulo é bastante similar ao anterior, considera-se tempo discretizado e o número de indivíduos como uma quantidade discreta sendo baseado no esquema dado na Figura 26.

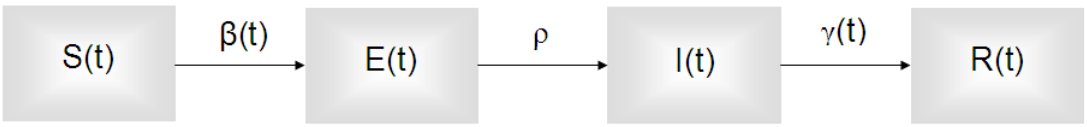


Figura 26: Modelo SEIR com taxa de remoção não homogênea

O modelo ainda é dado pelas equações (3.1) - (3.4) apresentadas no capítulo anterior sob as mesmas considerações. Entretanto, as equações (3.2) e (3.3) são modificadas pela distribuição da série de dados $D(t)$,

$$D(t)\dot{\sim}Binomial(I(t), p_D(t)), \tag{4.1}$$

em que

$$p_D(t) \;\; = \;\; 1 - exp(-\gamma(t)h). \tag{4.2}$$

O parâmetro $\gamma(t)$ representa a taxa média de remoção tal que $\dfrac{1}{\gamma(t)}$ é o período médio de remoção no período t.

A suposição de distribuição *Exponencial* para o período de remoção é frequentemente utilizada, pois basta conhecer o número de indivíduos em cada compartimento no período t. Esta suposição implica que todos os indivíduos têm a mesma taxa de transição entre os compartimentos e ela é constante no tempo durante todo o período da epidemia, todavia, esta suposição nem sempre é plausível pelas características da doença e a presença de medidas de intervenção.

A modificação necessária para eliminar tais implicações exige a utilização de distribuições diferentes da distribuição *Exponencial* e também a informação sobre o período de transição entre os compartimentos para cada indivíduo e não somente o número de indivíduos em cada compartimento no período t.

Esta última informação dificilmente é disponível, entretanto, ainda é possível eliminar a segunda implicação de que a taxa de transição entre determinados compartimentos seja constante no tempo durante todo o período da epidemia. Para tanto, pode-se utilizar a distribuição *Exponencial por Partes* apresentada no Capítulo 2.

O modelo de Lekone e Finkenstädt (2006) já considera a distribuição *Exponencial por Partes* para modelar o período de permanência dos indivíduos no compartimento dos Suscetíveis a fim de considerar as medidas de intervenção.

Contudo, a introdução de medidas de intervenção não tem impacto somente na taxa de contato entre os indivíduos suscetíveis e infectantes, mas também na taxa de remoção dos indivíduos, por isso, a probabilidade de transição entre os compartimentos dos infectantes e removidos, dada por $p_S(t)$, não é mais considerada constante no tempo a partir de,

$$\gamma(t) \quad = \quad \begin{cases} \gamma_0 & \text{se} \quad s_0 \leq t < s_1 \\ \gamma_1 & \text{se} \quad s_1 \leq t < s_2 \\ \vdots & \\ \gamma_k & \text{se} \quad s_k \leq t < s_{k+1}, \end{cases} \tag{4.3}$$

para k conhecido, $s_0 = 1$ e $s_{k+1} = \tau^* - 1$.

Note que as posições dos pontos de mudança dada por $s = (s_1, \ldots, s_k)$ são desconhecidas na equação (4.3), diferentemente da equação (3.4), a fim de tentar captar alterações no período de remoção provenientes de outras fontes além da introdução das medidas de intervenção no período τ, afinal, o impacto efetivo das medidas de intervenção pode ter sido posterior à introdução destas ou anterior já que medidas prévias poderiam estar em prática frente às suspeitas da epidemia.

Esta consideração poderia também ser feita para a taxa de contato $\beta(t)$, todavia, o desconhecimento quase total da série de dados $\mathbf{B}$ torna esta possibilidade inviável.

4.2 Estimação dos parâmetros

4.2.1 A função de verossimilhança

O modelo dado pelo conjunto de equações (3.1) - (3.4) adicionando as modificações dadas pelas equações (4.1) - (4.3) é definido a partir do conjunto de parâmetros $\boldsymbol{\theta} = \{\beta, \rho, q, \gamma, s\}$ sendo $\gamma = (\gamma_0, \ldots, \gamma_k)$, $s = (s_1, \ldots, s_k)$ e pode ser estimado a partir do conhecimento das condições iniciais, tamanho da população e as séries diárias {**B**, **C**, **D**}.

Além disso, note que $\{S(t),\ E(t),\ I(t)\}$ são completamente determinados a partir do sistema de equações (3.1) e como {**B**, **C**, **D**} são condicionalmente independentes, a função verossimilhança dos dados é definida por,

$$L(\mathbf{B}, \mathbf{C}, \mathbf{D}|\boldsymbol{\theta}) = \prod_{t=1}^{\tau^*-1} g_1(B(t)|\boldsymbol{\theta}, I_t) g_2(C(t)|\boldsymbol{\theta}, I_t) g_3(D(t)|\boldsymbol{\theta}, I_t), \tag{4.4}$$

onde g_1, g_2 e g_3 são as funções de probabilidades estabelecidas em (3.2) - (3.3) adicionando as modificações dadas pelas equações (4.1) - (4.3) condicionadas em $\boldsymbol{\theta}$ e toda informação disponível no período t, denotada por I_t. Então,

$$\begin{aligned}
L(\mathbf{B}, \mathbf{C}, \mathbf{D}|\boldsymbol{\theta}) &= \prod_{t=1}^{\tau-1} \binom{S(t)}{B(t)} \left(1 - \exp\left(-\frac{\beta}{N} I(t)\right)\right)^{B(t)} \exp\left(-\frac{\beta}{N} I(t)(S(t) - B(t))\right) \times \\
&\times \prod_{t=\tau}^{\tau^*-1} \binom{S(t)}{B(t)} \left(1 - \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right)^{B(t)} \times \\
&\times \exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)(S(t) - B(t))\right) \times \\
&\times \prod_{t=1}^{\tau^*-1} \binom{E(t)}{C(t)} (1 - e^{-\rho})^{C(t)} e^{-\rho(E(t)-C(t))} \times \\
&\times \prod_{i=0}^{k} \prod_{t=s_i}^{s_{i+1}-1} \binom{I(t)}{D(t)} (1 - e^{-\gamma_i})^{D(t)} e^{-\gamma_i(I(t)-D(t))}.
\end{aligned} \tag{4.5}$$

Pelas mesmas considerações feitas no capítulo anterior, é possível considerar a

função de verossimilhança aproximada,

$$\begin{aligned} L(\mathbf{B},\mathbf{C},\mathbf{D}|\boldsymbol{\theta}) &= \prod_{t=1}^{\tau-1}\frac{1}{B(t)!}\exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right)\times \\ &\times \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}I(t)\right)\right)\right]^{B(t)}\times \\ &\times \prod_{t=\tau}^{\tau^*-1}\frac{1}{B(t)!}\exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right)\times \\ &\times \left[S(t)\left(1-\exp\left(-\frac{\beta}{N}e^{-q(t-\tau)}I(t)\right)\right)\right]^{B(t)}\times \\ &\times \prod_{t=1}^{\tau^*-1}\binom{E(t)}{C(t)}(1-e^{-\rho})^{C(t)}e^{-\rho(E(t)-C(t))}\times \\ &\times \prod_{i=0}^{k}\prod_{t=s_i}^{s_{i+1}-1}\binom{I(t)}{D(t)}(1-e^{-\gamma_i})^{D(t)}e^{-\gamma_i(I(t)-D(t))}. \end{aligned} \tag{4.6}$$

E também é necessário considerar os dados aumentandos sendo que a função de verossimilhança para os dados aumentados é ainda dada pela expressão (4.6) e o próximo passo, segundo a abordagem bayesiana, é estabelecer distribuições a priori para $\boldsymbol{\theta}$.

4.2.2 As distribuições a priori

As considerações feitas no Capítulo 3 ainda podem ser utilizadas, desta forma, a distribuição a priori $Gama(a_\lambda, b_\lambda)$ é escolhida para $\lambda \in \{\beta, \rho, q\}$. E basta considerar a distribuição a priori não-informativa, isto é, $((a_\beta, b_\beta), (a_\rho, b_\rho), (a_q, b_q)) = ((2, 10), (2, 12), (2, 10))$ já que, a distribuição a priori não informativa não é concentrada em um intervalo tão restrito quanto a distribuição a priori informativa e é concordante com a informação sobre os parâmetros dada na literatura, diferentemente da distribuição a priori não centrada.

Para o vetor de taxas de remoção γ, considera-se que as componentes são variáveis aleatórias independentes e identicamente distribuídas e pelo capítulo anterior, uma distribuição $Gama(a_{\gamma_i}, b_{\gamma_i})$ é atribuída como distribuição a priori para a i-ésima com-

ponente, então,

$$p(\gamma) = \prod_{i=0}^{k} \frac{b_{\gamma_i}^{a_{\gamma_i}}}{\Gamma(a_{\gamma_i})} \gamma_i^{a_{\gamma_i}-1} e^{-b_{\gamma_i}\gamma_i} I_{(0,\infty)}(\gamma_i). \tag{4.7}$$

No caso de $k = 1$, define-se $a_{\gamma_0} = 2, b_{\gamma_0} = 14$ similarmente ao capítulo anterior e escolhe-se $a_{\gamma_1} = 4, b_{\gamma_1} = 14$ a partir da suposição de que as medidas de intervenção tem por objetivo a redução do período de remoção, ou seja, o aumento da taxa de remoção.

Para o vetor de pontos de mudança s, uma escolha de distribuição a priori é $U\{a_{s_i}, \ldots, b_{s_i}\}$ para cada componente do vetor s tal que $\cap_{i=1}^{k}\{a_{s_i}, \ldots, b_{s_i}\} = \emptyset$.

No caso de $k = 1$, há somente o ponto de mudança s_1, portanto, é possível atribuir duas distribuições a priori baseadas na distribuição Uniforme: a primeira distribuição a priori é uma distribuição não informativa sobre todo o período da epidemia, isto é, $s_1 \sim U\{s_0 + 1, s_2 - 1\}$; a segunda distribuição a priori utiliza a informação de que as medidas de intervenção foram introduzidas no dia 130 e uma possível escolha é dada por $U\{100, s_2 - 1\}$.

4.2.3 A distribuição a posteriori

A distribuição a posteriori para o conjunto de parâmetros $\boldsymbol{\theta}$ e os dados aumentados $\mathbf{B}$, $\mathbf{C}$ e $\mathbf{D}$ a partir de $\mathbf{B}_p$, $\mathbf{C}_p$ e $\mathbf{D}_p$ segue,

$$\pi(\boldsymbol{\theta}, \mathbf{B}, \mathbf{C}, \mathbf{D} | \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) \propto L(\mathbf{B}, \mathbf{C}, \mathbf{D} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \tag{4.8}$$

Então,

$$
\begin{aligned}
\pi(\boldsymbol{\theta}, \mathbf{B}, \mathbf{C}, \mathbf{D} | \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) \propto & \prod_{t=1}^{\tau-1} \frac{1}{B(t)!} \exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N} I(t)\right)\right)\right) \times \\
& \times \left[S(t)\left(1-\exp\left(-\frac{\beta}{N} I(t)\right)\right)\right]^{B(t)} \times \\
& \times \prod_{t=\tau}^{\tau^*-1} \frac{1}{B(t)!} \exp\left(-S(t)\left(1-\exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right)\right) \times \\
& \times \left[S(t)\left(1-\exp\left(-\frac{\beta}{N} e^{-q(t-\tau)} I(t)\right)\right)\right]^{B(t)} \times \\
& \times \prod_{t=1}^{\tau^*-1} \binom{E(t)}{C(t)} (1-e^{-\rho})^{C(t)} e^{-\rho(E(t)-C(t))} \times \\
& \times \prod_{i=0}^{k} \prod_{t=s_i}^{s_{i+1}-1} \binom{I(t)}{D(t)} (1-e^{-\gamma_i})^{D(t)} e^{-\gamma_i (I(t)-D(t))} \times \\
& \times \beta^{a_\beta - 1} e^{-b_\beta \beta} \rho^{a_\rho - 1} e^{-b_\rho \rho} q^{a_q - 1} e^{-b_q q} \times \\
& \times \prod_{i=0}^{k} \frac{b_\gamma^{a_\gamma}}{\Gamma(a_\gamma)} \gamma_i^{a_\gamma - 1} e^{-b_\gamma \gamma_i} \times \\
& \times I(\beta)_{(0,\infty)} I(\rho)_{(0,\infty)} I(q)_{(0,\infty)} \times \\
& \times \prod_{i=0}^{k} I_{(0,\infty)}(\gamma_i) \prod_{i=1}^{k} I_{\{a_{s_1}, \ldots, b_{s_1}\}}(s_i). \qquad (4.9)
\end{aligned}
$$

Como a distribuição a posteriori (4.9) não é simples de ser trabalhada analiticamente, a saída é a utilização de métodos computacionais MCMC.

4.3 Métodos MCMC

O método utilizado para a estimação dos parâmetros e simulação dos dados aumentados é, novamente, o Metropolis-Hastings por Blocos. O algoritmo geral é o mesmo apresentado na seção 3.3 do Capítulo 3 assim como também a inicialização de **B**, **C** e **D** dada na subseção 3.3.1 e a atualização de **B**, **C** e **D** na subseção 3.3.3.

4.3.1 Inicialização de $\boldsymbol{\theta}$

A inicialização de $\boldsymbol{\theta}$ também pode ser feita do mesmo modo que no Capítulo 3 para os parâmetros β, ρ e q, isto é, $\beta_0 = 0,5$, $\rho_0 = \dfrac{1}{6,1}$ e $q_0 = 0,5$

O vetor γ tem dimensão $(k+1)\times 1$, no caso de k = 1, $\gamma_{00} = \dfrac{1}{7,2}$ pela mesma discussão do capítulo anterior e $\gamma_{10} = \dfrac{2}{7,2}$, pois é esperado que as medidas de intervenção aumente a taxa de remoção. Por último, o vetor s tem dimensão $(k+2)\times 1$ e para o caso de $k = 1$, uma escolha coerente é $s = (0, \tau, \tau^* - 1)$ já que as medidas de intervenção foram introduzidas no período τ.

4.3.2 Atualização de $\boldsymbol{\theta}$

A atualização de $\boldsymbol{\theta}$ segue somente parcialmente o procedimento apresentado no capítulo anterior. Mais uma vez, é necessário a utilização do algoritmo Metropolis-Hastings por Blocos e a implementação pode ser feita de diversas formas.

Para β, ρ e q, as discussões acerca das densidades propostas são as mesmas, contudo, tais discussões não são diretamente aplicadas na atualização de γ e s.

A atualização de γ e s, ou seja, da função escada $\gamma(t)$ é realizada por partes,

- Atualizar γ;
- Atualizar s;

4.3.2.1 Atualização de $\boldsymbol{\gamma}$

A atualização de γ é construída componente a componente e, por isso, é bastante similar ao capítulo anterior e é possível aplicar as discussões consideradas sobre densidades propostas univariadas para cada componente.

- Atualizar γ_0;

  $\vdots$

- Atualizar γ_k;

Para tanto, é necessário conhecer a densidade condicional completa de γ,

$$\begin{aligned}\pi(\gamma|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) &\propto \prod_{i=0}^{k}\prod_{t=s_i}^{s_{i+1}-1}\binom{I(t)}{D(t)}(1-e^{-\gamma_i})^{D(t)}e^{-\gamma_i(I(t)-D(t))} \times \\ &\times \prod_{i=0}^{k}\frac{b_\gamma^{a_\gamma}}{\Gamma(a_\gamma)}\gamma_i^{a_\gamma-1}e^{-b_\gamma\gamma_i} \times \\ &\times \prod_{i=0}^{k} I_{(0,\infty)}(\gamma_i). \end{aligned} \tag{4.10}$$

Seguindo a discussão do Capítulo 3, é necessário maximizar $log(\pi(\gamma|\mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p, \boldsymbol{\theta}_\gamma))$ em relação a γ_j,

$$\begin{aligned} log(\pi(\gamma|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) &\propto \sum_{i=0}^{k}\sum_{t=s_i}^{s_{i+1}-1} log\binom{I(t)}{D(t)} + D(t)log(1-e^{-\gamma_i}) - \\ &- \gamma_i(I(t)-D(t)) + \sum_{i=0}^{k} a_\gamma log(b_\gamma) - log(\Gamma(a_\gamma)) + \\ &+ (a_\gamma - 1)log(\gamma_i) - b_\gamma\gamma_i. \end{aligned} \tag{4.11}$$

E a razão para o cálculo da probabilidade de aceitação de γ' que apresenta a componente γ'_j pode ser bastante simplificada, pois somente uma componente de γ é atualizada por vez,

$$\begin{aligned} R_\gamma &= \frac{\pi(\gamma'|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)q(\gamma', \gamma)}{\pi(\gamma|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)q(\gamma, \gamma')} \\ &= \exp(log(\pi(\gamma'|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) - log(\pi(\gamma|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) + \\ &+ log(q(\gamma', \gamma)) - log(q(\gamma, \gamma'))), \end{aligned} \tag{4.12}$$

sendo que

$$\begin{aligned} & log(\pi(\gamma'|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) - log(\pi(\gamma|\boldsymbol{\theta}_\gamma, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) = \\ = & \sum_{t=s_j}^{s_{j+1}-1} D(t)(log(1-e^{-\gamma'_j}) - log(1-e^{-\gamma_j})) - (\gamma'_j - \gamma_j)(I(t) - D(t)) + \\ + & (a_\gamma - 1)(log(\gamma'_j) - log(\gamma_j)) - b_\gamma(\gamma'_j - \gamma_j). \end{aligned} \tag{4.13}$$

A simplicação ocorre ao considerar que $\gamma'_i = \gamma_i \quad \forall i \neq j$.

4.3.2.2 Atualização de s

A atualização de s também é realizada componente por componente,

- Atualizar s_1;
  $\vdots$
- Atualizar s_k;

Na atualização da componente s_j para $j \in \{1, \ldots, k\}$, um candidato s'_j é gerado a partir de uma função de probabilidade de uma variável aleatória $U\{a_{s_j}, \ldots, b_{s_j}\}$, ou seja, é considerada uma densidade proposta de independência.

O cálculo da probabilidade de aceitação de s' que contém a componente s'_j exige a densidade a posteriori de s,

$$\begin{aligned} \pi(s|\boldsymbol{\theta}_s \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) \quad \propto & \quad \prod_{i=0}^{k} \prod_{t=s_i}^{s_{i+1}-1} \binom{I(t)}{D(t)} (1-e^{-\gamma_i})^{D(t)} e^{-\gamma_i(I(t)-D(t))} \times \\ & \times \quad \prod_{i=1}^{k} I_{\{a_{s_i}, \ldots, b_{s_i}\}}(s_i). \end{aligned} \tag{4.14}$$

E o logaritmo de $\pi(s|\boldsymbol{\theta}_s \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)$,

$$\begin{aligned} log(\pi(s|\boldsymbol{\theta}_s \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) \quad \propto & \quad \sum_{i=0}^{k} \sum_{t=s_i}^{s_{i+1}-1} log\binom{I(t)}{D(t)} + D(t) log(1-e^{-\gamma_i}) - \\ & - \quad \gamma_i(I(t) - D(t)). \end{aligned} \tag{4.15}$$

A razão para o cálculo da probabilidade de aceitação também pode ser simplificada,

$$\begin{aligned} R_s &= \frac{\pi(s'|\boldsymbol{\theta}_s, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) q(s', s)}{\pi(s|\boldsymbol{\theta}_s, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p) q(s, s')} \\ &= \exp(log(\pi(s'|\boldsymbol{\theta}_s, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) - log(\pi(s|\boldsymbol{\theta}_s, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) + \\ &+ log(q(s', s)) - log(q(s, s'))). \end{aligned} \tag{4.16}$$

Note que

$$q(s', s) = q(s, s') \tag{4.17}$$

Então,

$$\begin{aligned} &log(\pi(s'|\boldsymbol{\theta}_s, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) - log(\pi(s|\boldsymbol{\theta}_s, \mathbf{B}_p, \mathbf{C}_p, \mathbf{D}_p)) = \\ &= \sum_{i=j-1}^{j} \sum_{t=s'_i}^{s'_{i+1}-1} log\binom{I(t)}{D(t)} + D(t) log(1 - e^{-\gamma_i}) - \gamma_i (I(t) - D(t)) - \\ &- \sum_{i=j-1}^{j} \sum_{t=s_i}^{s_{i+1}-1} log\binom{I(t)}{D(t)} + D(t) log(1 - e^{-\gamma_i}) - \gamma_i (I(t) - D(t)). \end{aligned} \tag{4.18}$$

a partir do fato de que $s'_i = s_i \quad \forall i \neq j$.

4.4 Simulações

A fim realizar as simulações desta seção, uma epidemia foi gerada segundo o modelo (3.1) - (3.4) adicionando as modificações dadas pelas equações (4.1) - (4.3) a partir dos seguintes parâmetros: $\beta = 0,2$; $\rho = \frac{1}{6,1} \approx 0,1639$, $q = 0,2$, $\gamma_1 = \frac{1}{7,2}$, $\gamma_2 = \frac{2}{7,2}$ e $s_1 = 130$. A epidemia simulada teve tamanho 121 com duração de 165 dias e medidas de intervenção introduzidas no dia 130.

Novamente, segundo Manfredini (2009), escolheu-se uma epidemia típica, isto é, não estar localizada nas caudas da distribuição do tamanho m da epidemia.

Para os parâmetros definidos nesta simulação, a distribuição do tamanho m da epidemia obtida via simulações segue na Figura 27.

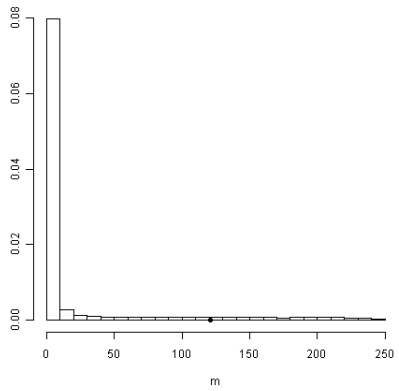


Figura 27: Distribuição do tamanho m da epidemia do modelo (3.1) - (3.4) com as modificações dadas por (4.1) - (4.3) até o quantil 0,75 através de 100000 simulações ; o ponto representa a epidemia escolhida

A epidemia simulada é o quantil 0,684 e é representada pelo ponto na distribuição dada pela Figura 4, que foi construída a partir de 100000 epidemias simuladas. Portanto, pode ser considerada uma epidemia típica, ou seja, com uma probabilidade significativa de ser observada a partir dos parâmetros definidos.

4.4.1 Dados completos

O primeiro cenário a ser considerado é a presença de todos os dados, ou seja, as séries de dados **B**, **C** e **D** são conhecidas.

A implementação do algoritmo foi realizada da mesma maneira do capítulo anterior com as considerações adicionais sobre γ e s, portanto, densidades normais propostas univariadas de passeio aleatório foram consideradas devido a constante multiplicativa da variância ser bem definida, $c_\lambda = 2{,}38$ para β, ρ, q, γ_1 e γ_2 e uma densidade proposta de independência dada pela distribuição a priori de s_1 para s_1.

Os resultados observados são razoáveis, apesar da estrutura de correlação entre β e q apresentar magnitude significativa de 0,5.

4.4.1.1 Análise de sensibilidade

As duas distribuições a priori apresentadas são consideradas. Note que as duas distribuições somente são diferenciadas pela distribuição do ponto de mudança s_1, a saber: não informativa e informativa. Portanto, as duas distribuições a priori são denotadas por distribuição a priori não informativa e informativa. Note que a nomenclatura é a mesma também no Capítulo 3, contudo, neste capítulo refere-se a objetos diferentes.

O procedimento descrito no capítulo anterior para a realização das simulações é o mesmo. Para a distribuição a priori não informativa, os gráficos para a definição das condições da simulação são dados na Figuras 80, 81 e 82. Para a distribuição informativa, as Figuras 83, 84 e 85 são consideradas. Os detalhes da simulação estão na Tabela 8.

Os resultados das simulações seguem nas Figuras 28, 29, 30, 31, respectivamente, para as distribuições a priori não informativa e informativa. As estatísticas descritivas estão na Tabela 9.

Tabela 8: Detalhes das simulações considerando dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

Priori	TAA	EA	TC (segundos)
Não informativa	8000	30	330,80
Informativa	8000	15	138,46
Parâmetros	**Estatística R**		**Taxas de aceitação (%)**
Priori não informativa			
β	1,001		44,35
ρ	1,001		47,06
q	1,004		47,16
γ_1	1,009		43,35
γ_2	1,002		43,58
s_1	1,001		17,21
Priori informativa			
β	1,000		44,25
ρ	1,000		46,71
q	1,001		47,21
γ_1	1,000		43,99
γ_2	1,001		43,15
s_1	1,000		25,75

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, TC = Tempo

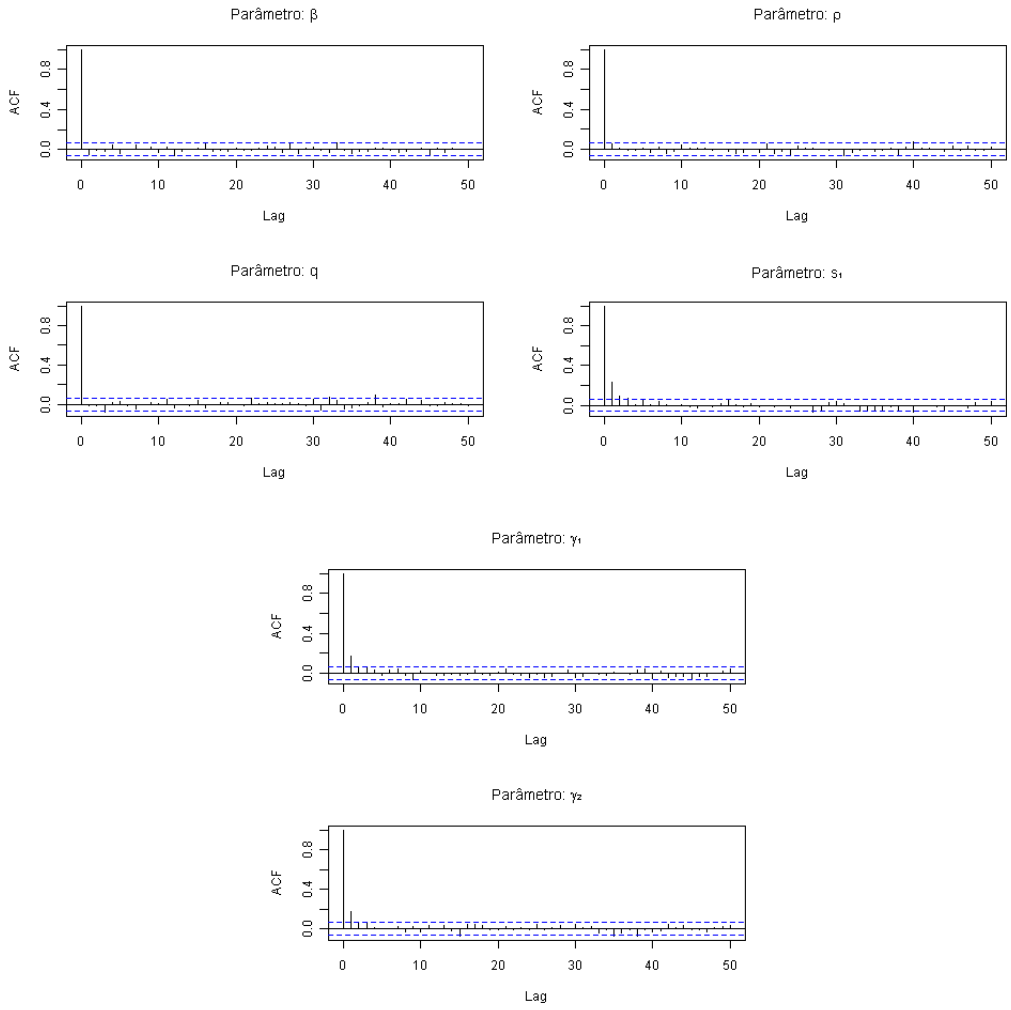


Figura 28: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

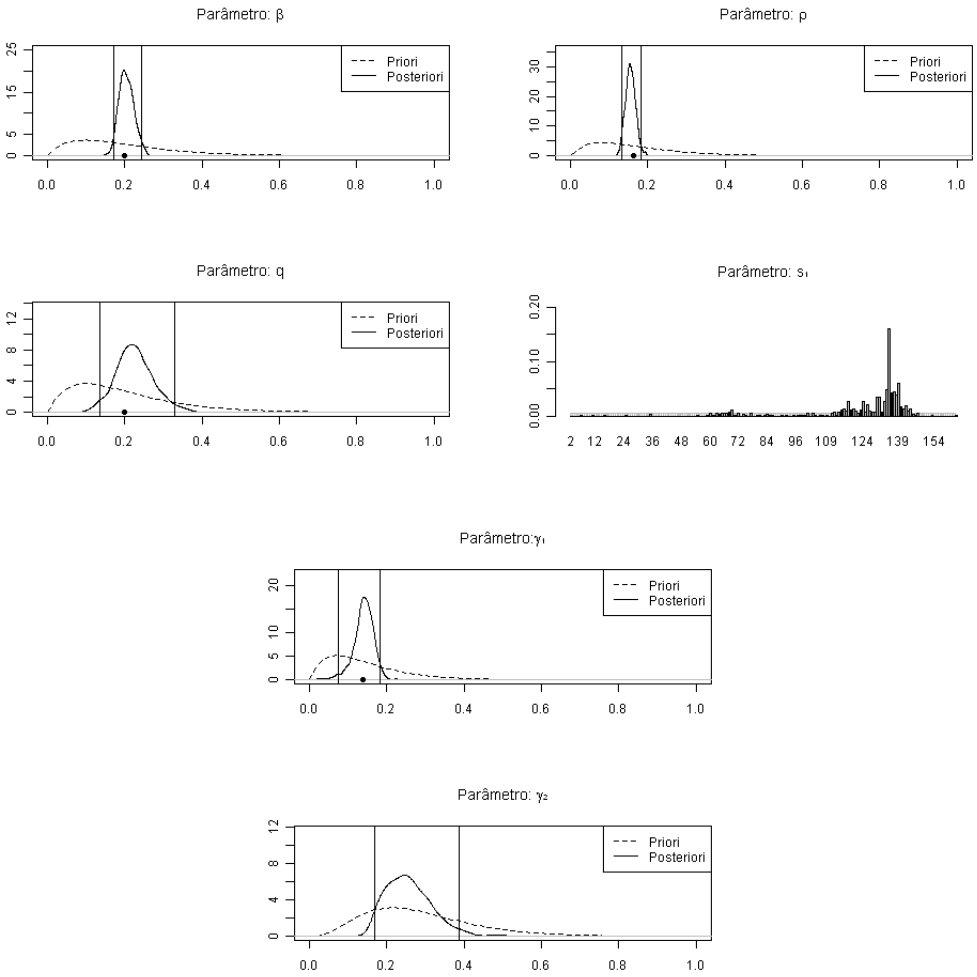


Figura 29: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

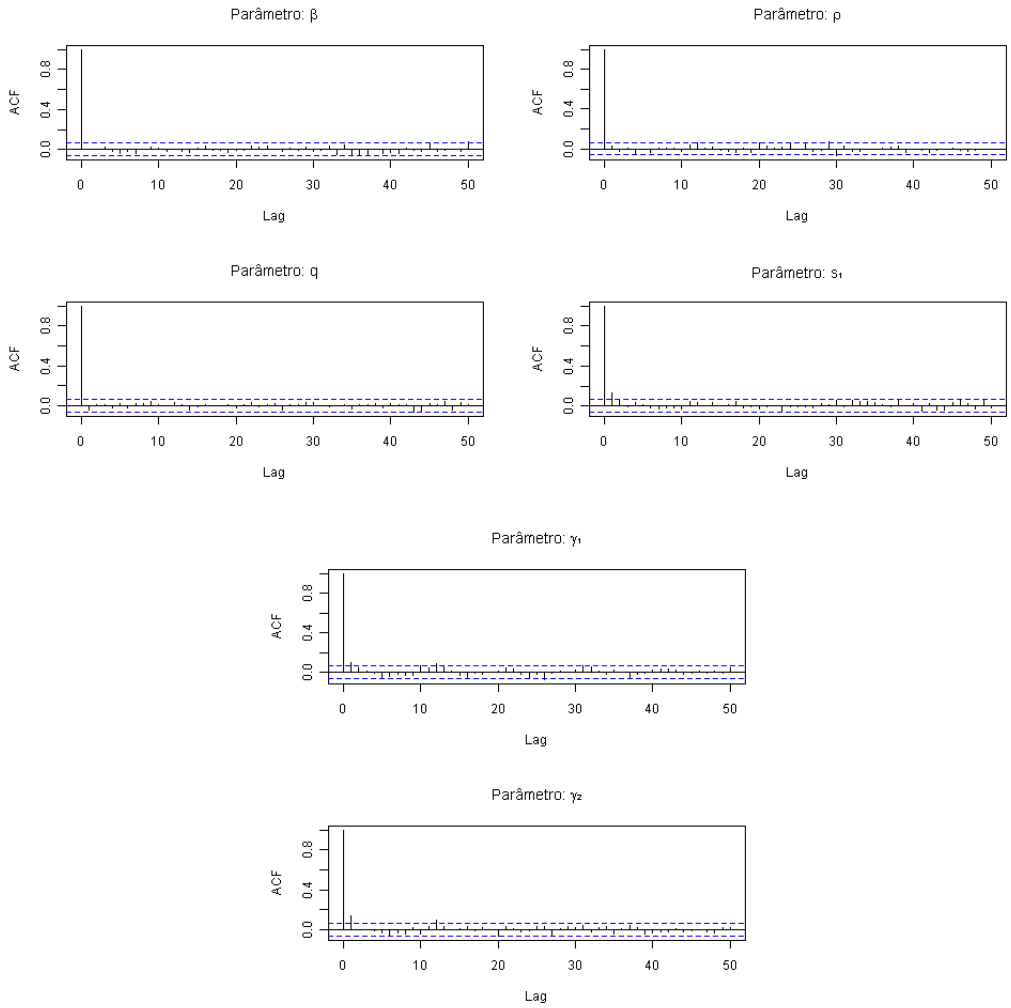


Figura 30: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

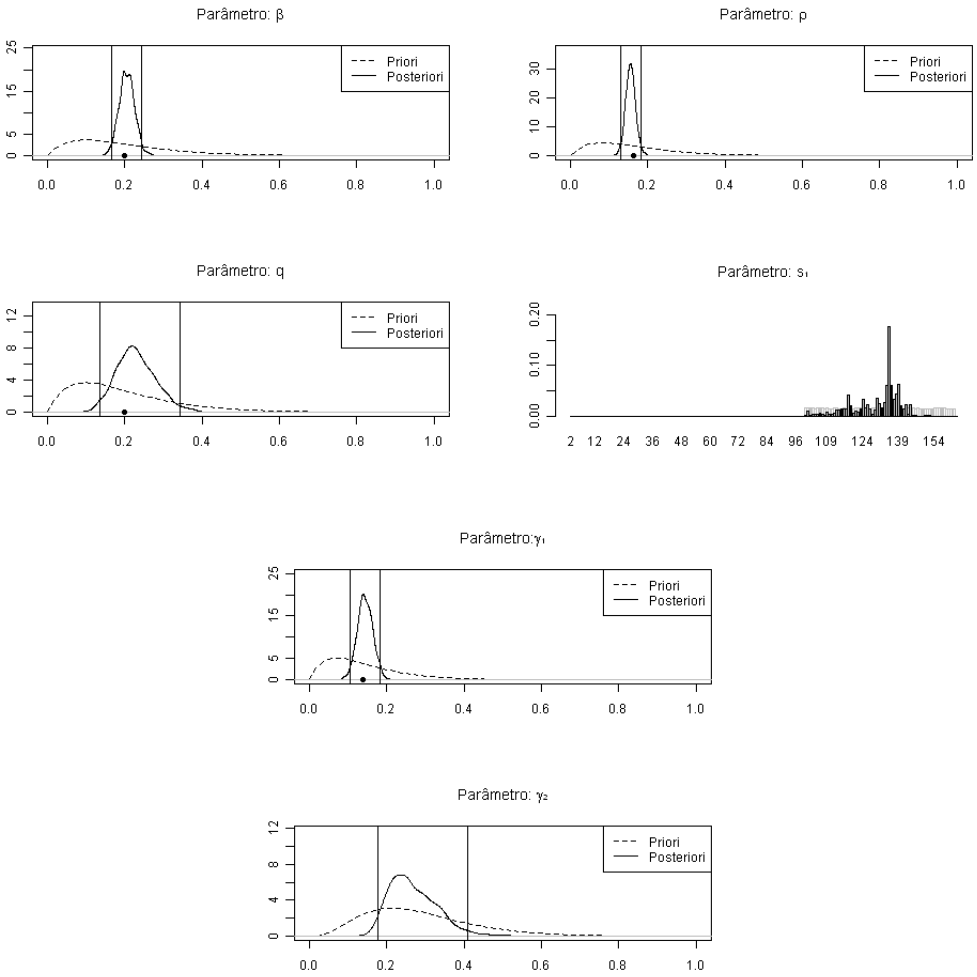


Figura 31: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

Tabela 9: Estatísticas descritivas das distribuições a posteriori considerando dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

β				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,204	0,019	(0,166 ; 0,244)	0,2
Informativa	0,205	0,020	(0,171 ; 0,244)	0,2
ρ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,155	0,013	(0,130 ; 0,182)	0,164
Informativa	0,155	0,013	(0,132 ; 0,181)	0,164
q				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,225	0,048	(0,137 ; 0,342)	0,2
Informativa	0,230	0,051	(0,135 ; 0,330)	0,2
γ_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,140	0,027	(0,105 ; 0,182)	0,139
Informativa	0,142	0,020	(0,073 ; 0,184)	0,139
γ_2				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,257	0,059	(0,176 ; 0,409)	0,278
Informativa	0,262	0,063	(0,168 ; 0,386)	0,278
s_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	120,606	28,452	(104,975 ; 144,000)	130
Informativa	130,054	10,384	(34,975 ; 144,000)	130

4.4.1.2 Conclusões: Análise de sensibilidade

A partir da Tabela 9, pode-se ver que o algoritmo apresenta estimativas razoáveis para qualquer escolha de distribuição a priori para s_1. Além disso, é fácil ver que as estimativas de β, ρ e q não sofreram modificações, como esperado, quando a distribuição a priori de s_1 foi modificada.

As estimativas de γ_1, γ_2 foram alteradas: o desvio padrão de γ_1 foi reduzido ao considerar uma distribuição informativa para s_1 e, consequentemente, o intervalo de credibilidade 95% tornou-se menor; por outro lado, o desvio padrão de γ_2 tornou-se ligeiramente um pouco maior ao considerar uma distribuição informativa para s_1, porém, o impacto não parece ser significativo.

A estimativa de s_1 teve grandes alterações, o valor pontual tornou-se muito mais próximo do valor verdadeiro ao considerar a distribuição informativa para s_1 assim como o desvio padrão tornou-se significativamente menor. Todavia, é válido notar que os intervalos de credibilidade 95% ainda são extensos em relação à distribuição a priori, isto é, o método de estimação tem grande dificuldade em encontrar boas estimativas para o ponto de mudança.

4.4.1.3 Erro de especificação

Uma importante etapa no processo de modelagem é a escolha do modelo mais apropriado ao conjunto de dados. Nesta dissertação, pode-se entender por modelo não apropriado como um modelo que considera uma taxa de remoção homogênea no tempo quando esta é, na verdade, não homogênea ou um modelo que considera uma taxa de remoção não homogênea no tempo quando esta é, de fato, homogênea.

Além disso, quando um modelo de taxa de remoção não homogênea no tempo é considerado apropriado, resta a discussão do número de pontos de mudança. Esta discussão não será feita já que somente o caso $k = 1$ é considerado. Mais detalhes para um número aleatório de pontos de mudança podem ser encontrados em Boys e Giles (2007), Green (1995) que utilizam o algoritmo *Reversible Jump*.

Então, considere o modelo do capítulo anterior dado por (3.1) - (3.4) e a epidemia

simulada nesta capítulo, ou seja, a situação em que o modelo de taxa de remoção homogênea no tempo é escolhido quando esta é, na verdade, não homogênea.

Para a distribuição a priori não informativa, as Figuras 86, 87 e 88 foram consideradas. Os detalhes da simulação estão na Tabela 10. Os resultados seguem nas Figuras 32, 33 e na Tabela 11.

Tabela 10: Detalhes das simulações considerando dados completos na análise de erro de especificação do modelo com taxa de remoção homogênea

Priori	TAA	EA	TC (segundos)
Não informativa	1000	5	44,09
Parâmetros	**Estatística R**		**Taxas de aceitação (%)**
	Priori não informativa		
β	1,000		44,30
ρ	1,000		44,30
γ	1,000		44,12
q	1,000		45,70

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, TC = Tempo

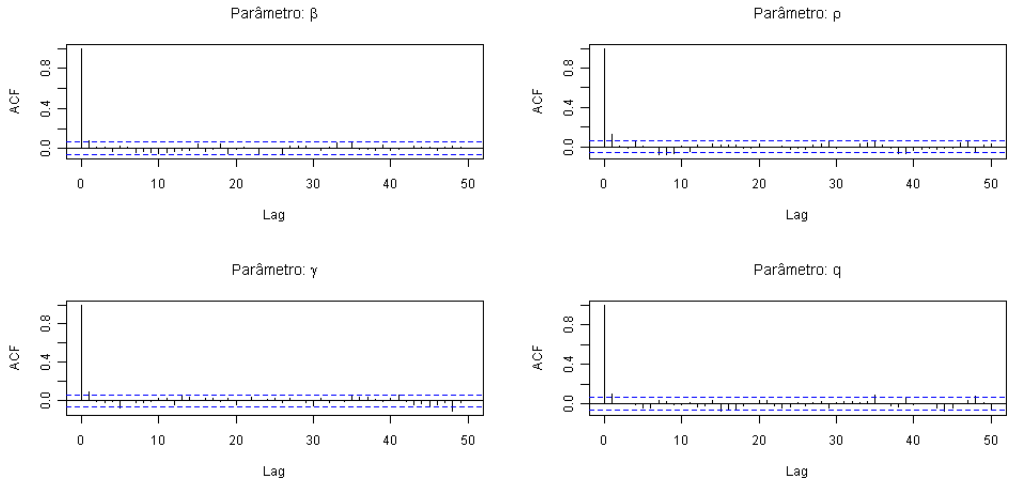


Figura 32: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção homogênea

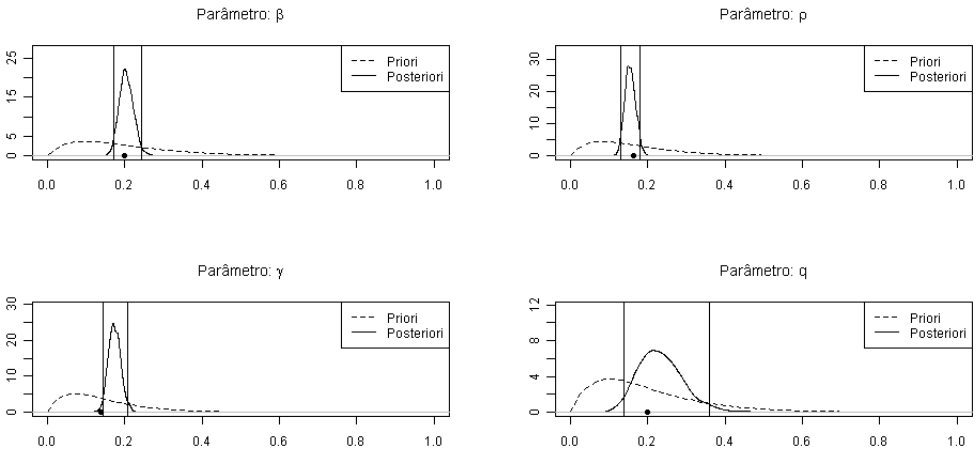


Figura 33: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção homogênea

Tabela 11: Estatísticas descritivas das distribuições a posteriori considerando dados completos na análise de erro de especificação do modelo com taxa de remoção homogênea

Parâmetro	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
β	0,204	0,018	(0,171 ; 0,243)	0,2
ρ	0,154	0,013	(0,129 ; 0,181)	0,164
γ	0,174	0,016	(0,144 ; 0,207)	-
q	0,234	0,057	(0,138 ; 0,360)	0,2

E, a fim de verificar o comportamento do modelo de taxa de remoção não homogênea quando não há pontos de mudança na taxa de remoção, considere a epidemia simulada no capítulo anterior. A epidemia simulada no capítulo anterior pelo modelo de Lekone e Finkenstädt (2006) a partir dos parâmetros $\beta = 0,2$; $\rho = \frac{1}{6,1}$, $\gamma = \frac{1}{7,2}$ e $q = 0,2$ tem tamanho igual a 72 e duração de 176 dias.

Para a distribuição a priori não informativa, os gráficos para a definição das condições da simulação são dados na Figuras 89, 90 e 91. Para a distribuição informativa, as Figuras 92, 93 e 94 são consideradas. Os detalhes das simulações estão na Tabela 12.

Os resultados das simulações seguem nas Figuras 34, 35, 36, 37, respectivamente, para as distribuições a priori não informativa e informativa. As estatísticas descritivas

estão na Tabela 13.

Tabela 12: Detalhes das simulações considerando dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

Priori	TAA	EA	TC (segundos)
Não informativa	8000	20	379,37
Informativa	2000	25	263,10

Parâmetros	Estatística R	Taxas de aceitação (%)
Priori não informativa		
β	1,008	44,41
ρ	1,000	46,69
q	1,009	46,82
γ_1	1,022	40,07
γ_2	1,034	41,42
s_1	1,072	34,53
Priori informativa		
β	1,000	44,37
ρ	1,000	46,54
q	1,000	47,21
γ_1	1,000	44,42
γ_2	1,000	44,86
s_1	1,000	36,06

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, TC = Tempo

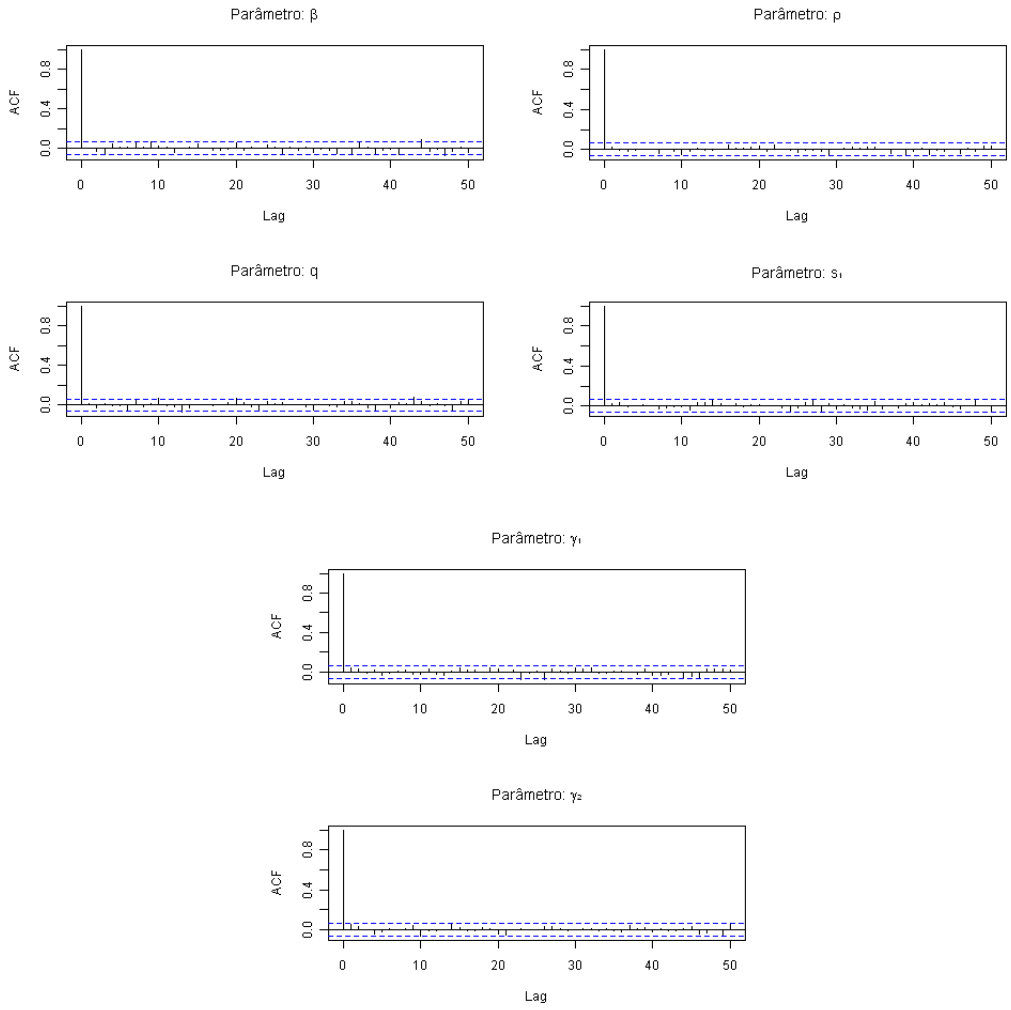


Figura 34: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

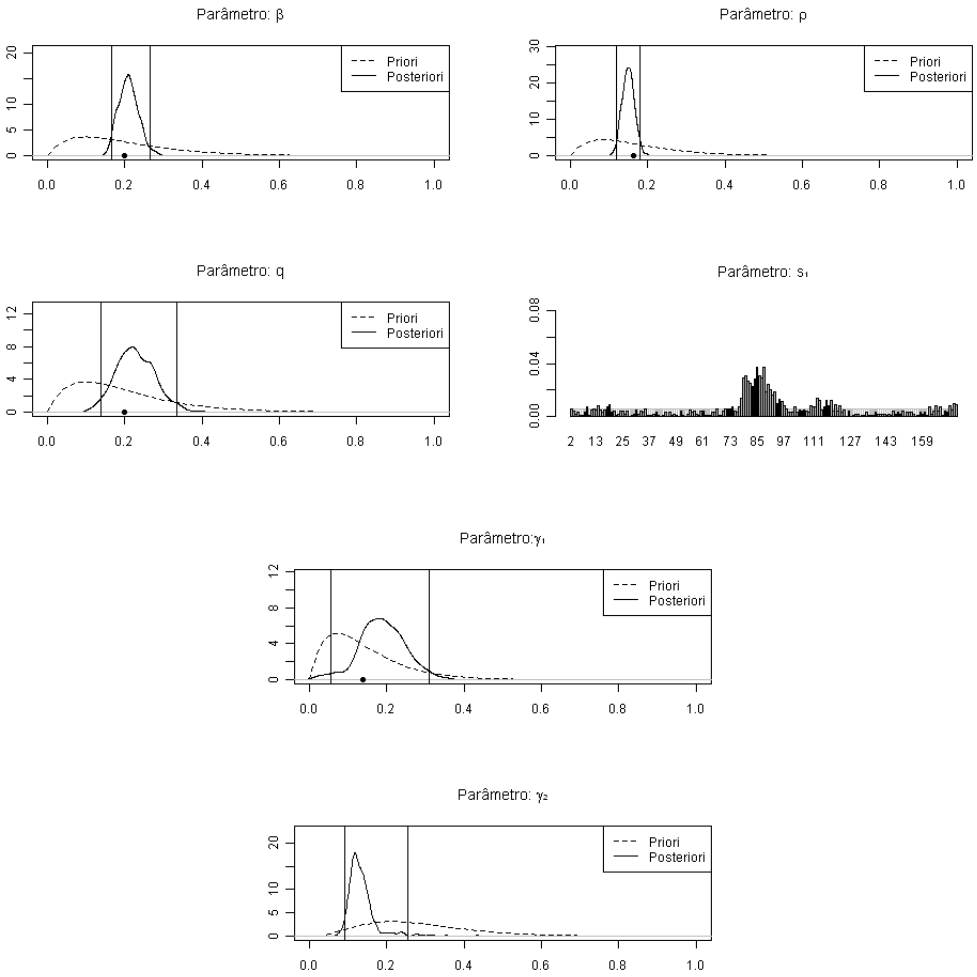


Figura 35: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

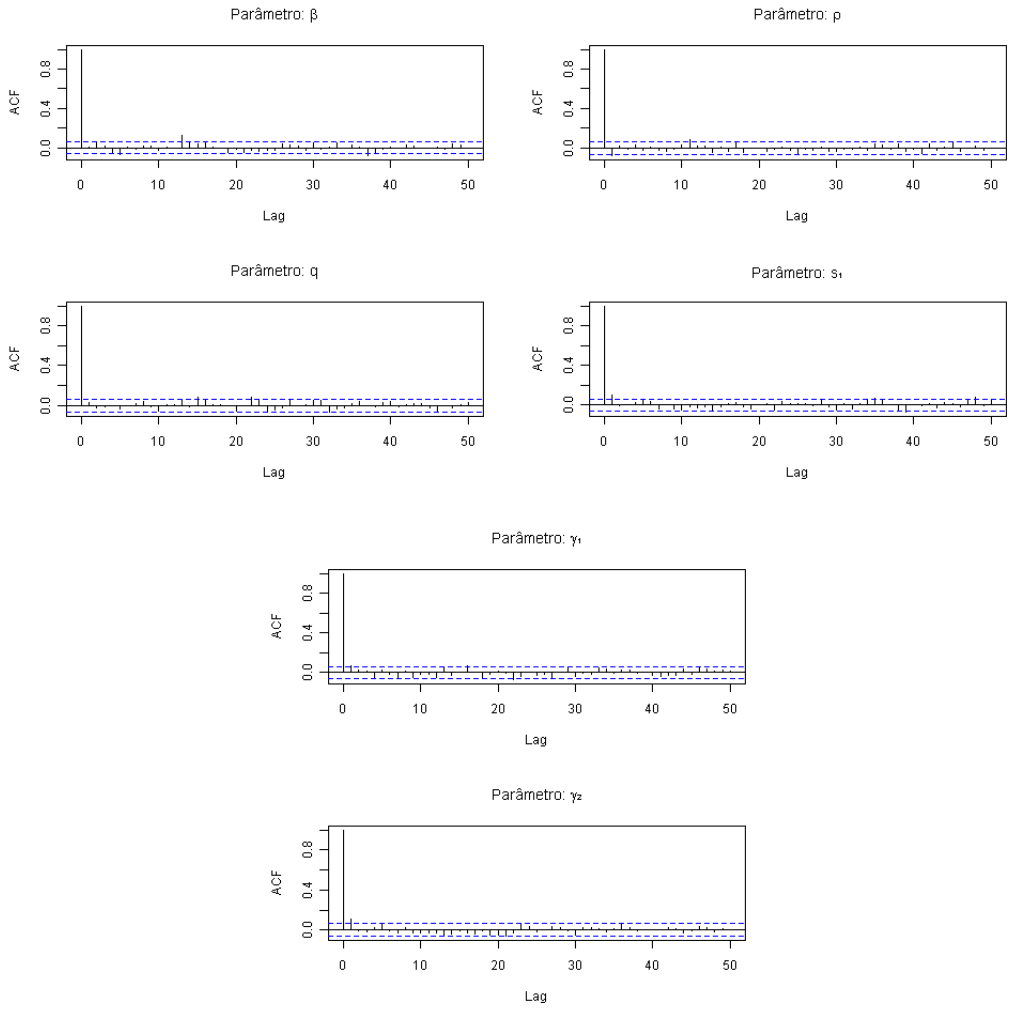


Figura 36: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

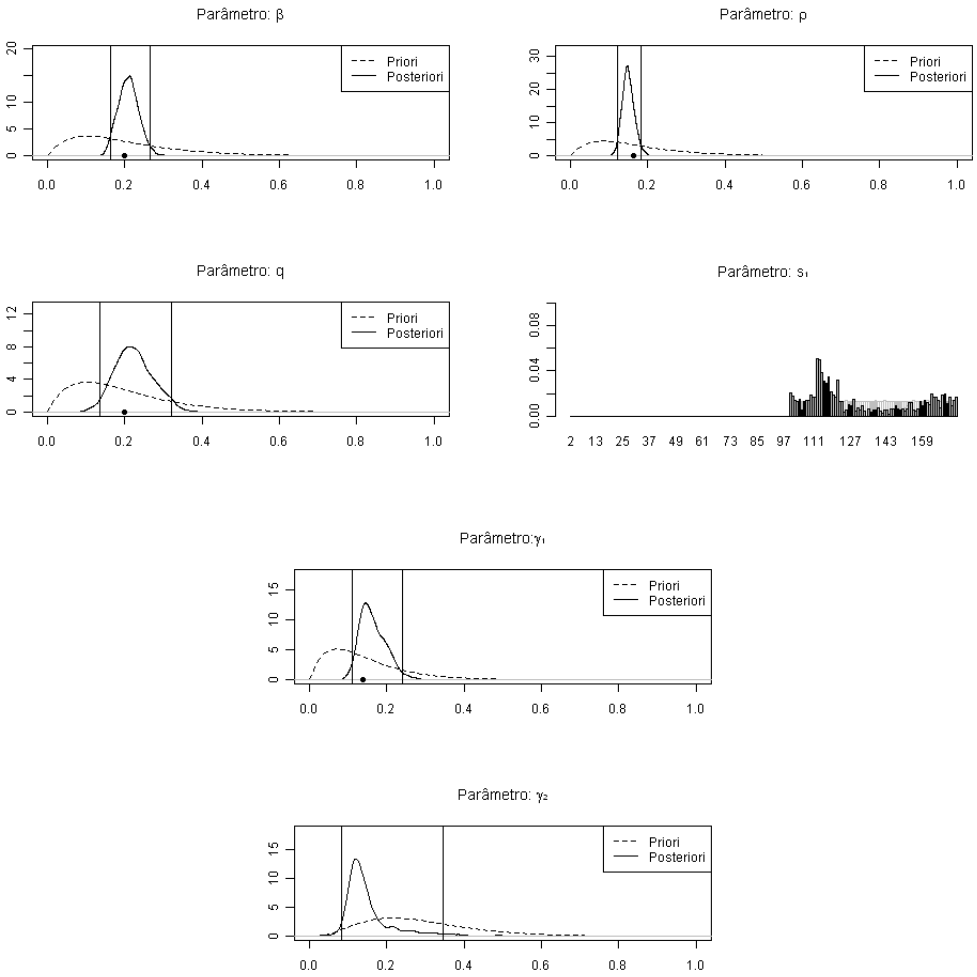


Figura 37: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

Tabela 13: Estatísticas descritivas das distribuições a posteriori considerando dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

β				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,210	0,026	(0,165 ; 0,265)	0,2
Informativa	0,209	0,026	(0,162 ; 0,265)	0,2
ρ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,150	0,016	(0,120 ; 0,180)	0,164
Informativa	0,149	0,015	(0,122 ; 0,182)	0,164
q				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,224	0,050	(0,138 ; 0,334)	0,2
Informativa	0,223	0,049	(0,136 ; 0,322)	0,2
γ_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,189	0,065	(0,056 ; 0,308)	0,139
Informativa	0,165	0,034	(0,111 ; 0,240)	0,139
γ_2				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,135	0,044	(0,091 ; 0,254)	-
Informativa	0,152	0,067	(0,084 ; 0,347)	-
s_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	86,649	40,233	(9 ; 171)	-
Informativa	131,861	23,477	(101 ; 173)	-

4.4.1.4 Conclusões: Erro de especificação

Inicialmente, considere a situação em que a taxa de remoção é não homogênea, porém, um modelo SEIR com taxa de remoção homogênea é ajustado: a partir da Tabela 11, pode-se ver que as estimativas pontuais de β, ρ e q são razoáveis e os valores verdadeiros dos parâmetros da simulação estão contidos nos intervalos de credibilidade 95%. Contudo, o mesmo não pode ser dito para a estimativa pontual de γ, a estimativa dada pelo modelo é uma média das duas taxas de remoção $\gamma_1 = 0.139$ e $\gamma_2 = 0.278$, consequentemente, o intervalo de credibilidade 95% não contém nenhum dos valores verdadeiros.

Na segunda situação, a taxa de remoção é homogênea, todavia, um modelo SEIR com taxa de remoção não homogênea é ajustado: a partir da Tabela 13, pode-se ver que as estimativas de β, ρ e q também são razoáveis tal que os intervalos de credibilidade contêm os valores verdadeiros. Entretanto, as estimativas pontuais para γ_1 e γ_2 não fazem sentido segundo a suposição de que a introdução de medidas de intervenção tem um impacto positivo ao reduzir o período de remoção, isto é, a estimativa de γ_2 deveria ser maior do que a estimativa de γ_1.

É possível dizer que a suposição sobre o impacto das medidas de intervenção esteja errada e, na verdade, as medidas de intervenção tenham sido ineficazes e o impacto negativo, entretanto, esta possibilidade é remota dada que a epidemia foi extinta em um período não muito posterior à introdução das medidas de intervenção. Além disso, os intervalos de credibilidade 95% devem ser considerados e estes não são mutuamente excludentes e ordenados, portanto, γ_1 pode ser menor do que γ_2.

Além disso, a estimativa pontual de s_1 também não é razoável e dada pela média do intervalo da distribuição a priori sendo que os intervalos de credibilidade 95% se extendem sobre quase todo o suporte da distribuição a priori.

Esta discussão possibilita concluir que, no contexto de dados completos, é sempre válido ajustar um modelo com taxa de remoção não homogênea no tempo e se este não for adequado, as estimativas dos parâmetros envolvidos com a taxa de remoção não homogênea não serão plaúsiveis e, portanto, um modelo com taxa de remoção homogênea deve ser ajustado. Observe que se somente o modelo de taxa de remoção

homogênea for ajustado, não há indicativos de que o modelo não é adequado.

Note que a discussão poderia ser complementada com a informação dada por critérios de seleção de modelos, como por exemplo, o DIC. Todavia, o DIC parece não ser um critério adequado, já que o modelo com taxa de remoção não homogênea apresenta um menor valor de DIC em qualquer uma das situações como se vê na Tabela 14.

Tabela 14: DIC

Situação: Taxa de remoção homogênea	
Modelo	DIC
Taxa de remoção homogênea	2173,694
Taxa de remoção não homogênea	2167,499
Situação: Taxa de remoção não homogênea	
Modelo	DIC
Taxa de remoção homogênea	2326,388
Taxa de remoção não homogênea	2312,604

E para confirmar a conclusão sobre o DIC, 200 epidemias foram simuladas tal que o período de extinção da epidemia ocorresse após o período de introdução das medidas de intervenção. Os modelos com taxa de remoção homogênea e não homogênea foram ajustados considerando o mesmo tamanho de amostra de aquecimento e espaçamento para todas simulações e a partir deste procedimento, observou-se que o modelo com taxa de remoção não homogênea sempre apresenta um valor do DIC inferior ao modelo com taxa de remoção homogênea.

4.4.2 B, C, D Incompletos

O segundo cenário considerado é na presença de dados faltantes em todas as séries de dados seguindo a proporção de 99,6%, 8% e 25% para, respectivamente, **B**, **C** e **D**. Note que o cenário em que somente a série de dados **B** apresenta dados faltantes não é considerada, pois, a utilização de **B** aumentado irá influenciar somente na estimação de β, q e ρ, por sua vez, **C** aumentado influencia todas as estimativas dos parâmetros e por fim, **D** influencia as estimativas de β, q e γ.

A implementação do algoritmo foi realizada da mesma maneira do capítulo anterior

com as considerações adicionais sobre γ e s, portanto, densidades propostas normais univariadas de passeio aleatório para β, γ_1, γ_2 e de independência normais com $c_\rho = c_q = 1,5$ para ρ e q, enquanto uma densidade proposta de independência dada pela distribuição a priori de s_1 para s_1.

Os resultados observados foram bons, apesar da estrutura de correlação amostral envolvendo γ e s apresentar valores significativos de magnitude 0,5.

4.4.2.1 Análise de sensibilidade

Para a distribuição a priori não informativa, os gráficos para a definição das condições da simulação são dados na Figuras 95, 96 e 97. Para a distribuição informativa, as Figuras 98, 99 e 100 são consideradas. Os detalhes da simulação estão na Tabela 15.

Os resultados das simulações seguem nas Figuras 38, 39, 40, 41, respectivamente, para as distribuições a priori não informativa e informativa. As estatísticas descritivas estão na Tabela 16.

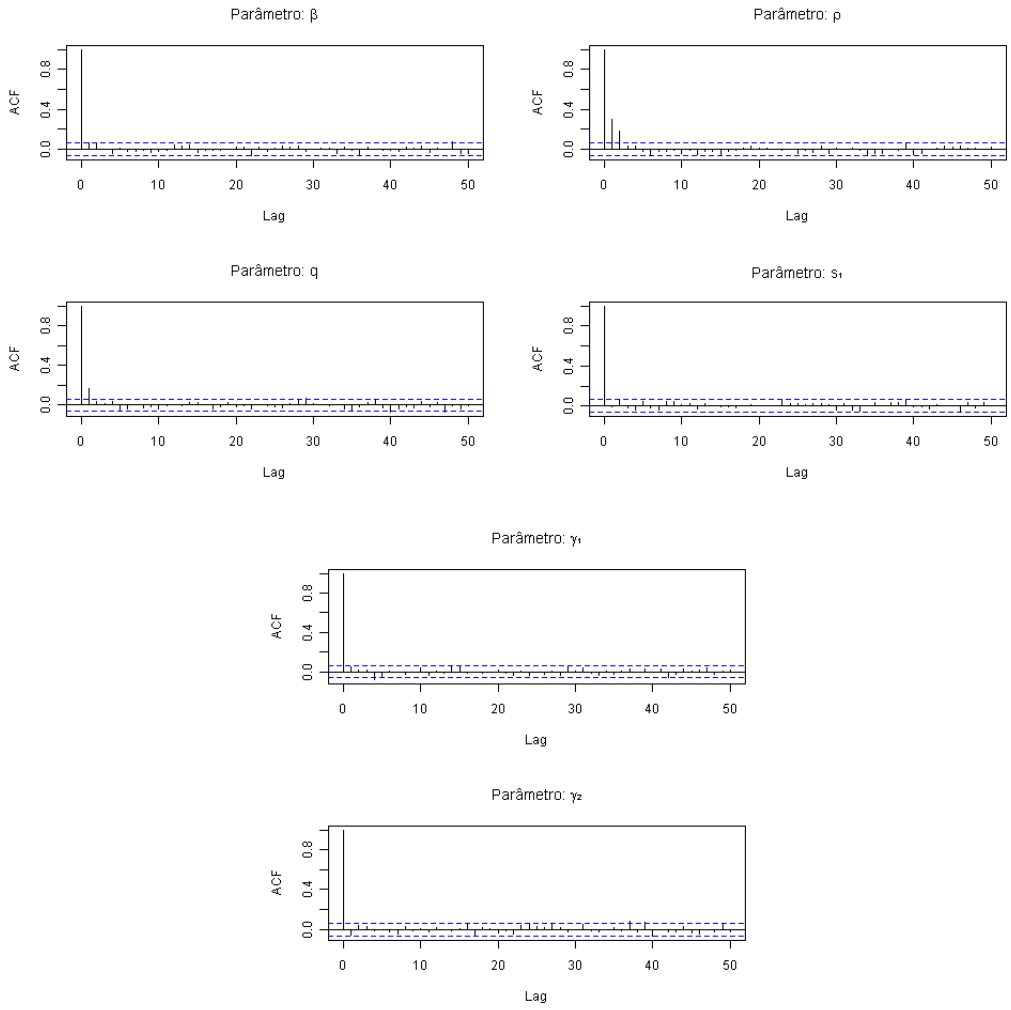


Figura 38: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

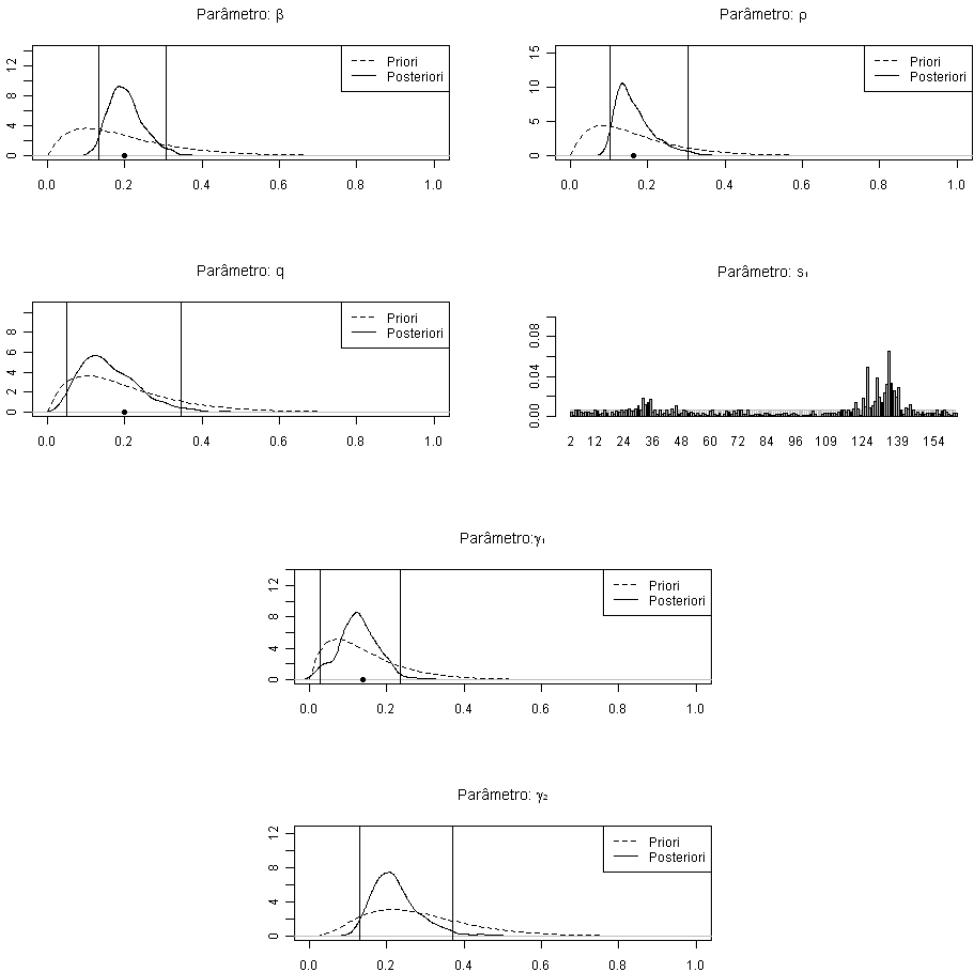


Figura 39: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

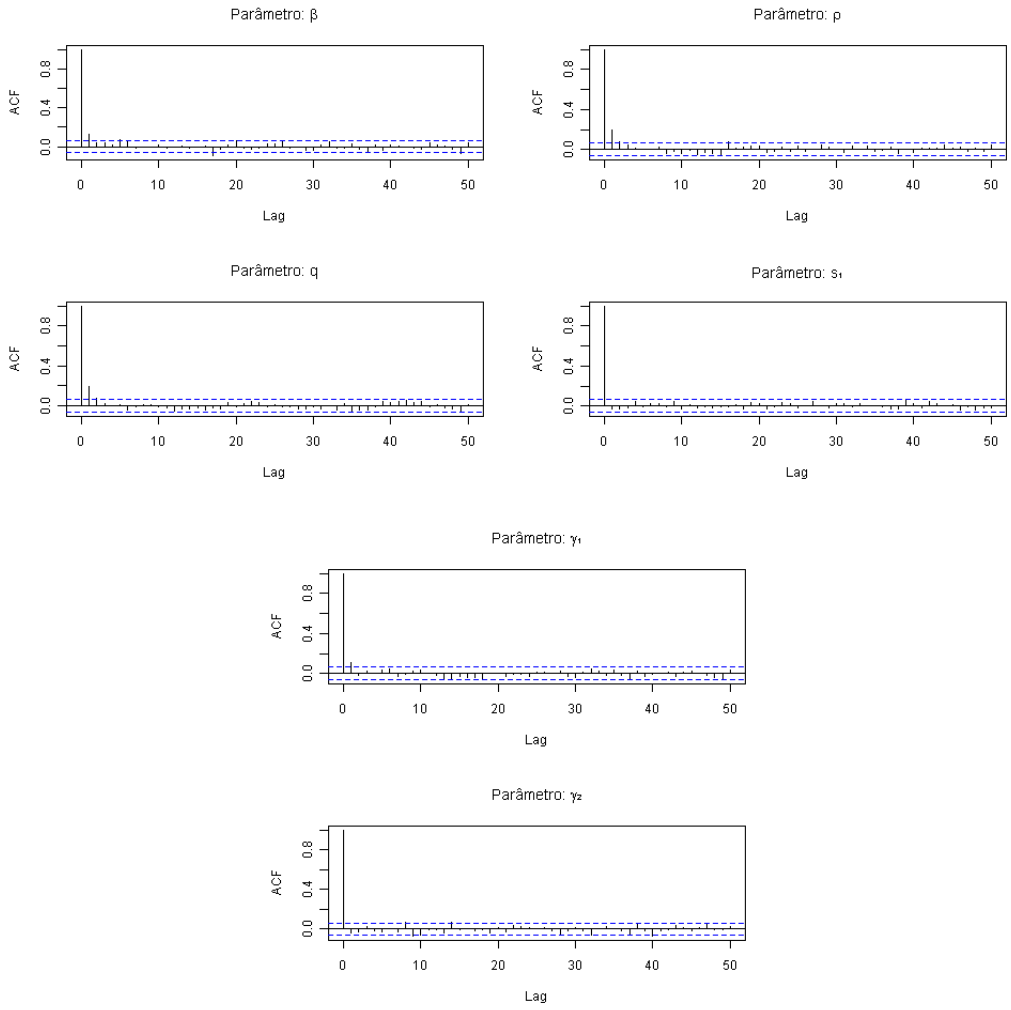


Figura 40: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

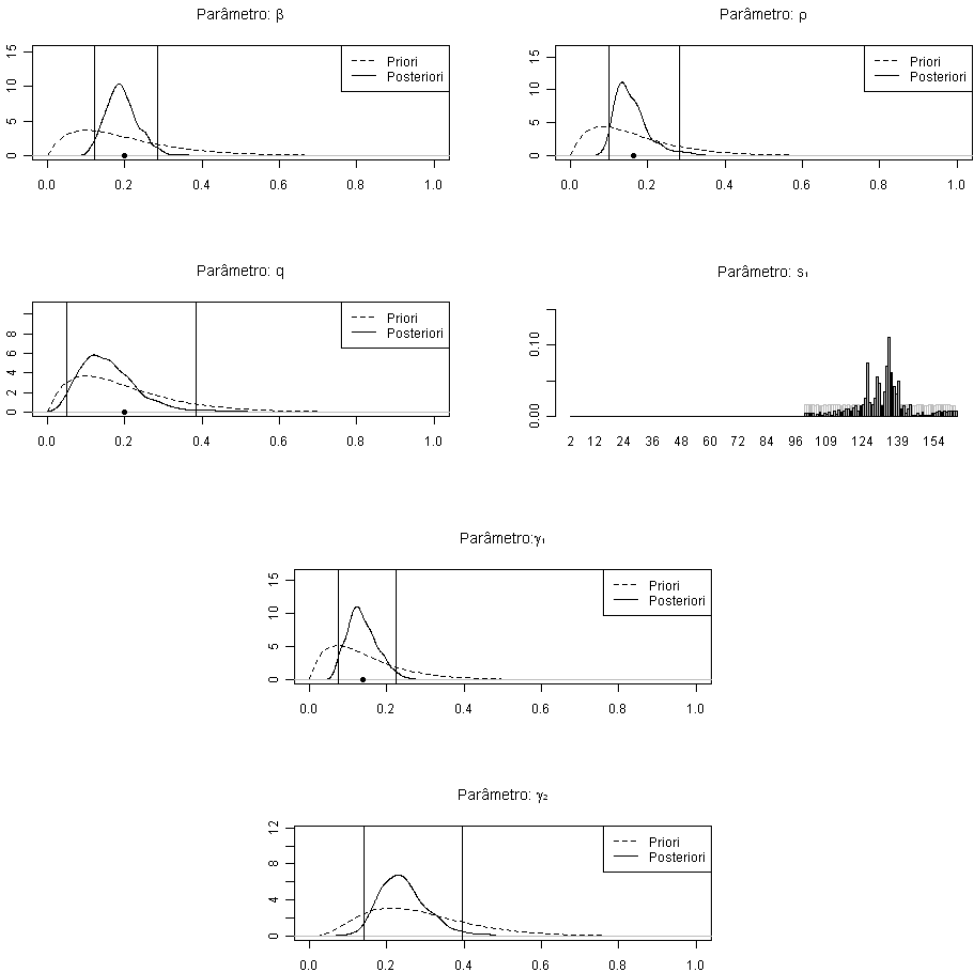


Figura 41: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

Tabela 15: Detalhes das simulações considerando **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

Priori	TAA	EA	T (horas)
Não informativa	100000	1000	9,83
Informativa	100000	750	7,40
Parâmetros	**Estatística R**		**Taxas de aceitação (%)**
Priori não informativa			
B	-		21,37
C	-		25,79
D	-		26,49
β	1,002		44,52
ρ	1,018		74,58
q	1,021		73,29
γ_1	1,000		40,06
γ_2	1,000		43,07
s_1	1,001		26,37
Priori informativa			
B	-		22,09
C	-		27,33
D	-		27,69
β	1,003		44,46
ρ	1,002		74,57
q	1,000		73,24
γ_1	1,003		44,21
γ_2	1,000		42,31
s_1	1,000		25,89

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, T = Tempo

4.4.2.2 Conclusões: Análise de sensibilidade

A partir da Tabela 16, pode-se ver que o algoritmo é funcional para qualquer escolha de distribuição a priori para s_1. Além disso, as estimativas de β, ρ e q sofreram pequenas mudanças quando a distribuição a priori de s_1 foi modificada, diferentemente do cenário anterior em que não houve mudanças. Tais mudanças surgem devido aos dados aumentados.

As estimativas de γ_1, γ_2 foram alteradas de maneira similar discutida no cenário de dados completos: o desvio padrão de γ_1 foi reduzido ao considerar uma distribuição informativa para s_1 e, consequentemente, o intervalo de credibilidade 95% tornou-se

Tabela 16: Estatísticas descritivas das distribuições a posteriori considerando **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

β				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,203	0,045	(0,134 ; 0,306)	0,2
Informativa	0,193	0,042	(0,121 ; 0,284)	0,2
ρ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,169	0,054	(0,103 ; 0,304)	0,164
Informativa	0,159	0,046	(0,098 ; 0,282)	0,164
q				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,162	0,078	(0,049 ; 0,346)	0,2
Informativa	0,163	0,083	(0,050 ; 0,385)	0,2
γ_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,129	0,056	(0,028 ; 0,235)	0,139
Informativa	0,137	0,039	(0,074 ; 0,223)	0,139
γ_2				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,223	0,063	(0,129 ; 0,370)	0,278
Informativa	0,244	0,064	(0,142 ; 0,394)	0,278
s_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	96,986	47,493	(8 ; 156)	130
Informativa	132,668	11,368	(106 ; 160)	130

menor; por outro lado, o desvio padrão de γ_2 tornou-se ligeiramente um pouco maior ao considerar uma distribuição informativa para s_1, porém, a modificação não parece ser significativa com exceção da estimativa pontual.

A estimativa de s_1 teve grandes alterações, o valor pontual tornou-se muito mais próximo do valor verdadeiro ao considerar a distribuição informativa para s_1 assim como o desvio padrão tornou-se significativamente menor. Todavia, é válido notar que o intervalo de credibilidade 95% ainda é muito extenso em relação ao suporte da distribuição a priori.

Por fim, se comparadas as Tabelas 9 e 16, é possível observar que houve um aumento significativo dos desvios padrão das distribuições posteriori quando os dados incompletos são considerados. Não houve modificações somente para o parâmetro γ_2 nas duas escolhas de distribuições a priori e para o parâmetro s_1 ao escolher a distribuição a priori informativa para o mesmo.

Além disso, as estimativas pontuais para q, γ_1, γ_2 e s_1 tornaram-se mais distantes do valor em verdadeiro no cenário de dados incompletos se comparado ao cenário de dados completos apresentado na Tabela 9.

4.4.2.3 Erro de especificação

Novamente, a discussão sobre erro de especificação de modelo deve ser feita. É necessário verificar se as idéias desenvolvidas para o cenário de dados completos ainda podem ser aplicadas no cenário de dados incompletos. O DIC não é considerado, pois já foi visto que não é um critério adequado para modelos SEIR com dados completos e, consequemente, para dados incompletos.

Cabe ressaltar que é possível calcular o DIC mesmo considerando dados incompletos como apresentado por Celeux et al. (2006), todavia, a exigência computacional pode ser bastante grande e, neste caso, não foi possível calculá-lo por esta razão.

Então, considere a situação em que um modelo de taxa de remoção homogênea no tempo quando esta é, na verdade, não homogênea.

Para a distribuição a priori não informativa, as Figuras 101, 102 e 103 foram con-

sideradas. Os detalhes da simulação estão na Tabela 17. Os resultados seguem nas Figuras 42, 43 e na Tabela 18.

Tabela 17: Detalhes das simulações considerando **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção homogênea

Priori	TAA	EA	T (horas)
Não informativa	100000	1000	10,63
Parâmetros	**Estatística R**		**Taxas de aceitação (%)**
Priori não informativa			
B	-		21,72
C	-		21,73
D	-		22,32
β	1,009		44,49
ρ	1,002		74,56
γ	1,011		44,36
q	1,000		74,98

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, T = Tempo

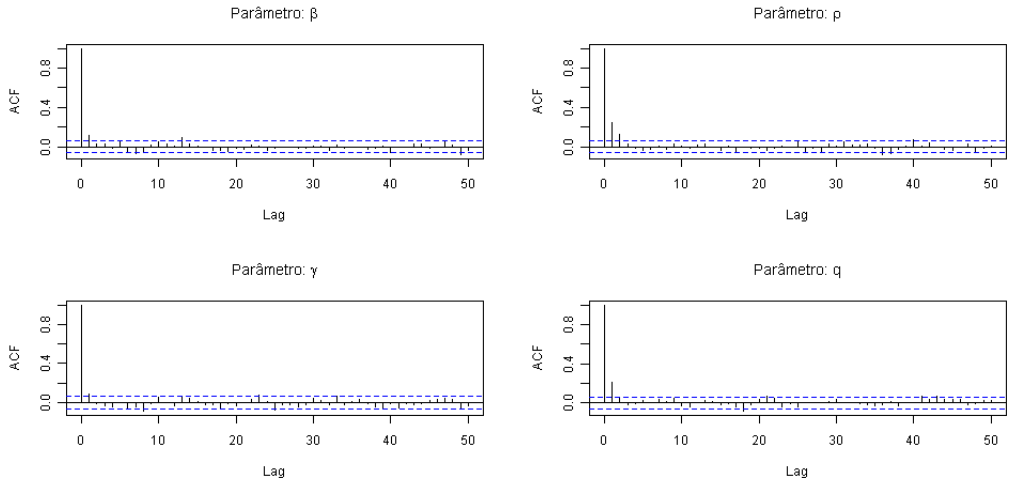


Figura 42: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção homogênea

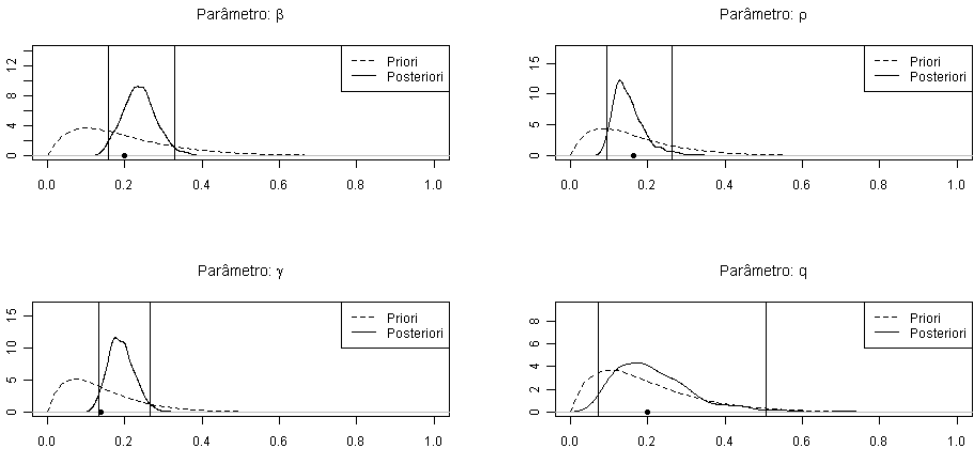


Figura 43: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção homogênea

Tabela 18: Estatísticas descritivas das distribuições a posteriori considerando **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção homogênea

Parâmetro	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
β	0,238	0,043	(0,158 ; 0,329)	0,2
ρ	0,152	0,043	(0,094 ; 0,264)	0,164
γ	0,191	0,034	(0,132 ; 0,264)	-
q	0,220	0,111	(0,073 ; 0,505)	0,2

E a segunda situação consiste em ajustar um modelo de taxa de remoção não homogênea quando não há pontos de mudança na taxa de remoção.

Novamente, considere a epidemia simulada no capítulo anterior pelo modelo de Lekone e Finkenstädt (2006) a partir dos parâmetros $\beta = 0,2$; $\rho = \frac{1}{6,1} \approx 0,1639$, $\gamma = \frac{1}{7,2} \approx 0,1389$ e $q = 0,2$ com tamanho igual a 72 e duração de 176 dias.

Para a distribuição a priori não informativa, os gráficos para a definição das condições da simulação são dados na Figuras 104, 105 e 106. Para a distribuição informativa, as Figuras 107, 108 e 109 são consideradas. Os detalhes das simulações estão na Tabela 19.

Os resultados das simulações seguem nas Figuras 44, 45, 46, 47, respectivamente,

para as distribuições a priori não informativa e informativa. As estatísticas descritivas estão na Tabela 20.

Tabela 19: Detalhes das simulações considerando **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

Priori	TAA	EA	T (horas)
Não informativa	100000	1100	14,48
Informativa	100000	1250	16,53

Parâmetros	Estatística R	Taxas de aceitação (%)
Priori não informativa		
B	-	15,68
C	-	24,23
D	-	23,48
β	1,000	44,57
ρ	1,000	74,27
q	1,000	72,35
γ_1	1,000	37,76
γ_2	1,000	40,92
s_1	1,000	33,94
Priori informativa		
B	-	15,42
C	-	23,68
D	-	23,29
β	1,006	44,49
ρ	1,000	74,30
q	1,010	72,36
γ_1	1,005	44,09
γ_2	1,001	38,33
s_1	1,002	34,92

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, T = Tempo

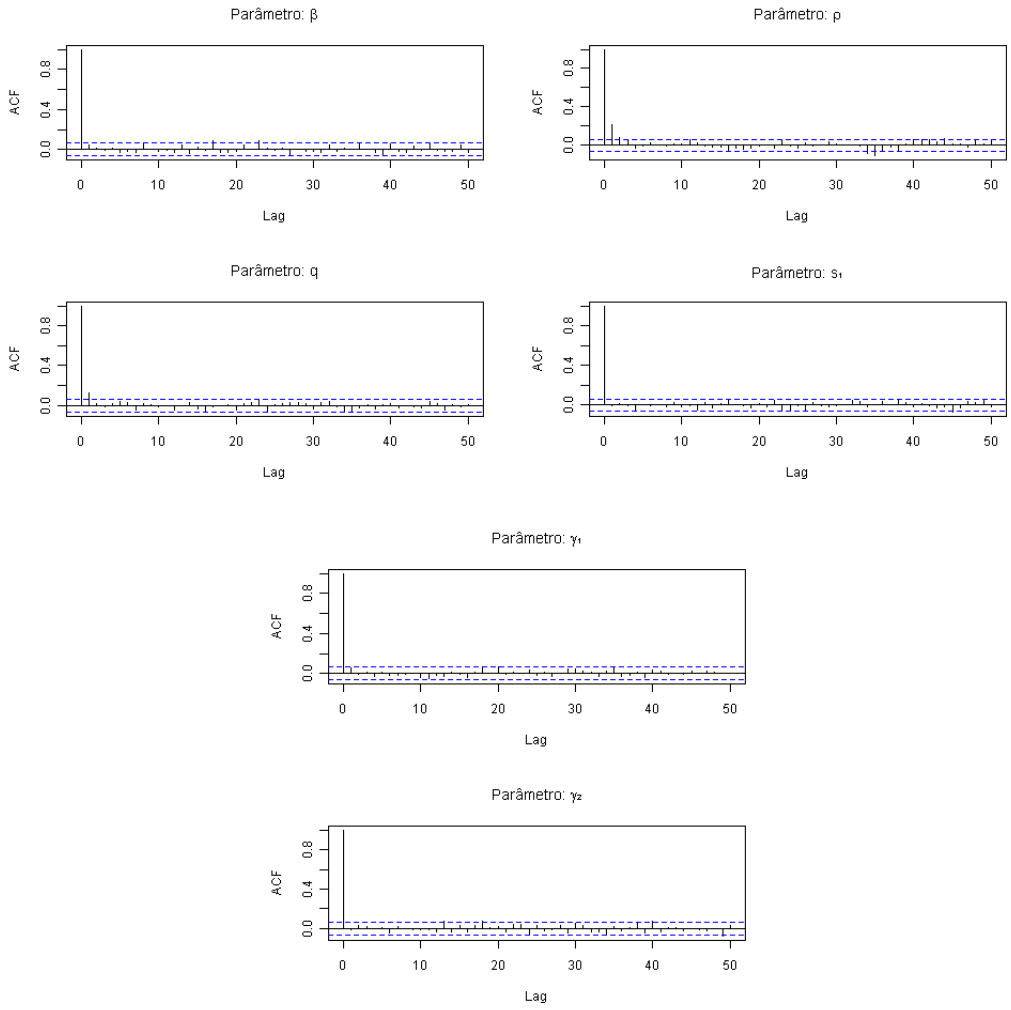


Figura 44: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

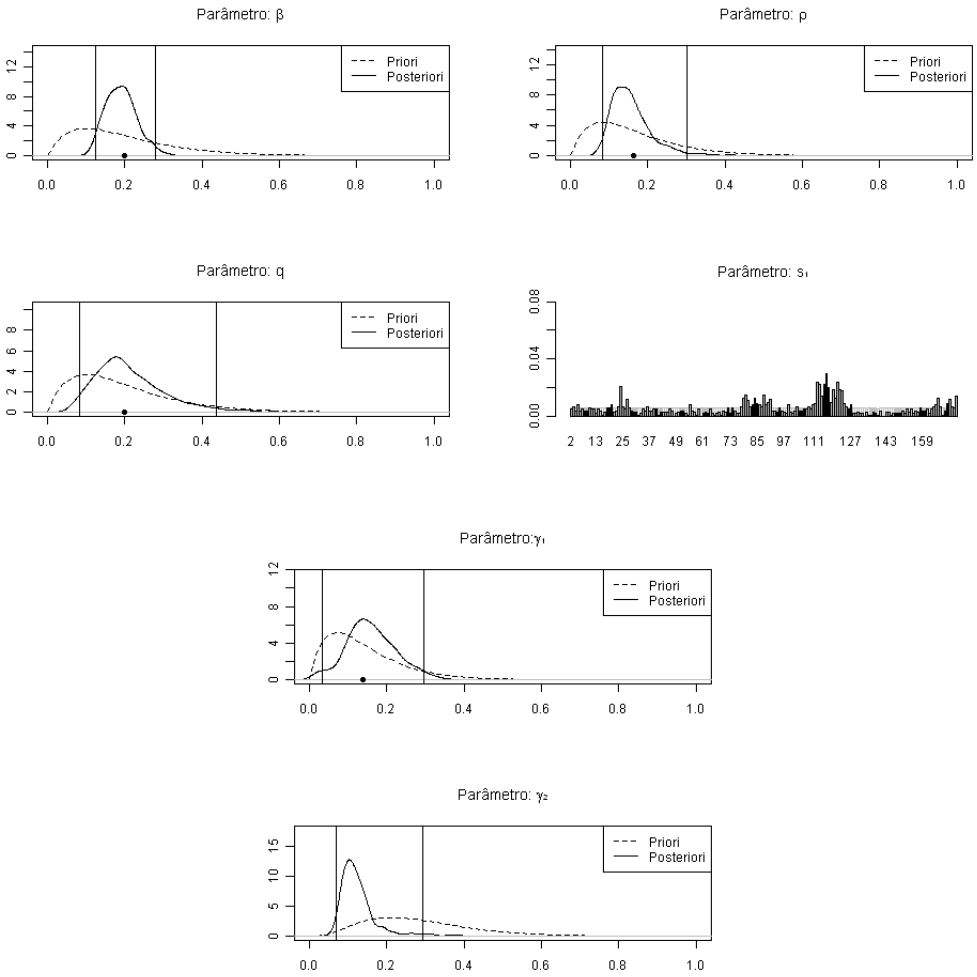


Figura 45: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

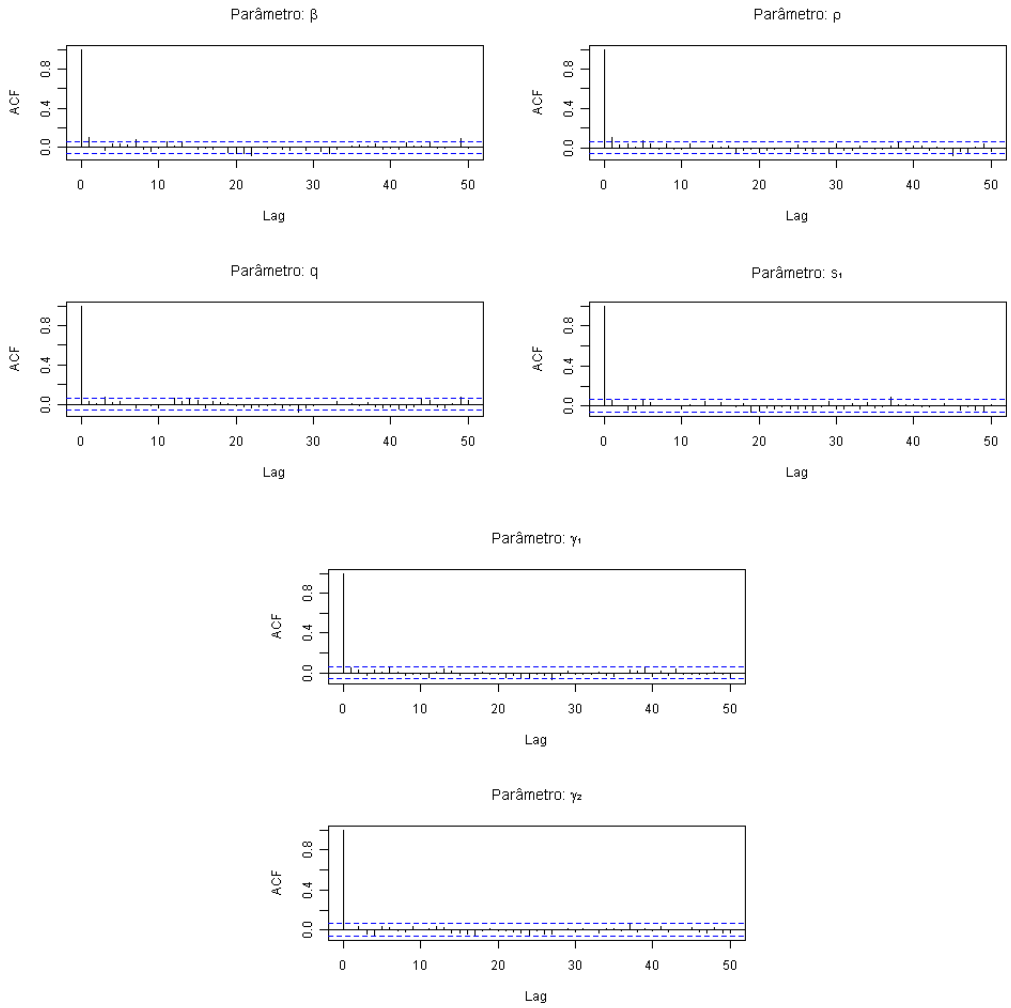


Figura 46: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

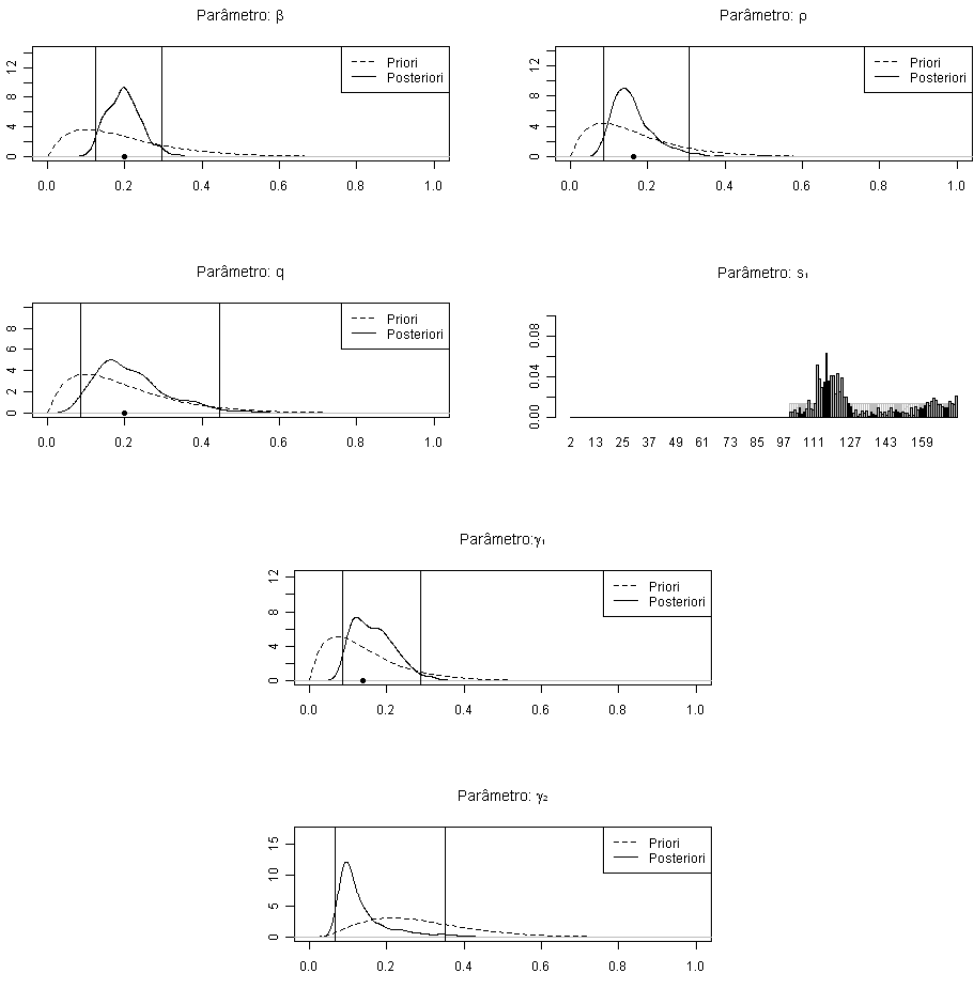


Figura 47: Densidades a priori e a posteriori e valores verdadeiros (pontos) para a distribuição a priori informativa com **B**, **C**, **D** incompletos

Tabela 20: Estatísticas descritivas das distribuições a posteriori considerando **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

β				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,191	0,041	(0,124 ; 0,278)	0,2
Informativa	0,198	0,044	(0,124 ; 0,295)	0,2
ρ				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,159	0,057	(0,083 ; 0,301)	0,164
Informativa	0,163	0,059	(0,085 ; 0,306)	0,164
q				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,214	0,092	(0,083 ; 0,438)	0,2
Informativa	0,219	0,096	(0,086 ; 0,446)	0,2
γ_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,160	0,065	(0,032 ; 0,296)	0,139
Informativa	0,166	0,054	(0,087 ; 0,287)	0,139
γ_2				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	0,127	0,055	(0,070 ; 0,294)	-
Informativa	0,137	0,074	(0,067 ; 0,350)	-
s_1				
Priori	Média	Desvio Padrão	IC 95%	Valor Verdadeiro
Não-Informativa	93,007	47,332	(6 ; 172)	-
Informativa	131,891	21,934	(104 ; 173)	-

4.4.2.4 Conclusões: Erro de especificação

As conclusões na situação em que a taxa de remoção é não homogênea, porém, um modelo SEIR com taxa de remoção homogênea é ajustado são similares ao cenário de dados completos. A partir da Tabela 18, as estimativas pontuais de β, ρ e q são razoáveis e a estimativa pontual de γ é uma média das duas taxas de remoção $\gamma_1 = 0.139$ e $\gamma_2 = 0.278$, consequentemente, o intervalo de credibilidade 95% não contém nenhum dos valores verdadeiros.

Na segunda situação, a taxa de remoção é homogênea, todavia, um modelo SEIR com taxa de remoção não homogênea é ajustado. As conclusões são também similares, todavia, não há tantas evidências de que o modelo com taxa de remoção não homogênea não é apropriado como no cenário de dados completos.

A partir da Tabela 20, as estimativas de β, ρ e q também são razoáveis tal que os intervalos de credibilidade 95% contêm os valores verdadeiros. Entretanto, as estimativas pontuais para γ_1 e γ_2, novamente, não são coerentes segundo a suposição de impacto positivo das medidas de intervenção.

A discussão desta última evidência para o contexto de dados completos é ainda válida, isto é, a suposição sobre o impacto positivo das medidas de intervenção pode estar errada e, na verdade, as medidas de intervenção tenham tido um impacto negativo, porém, esta possibilidade é remota dada que a epidemia foi extinta em um período não muito posterior à introdução das medidas de intervenção. Além disso, os intervalos de credibilidade 95% devem ser considerados e estes não são mutuamente excludentes ou ordenados, portanto, γ_1 pode ser menor do que γ_2.

A estimativa pontual de s_1 também não é razoável e dada pela média do intervalo da distribuição a priori sendo que os intervalos de credibilidade 95% se extendem sobre quase todo o suporte da distribuição a priori. Entretanto, estes resultados são bastante similares aos resultados dados na Tabela 16 quando um modelo de taxa de remoção não homogêna foi ajustado numa situação que, de fato, a taxa de remoção é não homogênea.

Esta discussão dá indícios de que a conclusão no contexto de dados completos é ainda válida, é interessante ajustar um modelo com taxa de remoção não homogênea

no tempo e se este não for adequado, as estimativas dos parâmetros envolvidos com a taxa de remoção não homogênea, principalmente as taxas de remoção, não serão razoáveis e, então, um modelo com taxa de remoção homogênea deve ser ajustado.

5 *Aplicação*

Este penúltimo capítulo apresenta os dois modelos estudados no Capítulo 3 e 4 aplicados ao conjunto de dados reais.

O conjunto de dados reais, apresentado no Capítulo 1, consiste em três séries diárias de dados referentes ao número de indivíduos suscetíveis que se tornaram expostos, expostos que se tornaram infectantes e infectantes que se tornaram removidos tal que as séries de dados são incompletas na proporção, respectivamente, de 99,6%, 8% e 25% sendo que o tamanho da epidemia é igual a 316 indivíduos.

O início da epidemia é no dia 31 de Dezembro de 1994 e o término é no dia 16 de Julho de 1996, isto é, a epidemia teve uma duração de 199 dias com medidas de intervenção introduzidas no dia 130, ou seja, 9 de Maio de 1995.

A distribuição a priori não informativa para todos os parâmetros é escolhida por apresentar bons resultados nas simulações e os dois modelos de taxa de remoção homogênea e taxa de remoção não homogênea são considerados.

Na seção 5.1, o modelo SEIR com taxa de remoção homogênea é aplicado; na seção 5.2, por sua vez, o modelo com taxa de remoção não homogêna é considerado; por fim, na seção 5.3, os resultados são discutidos.

5.1 Modelo SEIR com taxa de remoção homogênea

O modelo SEIR com taxa de remoção homogênea apresentado no Capítulo 3 e dado por (3.1) - (3.4) é considerado.

Os gráficos para a definição das condições da simulação são dados na Figuras 111,

110 e 112. Os detalhes das simulações estão na Tabela 21. Os resultados das simulações seguem nas Figuras 48, 49.

Tabela 21: Detalhes das simulações considerando o conjunto de dados reais no modelo com taxa de remoção homogênea

Priori	TAA	EA	TC (horas)
Não informativa	100000	8000	80,56
Parâmetros	**Estatística R**		**Taxas de aceitação (%)**
Priori não informativa			
B	-		14,62
C	-		24,17
D	-		31,25
β	1,022		44,44
ρ	1,041		74,65
γ	1,000		44,40
q	1,068		74,95

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, TC = Tempo

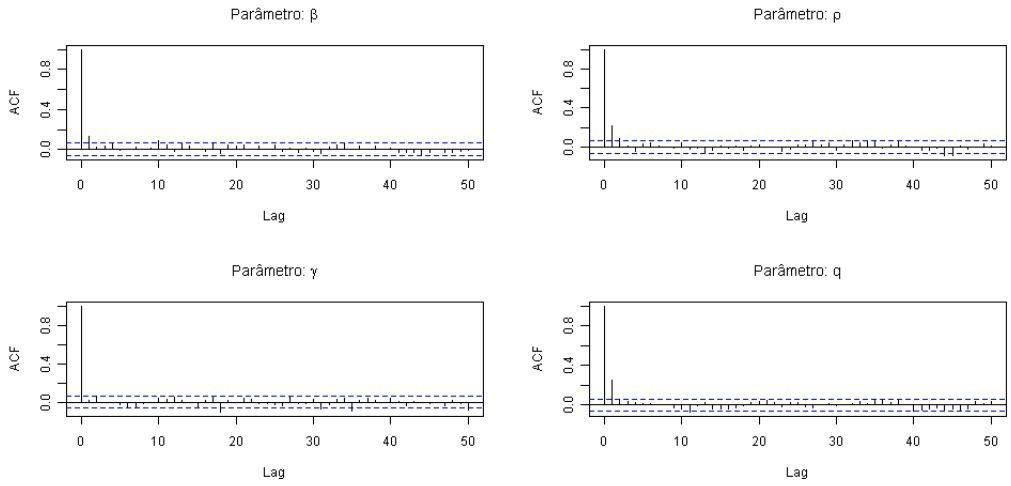


Figura 48: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção homogênea

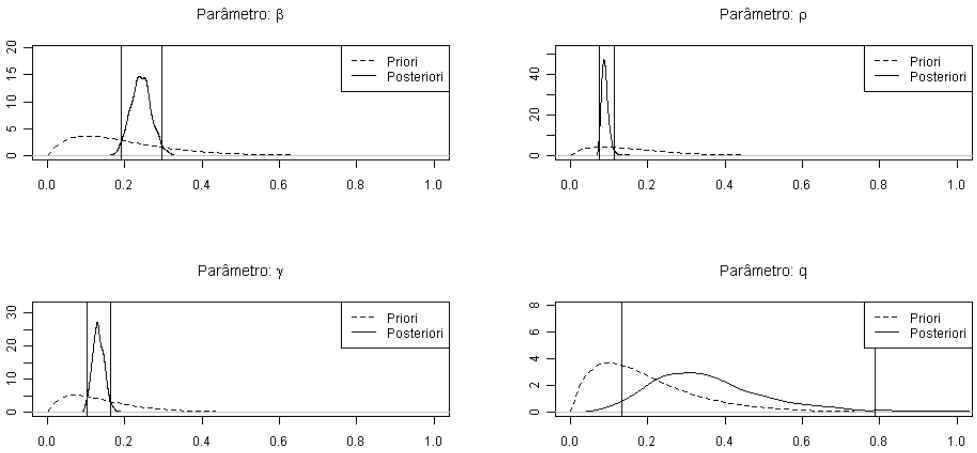


Figura 49: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção homogênea

5.2 Modelo SEIR com taxa de remoção não homogênea

Agora, o modelo SEIR com taxa de remoção não homogênea apresentado no Capítulo 3 em (3.1) - (3.4) adicionando as modificações dadas pelas equações (4.1) - (4.3) do Capítulo 4 é utilizado.

Os gráficos para a definição das condições da simulação são dados na Figuras 114, 113 e 115. Os detalhes das simulações estão na Tabela 22. Os resultados das simulações seguem nas Figuras 50, 51.

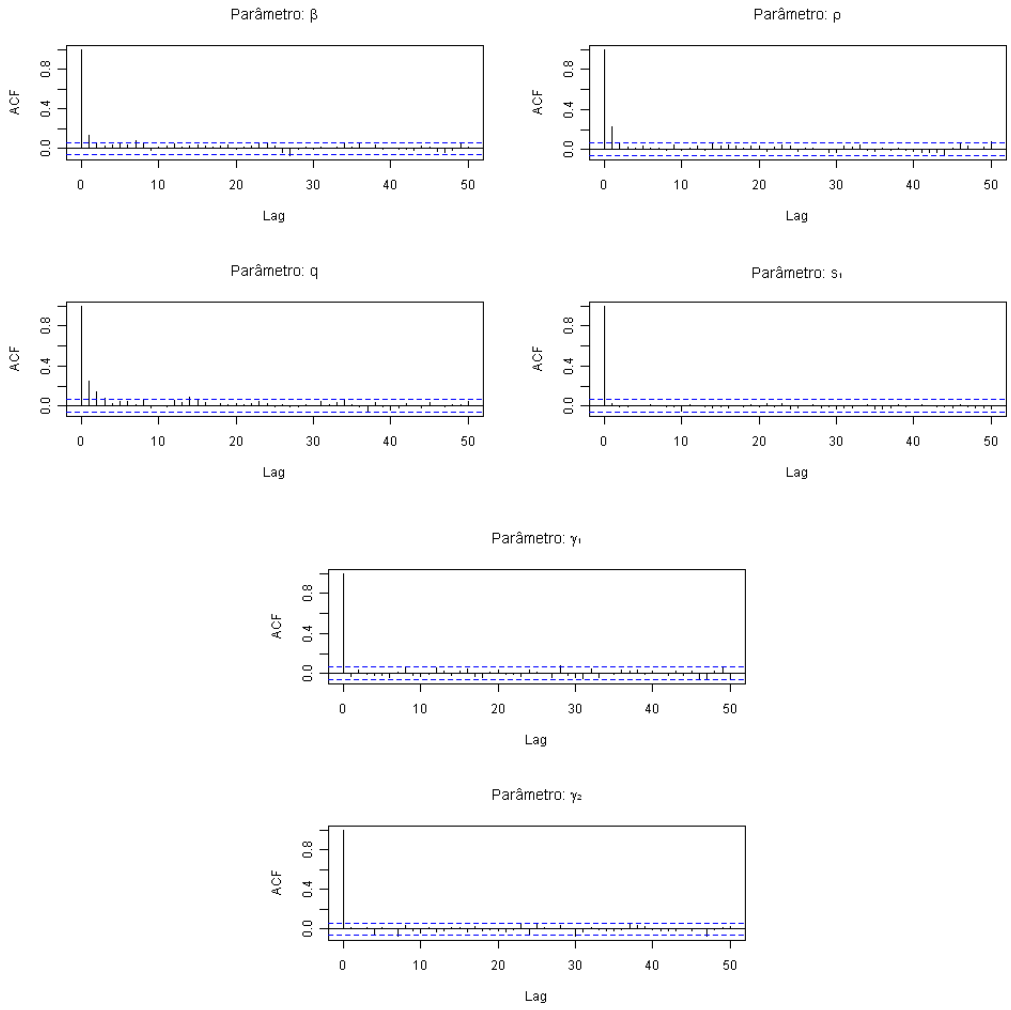


Figura 50: Autocorrelações para as amostras geradas dos parâmetros considerando espaçamento amostral e a distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção não homogênea

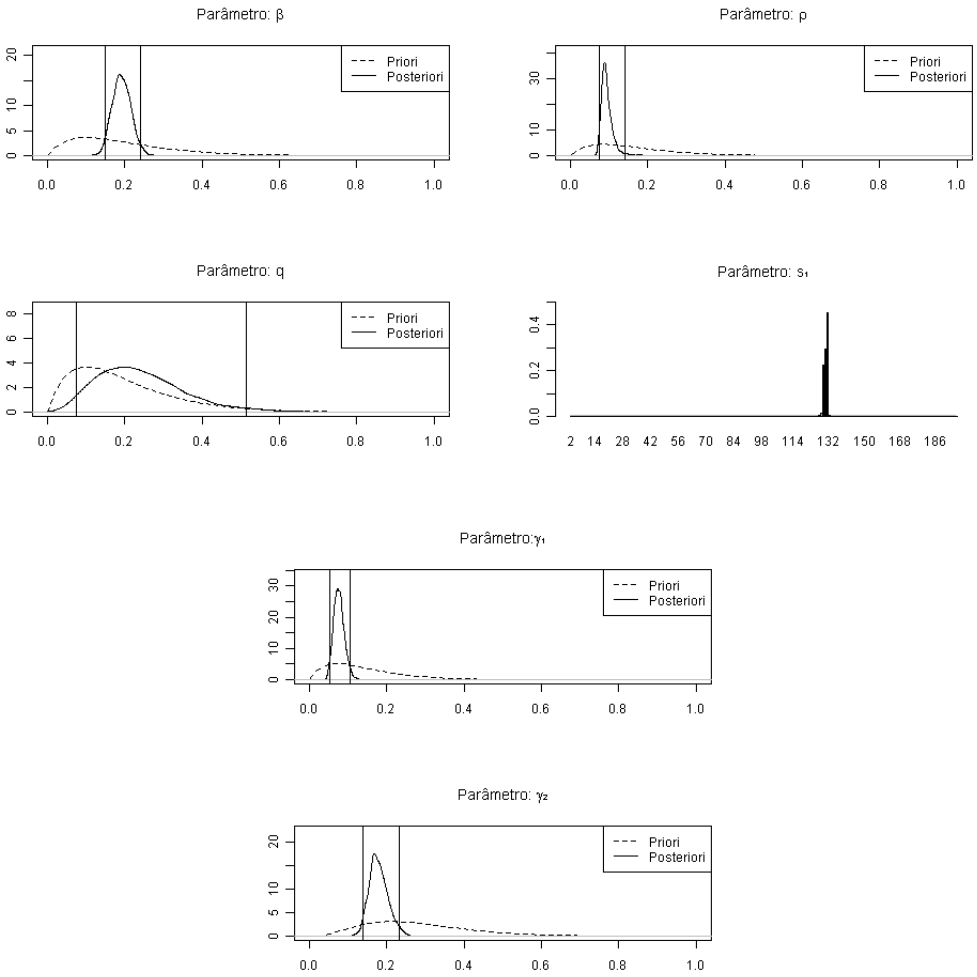


Figura 51: Densidades a priori e a posteriori, valores verdadeiros (pontos) para os parâmetros considerando distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção não homogênea

Tabela 22: Detalhes das simulações considerando o conjunto de dados reais no modelo com taxa de remoção não homogênea

Priori	TAA	EA	TC (horas)
Não informativa	100000	8000	80,77
Parâmetros	Estatística R		Taxas de aceitação (%)
Priori não informativa			
B	-		17,47
C	-		30,94
D	-		26,25
β	1,000		44,46
ρ	1,010		74,67
q	1,000		73,81
γ_1	1,001		44,36
γ_2	1,001		44,23
s_1	1,000		1,05

TAA = Tamanho da amostra de aquecimento, EA = Espaçamento amostral, TC = Tempo

5.3 Conclusões

A Tabela 23 apresenta as estatísticas descritivas obtidas para cada modelo.

A seleção para modelos SEIR é um atual desafio como apontado por O'Neill (2010) e como visto, o DIC também não é adequado para os modelos considerados. Boys e Giles (2007) utilizam o algoritmo *Reversible Jump* para a seleção de modelos, todavia, optou-se em não utilizá-lo pelas dificuldades computacionais enfrentadas. Portanto, a seleção dos modelos ocorreu via a distribuição a posteriori do ponto de mudança e das taxas de remoção.

Sob este critério, é fácil ver que o modelo com taxa de remoção não homogênea é adequado ao conjunto de dados reais devido ao pequeno desvio padrão da distribuição a posteriori do ponto de mudança dada na Figura 51, mesmo sendo considerada uma distribuição a priori não informativa para s_1.

Ao comparar as estimativas em comum dos dois modelos, pode-se ver que a estimativa pontual de β do modelo com taxa de remoção não homogênea é menor do que a estimativa dado pelo modelo com taxa de remoção homogênea, assim como os desvios padrão de suas respectivas distribuições a posteriori; a estimativa pontual de ρ nos

Tabela 23: Estatísticas descritivas das distribuições a posteriori considerando o conjunto de dados real para os modelos SEIR com taxa de remoção homogênea e não homogênea

β			
Modelo	Média	Desvio Padrão	IC 95%
Homogênea	0,242	0,027	(0,199 ; 0,286)
Não-Homogênea	0,192	0,024	(0,155 ; 0,231)
ρ			
Modelo	Média	Desvio Padrão	IC 95%
Homogênea	0,090	0,010	(0,077 ; 0,107)
Não-Homogênea	0,095	0,018	(0,076 ; 0,123)
γ			
Modelo	Média	Desvio Padrão	IC 95%
Homogênea: γ	0,132	0,015	(0,108 ; 0,157)
Não-Homogênea: γ_1	0,076	0,014	(0,055 ; 0,100)
Não-Homogênea: γ_2	0,179	0,025	(0,143 ; 0,223)
q			
Modelo	Média	Desvio Padrão	IC 95%
Homogênea	0,365	0,163	(0,159 ; 0,670)
Não-Homogênea	0,242	0,114	(0,089 ; 0,459)
s			
Modelo	Média	Desvio Padrão	IC 95%
Não-Homogênea	130,254	2,202	(129 ; 131)

dois modelos são similares, todavia, o desvio padrão da distribuição a posteriori para o modelo com taxa de remoção não homogênea é muito maior do que o desvio padrão para o modelo com taxa de remoção homogênea.

As estimativas de q são bastante distintas: a estimativa de q no modelo com taxa de remoção homogênea é muito maior do que a estimativa no modelo com taxa de remoção não homogênea, pois no primeiro modelo q agrega todo o impacto das medidas de intervenção na taxa de contato, enquanto, no segundo modelo, o impacto das medidas de intervenção é considerado tanto na taxa de contato quanto na taxa de remoção. Note também a grande redução do desvio padrão da distribuição a posteriori.

As estimativas de γ não são diretamente comparáveis, contudo, é possível observar que a estimativa de γ para o modelo com taxa de remoção homogênea é a média das estimativas de γ_1 e γ_2 do modelo com taxa de remoção não homogênea.

Por fim, para avaliar o impacto de tais mudanças nas estimativas destes parâmetros é importante avaliar o número de reprodutibilidade efetivo. O número de reprodutibilidade efetivo $R(t)$ é o número esperado de novos casos secundários que um caso primário é capaz de produzir em uma população totalmente suscetível no período t.

Para os modelos considerados, $R(t)$ é dado segundo Chowell et al. (2004) por

$$R_0(t) = \frac{\beta(t)}{\gamma(t)} \tag{5.1}$$

Se $R_0(t) < 1$, a epidemia será extinta em um tempo finito, se $R(t) > 1$ a epidemia será desencadeada com probabilidade positiva, por isso, $R(t)$ é uma importante quantidade para os epidemiologistas a fim de analisar o impacto das medidas de interveção.

A distribuição a posteriori de $R(t)$ pode ser calculada a partir da amostra obtida dos outros parâmetros. Então, as estimativas pontuais dada pela média e os desvios padrão das distribuições a posteriori de $R(t)$ para cada período t considerando os dois modelos são apresentadas na Figura 52.

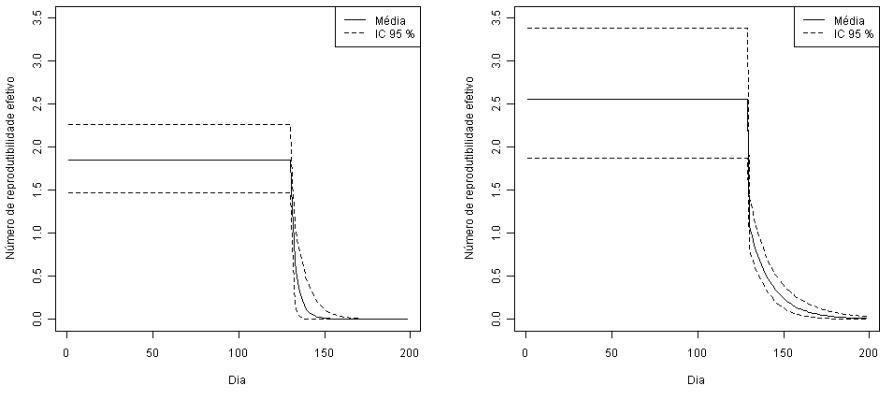


Figura 52: Número de reprodutibilidade efetivo para o modelo com taxa de remoção homogênea e com taxa de remoção não homogênea, respectivamente

A partir do modelo de taxa de remoção não homogênea, observa-se que a estimativa pontual do número de reprodutibilidade efetivo antes da introdução das medidas de interveção assumia um valor muito superior à estimativa dada pelo modelo com taxa de remoção homogênea.

Nos dois modelos, $R(t)$ torna-se menor do que 1 somente a partir do dia 132, sendo que as medidas de intervenção foram introduzidas no dia 130, portanto, o modelo de taxa de remoção não homogênea permite afirmar que o impacto inicial estimado das medidas de intervenção é maior do que o impacto inicial estimado pelo modelo com taxa de remoção homogênea, pois a estimativa do número de reprodutibilidade efetivo dado pelo modelo de taxa de remoção não homogênea é maior do que a estimativa dada pelo modelo de taxa de remoção homogênea antes das medidas de intervenção.

Por fim, é possível observar que o decaimento do número de reprodutibilidade efetivo após a introdução das medidas de intervenção é menor no modelo de taxa de remoção não homogênea, indicando que o impacto posterior das medidas de intervenção foi menor segundo o modelo de taxa de remoção não homogênea em relação ao modelo de taxa de remoção homogênea.

6 *Conclusões e Trabalhos Futuros*

Neste trabalho, em um primeiro instante, o modelo de Lekone e Finkenstädt (2006) foi discutido e, a partir das simulações com dados completos e incompletos, pode-se observar que o método de estimação segundo a abordagem bayesiana e os métodos MCMC é funcional no sentido de obter estimativas similares aos valores verdadeiros e com pequenos desvios padrão das distribuições a posteriori, desde que a probabilidade a priori de uma vizinhança dos valores verdadeiros não seja próxima de zero, como é o caso da distribuição a priori não informativa e informativa estabelecidas no Capítulo 3.

Se a distribuição a priori apresenta tal probabilidade próxima de zero, como é o caso da distribuição a priori não centrada definida no Capítulo 3, as estimativas pontuais não são próximas dos valores verdadeiros e alguns intervalos de credibilidade 95% não contém tais valores. Todavia, o método ainda é funcional, pois é possível observar uma grande mudança da distribuição a posteriori em relação a distribuição a priori na direção dos valores verdadeiros.

Cabe ressaltar, não foi possível a utilização da distribuição uniforme como distribuição a priori dos parâmetros devido a problemas de maximização da distribuição a posteriori do parâmetro q na presença de dados incompletos para a implementação do algoritmo de Metropolis-Hastings.

Além disso, a introdução da incerteza devido aos dados incompletos pode ser observada no aumento do desvio padrão das distribuições a posteriori de todos os parâmetros assim como também o aumento da exigência e do tempo computacional.

É importante observar que há uma grande dificuldade na estimação do parâmetro

q, que representa o decaimento da taxa de contato após a introdução das medidas de intervenção, pois a quantidade de informação dos dados é menor para este parâmetro do que em relação aos outros, já que a informação dada pelos dados sobre q limita-se ao período posterior a introdução das medidas de intervenção.

Em um segundo instante, o modelo de Lekone e Finkenstädt (2006) foi modificado similarmente a Boys e Giles (2007) que consideram uma taxa de remoção não homogênea com um número aleatório de pontos de mudança assim como suas respectivas posições. Neste trabalho, somente um ponto de mudança de posição aleatória foi considerado a fim de também considerar o impacto das medidas de intervenção na taxa de remoção, tornando o modelo mais razoável na presença das medidas de intervenção.

Desta forma, duas distribuições a priori foram consideradas e simulações foram realizadas com dados completos e incompletos no Capítulo 4. O método de estimação segundo a abordagem bayesiana e os métodos MCMC é funcional, contudo, pode-se observar que tanto na presença de dados completos como dados incompletos e uma distribuição a priori não informativa para o ponto de mudança, há dificuldades na estimação do ponto de mudança, sendo que a estimativa pontual no caso de dados incompletos é distante do valor verdadeiro e a distribuição a posteriori deste parâmetro apresentou um grande desvio padrão em ambos os casos.

A utilização de uma distribuição a priori informativa para o ponto de mudança melhora significativamente este cenário em relação a estimativa pontual, todavia, o desvio padrão da distribuição a posteriori ainda permanece muito similar ao desvio padrão da distribuição a priori.

E para a escolha entre os dois modelos com taxa de remoção homogênea e taxa de remoção não homogênea, uma discussão sobre erro de especificação foi realizada no Capítulo 4, tal que foi possível observar que as distribuições a posteriori das taxas de remoção e do ponto de mudança no modelo com taxa de remoção não homogênea são indicadores razoáveis da necessidade de um ponto de mudança na taxa de remoção.

Portanto, um modelo com taxa de remoção não homogênea deve ser ajustado e caso, as distribuições a posteriori relacionadas a taxa de remoção não sejam coerentes, um modelo com taxa de remoção homogênea deve ser escolhido. É importante dizer

que houve a tentativa da utilização do DIC, contudo, foi verificado que este não é um bom critério para a seleção neste tipo de modelo.

Por fim, os dois modelos com taxa de remoção homogênea e não homogênea foram aplicados aos dados reais e foi possível observar um ótimo ajute do modelo com taxa de remoção não homogênea aos dados reais, diferentemente de como foi observado nas simulações, o desvio padrão da distribuição a posteriori do ponto de mudança foi pequeno em relação ao desvio padrão da distribuição a priori não informativa.

A utilização de um modelo com taxa de remoção não homogênea apresentou uma mudança significativa na análise do número de reprodutibilidade efetivo, tal que as medidas de intervenção apresentaram um impacto inicial mais significativo e um impacto posterior menos significativo do que aqueles estimados pelo modelo com taxa de remoção homogênea. Por isso, um modelo que permita descrever o impacto das medidas de intervenção mais precisamente assim como o modelo SEIR com taxa de remoção não homogênea e também taxa de contato não homogênea tem grande importância na quantificação das consequências das medidas de intervenção introduzidas em uma epidemia.

O'Neill (2010) aponta os vários desafios a serem enfrentados na modelagem de epidemias. Nesta dissertação, também foi possível observar as dificuldades citadas por O'Neill (2010): a utilização de dados aumentados não é eficiente computacionalmente e não há métodos adequados para a seleção e o diagnóstico dos modelos SEIR.

O modelo SEIR discutido considera somente um ponto de mudança na taxa de remoção, sendo que é interessante permitir a existência de mais pontos de mudança tal que o número de pontos de mudança seja aleatório exatamente como em Boys e Giles (2007). Nesta direção, a utilização de outras distribuições além da distribuição Exponencial surge como um grande desafio e Stretfaris e Gibson (2004) apontam caminhos nesta direção.

Por fim, o método de estimação apresentou dificuldades para estimar o parâmetro q que representa o decaimento da taxa de contato após as medidas de intervenção e do ponto de mudança tal que para este último, a dificuldade somente foi encontrada nas simulações, portanto, a modelagem da introdução das medidas deve ser discutida.

Referências

ABBEY, H. An examination of the reed-frost theory of epidemics. *Human Biology*, v. 29, p. 66 – 77, 1952.

AITCHINSON, J.; DUNSMORE, I. R. *Statiscal prediction analysis*. Cambridge: Cambridge University Press, 1975.

ALBERT, J. *Bayesian computation with R (Use R)*. 2ª. ed. New York: Springer, 2004.

ANDERSON, R. M.; MAY, R. M. *Infectious diseases of humans*. 1ª. ed. New York: Oxford University Press, 1991.

BAILEY, N. T. J. *The mathematical theory of infections diseases and its applications*. 2ª. ed. London: Griffin, 1975.

BARTLETT, M. S. Some evolutionary stochastic processes. *Journal Royal Statistical Society*, v. 11, p. 211 – 229, 1949.

BECKER, N. G. *Analysis of infectious disease data*. London: Chapman & Hall, 1989.

BENNETT, J. E.; RACINE-POON, A.; WAKEFIELD, J. C. *MCMC for nonlinear hierarchical models*. London: Chapman & Hall, 1996. 339 - 357 p.

BERNARDO, J. M. Reference posterior distributions for bayesian inference (with discussion). *Journal of the Royal Statistical Society*, v. 41, p. 113 – 147, 1979.

BERNARDO, J. M.; SMITH, A. F. M. *Bayesian theory*. New York: Wiley, 1994.

BIRMINGHAM, K.; COONEY, S. Ebola: small, but real progress (news feature). *Nature Medicine*, v. 8, p. 313, 2002.

BOYS, R. J.; GILES, P. R. Bayesian inference for stochastic epidemic models with time-inhomogeneous removal rates. *Journal Mathematical Biology*, v. 55, p. 233 – 247, 2007.

BREMAN, J. G.; PIOT, P.; JOHNSON, K. M. The epidemiology of ebola hemorrhagic fever in zaire, 1976. In: *Proceedings of the International Colloquium on Ebola Virus Infections*. Antwerp, Belgium: Elsevier/North Holland Biomedical Press, 1977. p. 103 – 124.

BRITTON, T.; BECKER, N. G. Estimating the immunity coverage required to prevent epidemics in a community of households. *Biostatistics*, v. 1, p. 389 – 420, 2000.

BROOKS, S. P.; ROBERTS, G. O. Convergence assessment techniques for markov chain monte carlo. *Biometrika*, v. 8, n. 3, p. 319 – 335, 1998.

BROWNLEE, J. Statistical studies in immunity: the theory of an epidemic. *Proceedings of the Royal Society of Edinburgh*, v. 26, p. 484 – 521, 1906.

CASELLA, G.; GEORGE, E. I. Explaining the gibbs sampler. *The American Statistician*, v. 46, n. 3, p. 167 – 174, 1992.

CDDP. *Ebola hemorrhagic fever information packet*. Centers for Disease Control and Prevention - E.U.A., 2009. Disponível em: <http://www.cdc.gov/ncidod/dvrd/spb-/mnpages/dispages/ebola.htm>.

CELEUX, G. et al. Deviance information criteria for missing data models. *Bayesian Analysis*, v. 1, p. 651 – 674, 2006.

CHIB, S.; GREENBERG, E. Bayes inference in regression models with arma (p, q) errors. *Journal of Econometrics*, v. 64, p. 183 – 206, 1994.

CHIB, S.; GREENBERG, E. Understanding the metropolis-hastings algorithm. *The American Statistician*, v. 49, n. 4, p. 327 – 335, 1995.

CHOWELL, G. et al. The basic reproductive number of ebola and the effects of public health measures the cases of congo and uganda. *Journal of Theorethical Biological*, v. 229, p. 119 – 126, 2004.

COLOSIMO, E. A.; GIOLO, S. R. *Análise de Sobrevivência Aplicada*. 1ª. ed. São Paulo: Edgar Blücher, 2006.

COWLES, M. K.; CARLIN, B. P. Markov chain monte carlo convergence diagnostic: a comparative review. *Journal of the American Statistical Association*, v. 91, p. 883 – 904, 1996.

DEMIRIS, N. *Bayesian inference for stochastic epidemic models using Markov chain Monte Carlo methods*. Tese (Doutorado) — University of Nottingham, 2004.

ERKANLI, A. Laplace approximations for posterior expectations when the mode occurs on the boundary of the parameter space. *Journal of the American Statistical Association*, v. 89, p. 250 – 258, 1994.

EVANS, G. H. Some arithmetical considerations of the progress of epidemics. *Transactions of the Epidemiological Society*, p. 551, 1875.

FARR, W. Progress of epidemics. *Second report of the Register General of England and Wales*, p. 91 – 98, 1840.

GAMERMAN, D.; LOPES, H. F. *Markov chain Monte Carlo: Stochastic simularion for Bayesian inference*. 2ª. ed. New York: Chapman & Hall/CRC, 2006.

GEISSER, S. *Predictive inference: an introduction*. London: Chapman & Hall, 1993.

GELFAND, A. E.; SMITH, A. F. M. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, v. 85, p. 398 – 409, 1990.

GELMAN, A. Iterative and non-iterative simulation algorithms. *Computing Science and Statistics (Interface Proceedings)*, v. 24, p. 433 – 438, 1992.

GELMAN, A. Inference and monitoring convergence. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Ed.). *Markov chain Monte Carlo in practice*. London: Chapman & Hall, 1996. p. 131 – 143.

GELMAN, A. *Bayesian data analysis*. 2ª. ed. Londres: Chapman & Hall/CRC, 2004.

GELMAN, A.; MENG, X.-L.; STERN, H. Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, v. 6, p. 733 – 807, 1996.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. *Statistical Science*, v. 7, n. 4, p. 457 – 472, 1992.

GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions and the bayesian restorating of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 6, p. 721 – 741, 1984.

GEWEKE, J. Evaluating the accuracy of sampling base approaches to the calculation of posterior moments (with discussion). In: BERNARDO, J. M. et al. (Ed.). *Bayesian Statistics 4*. Oxford: Oxford University Press, 1992. p. 169 – 193.

GILKS, W. R.; RICHARDSON, R.; SPIEGELHALTER, D. J. *Markov chain Monte Carlo in practice*. 1ª. ed. London: Chapman & Hall, 1996.

GREEN, P. J. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, v. 82, p. 711 – 732, 1995.

GUTTORP, P. *Stochastic modelling of scientific data*. London: Chapman & Hall, 1995.

HAMMER, W. H. Epidemic disease in england. *Lancet*, v. 1, p. 733 – 739, 1906.

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, v. 57, p. 97 – 109, 1970.

HöHLE, M.; JøRGENSEN, E. *Dina research report 102: Estimating parameters for stochastic epidemics*. 2002. Disponível em: <http://www.husdyr.kvl.dk/˜ark/pub-/dina/report102.pdf>.

HOEL, P. G.; PORT, S. C.; STONE, C. J. *Introduction to Stochastic Processes*. 1ª. ed. New York: Houghton Mifflin Company, 1972.

JEFFREY, H. *Theory of probability*. 3ª. ed. Oxford: Oxford University Press, 1961.

JEWELL, C. P. et al. Bayesian analysis for emerging infectious diseases. *Bayesian Analysis*, v. 4, n. 2, p. 191 – 222, 2009.

KERMARCK, W. O.; MCKENDRICK, A. G. Contributions to the mathematical theory of epidemics. *Proceedings of the Royality Society*, v. 115, p. 700–721, 1927.

KHAN, A. S. et al. The reemergence of ebola hemorrhagic fever, democratic republic of the congo, 1995. *Journal of Infectious Diseases*, v. 179, p. S76 – S86, 1999.

KIM, J. J.; GOLDIE, S. J. Health and economic implications of hpv vaccination in the united states. *New England Journal of Medicine*, v. 359, p. 821 – 832, 2008.

KYPRAIOS, T. *Efficient Bayesian inference for partially observed stochastic epidemics and a new class of semi parametric time series models*. Tese (Doutorado) — Lancaster University, 2007.

LEKONE, P. E.; FINKENSTäDT, B. F. Statistical inference in a stochastic epidemic seir model with control intervention ebola as a case study. *Biometrics*, v. 62, p. 1170 – 1177, 2006.

LIU, J. S.; WONG, W.; KONG, A. Correlation structure and convergence rate of the gibbs sampler: applications to the comparison of estimators and augmentation schemes. *Biometrika*, v. 4, p. 27 – 40, 1994.

MANFREDINI, R. B. *Estimação bayesiana e por máxima verossimilhança de modelos SIR estocásticos*. Dissertação (Mestrado) — Universidade Estadual de Campinas, 2009.

MCBRYDE, E. S. et al. Early transmission characteristics of influenza a(h1n1)v in australia: Victorian state, 16 - may - 3 june 2009. *Eurosurveillance*, v. 14, n. 42, 2009.

MCKENDRICK, A. G. Applications of mathematics to medical problems. *Proceedings Edinburgh Mathematical Society*, v. 14, p. 98 – 130, 1926.

METROPOLIS, N. et al. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, v. 21, p. 1087 – 1092, 1953.

MEYN, S.; TWEEDIE, R. L. *Markov Chains and stochastic stability*. 2ª. ed. New York: Cambrigde University Press, 2009.

MEYN, S. P.; TWEEDIE, R. L. Computable bounds for convergence rates of markov chains. *Annals of Applied Probability*, v. 4, p. 124 – 148, 1994.

MIGNON, H. S.; GAMERMAN, D. *Statistical inference: an integrated approach*. London: Arnold, 1999.

MULLER, P. *Metropolis based posterior integration schemes*. 1991.

NEAL, P.; ROBERTS, G. Optimal scaling for partially updating mcmc algorithms. *The Annals of Applied Probability*, v. 16, n. 2, p. 475 – 515, 2006.

NTZOUFRAS, I. *Bayesian models using Winbugs*. 1ª. ed. New Jersey: Jogn Wiley & Sons, 2009.

O'NEILL, P. D. Introduction and snapshot review: Relating infectious disease transmission models to data. *Statistics in Medicine*, v. 29, p. 2069 – 2077, 2010.

O'NEILL, P. D.; ROBERTS, G. O. Bayesian inference for partially observed stochastic epidemics. *Journal Royal Statistical Society*, v. 162, p. 121 – 129, 1999.

PATZ, R. J.; JUNKER, B. W. A straightforward approach to markov chain monte carlo methods for item response models. *Journal of Educational and Behavioral Statistics*, v. 24, n. 2, p. 146 – 148, 1999.

PAULINO, C. D.; SILVA, G.; ACHCAR, J. A. Bayesian analysis of correlated misclassified binary data. *Computational Statistics & Data Analysis*, v. 49, p. 1120 – 1131, 2005.

PAULINO, C. D.; TURKAMAN, A. A.; MURTEIRA, B. *Estatística bayesiana*. Lisboa: Fundação Calouste Gulbekian, 2003.

PLUIMERS, F. H. et al. Classical swine fever in the netherlands 1997-1998: a description of organisation and measures to eradicate the disease. *Preventive Veterinary Medicine*, v. 42, n. 3 - 4, p. 139 – 155, 1999.

POLSON, N. G. *Convergence of Markov chain Monte Carlo algorithms (with discussion)*. Oxford: Oxford University Press, 1996. 297 - 321 p.

RAFTERY, A. E.; LEWIS, S. How many iterations in the gibbs sampler? In: BERNANDO, J. M. et al. (Ed.). *Bayesian Statistics 4*. Oxford: Oxford University Press, 1992. p. 763 – 773.

ROBERT, C. P.; CASELLA, R. *Monte Carlo statistical methods*. 2ª. ed. New York: Springer, 2004.

ROBERTS, G. O.; GELMAN, A.; GILKS, W. R. Weak convergence and optimal scaling of random walk metropolis algorithm. *The Annals of Applied Probability*, v. 7, p. 110 – 120, 1997.

ROBERTS, G. O.; ROSENTHAL, J. S. Optimal scaling for various metropolis-hastings algorithms. *Statistical Science*, v. 16, n. 4, p. 351 – 367, 2001.

ROSS, R. *Report on the prevention of Malaria in Mauritius*. London: [s.n.], 1908.

ROSS, R. *The presention of Malaria*. 2ª. ed. London: Murray, 1911.

ROSS, S. *Introduction to probability models*. 9ª. ed. California: Elsevier, 2007.

STRETFARIS, G.; GIBSON, G. J. Bayesian inference for stochastic epidemics in closed populations. *Statistical Modelling*, v. 4, p. 63 – 75, 2004.

TANNER, M. A.; WONG, W. The calculation of posterior distributions by data augmentation (with discussion). *Journal of the American Statistical Association*, v. 82, p. 528 – 550, 1987.

TIERNEY, L. Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics*, v. 22, p. 1701 – 1762, 1994.

TONY, T. et al. Approximate bayesian computation scheme for parameter inference and model selection in dynamic systems. *Journal of Royal Society Interface*, v. 9, p. 187 – 202, 2009.

WHO. *World Organization Health: Ebola hemorrhagic fever*. 2009. Disponível em: <http://www.who.int/mediacentre/factsheets/fs103/en/>.

YANG, H. M. *Epidemiologia matematica: estudos dos efeitos da vacinação em doençcas de transmissão direta*. 1ª. ed. Campinas: Unicamp, 2001.

APÊNDICE A – Gráficos do Capítulo 3

Este apêndice contém as figuras rerente ao Capítulo 3.

A.1 Dados completos

A.1.1 Distribuição a priori não informativa

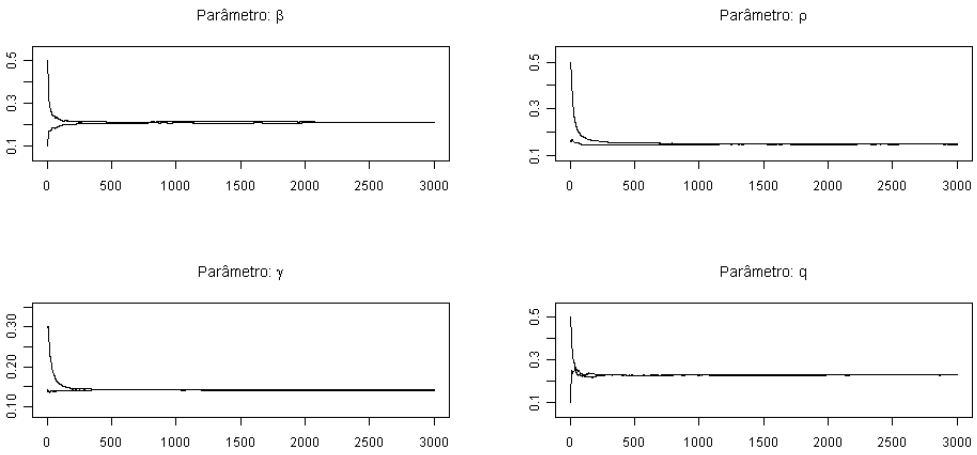


Figura 53: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

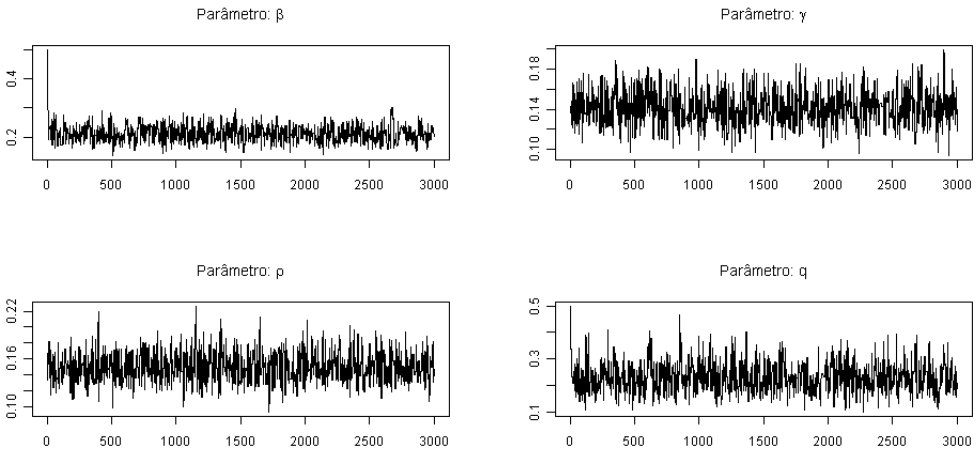


Figura 54: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

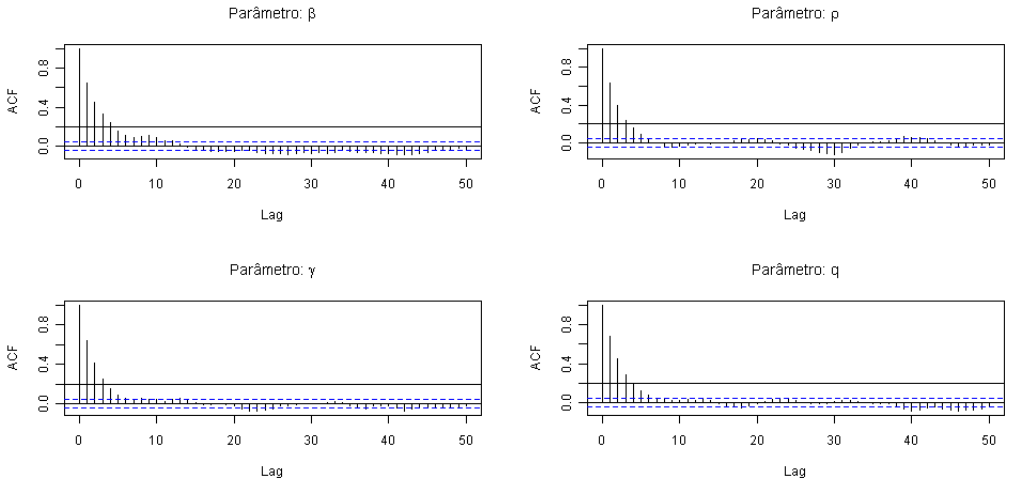


Figura 55: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

A.1.2 Distribuição a priori informativa

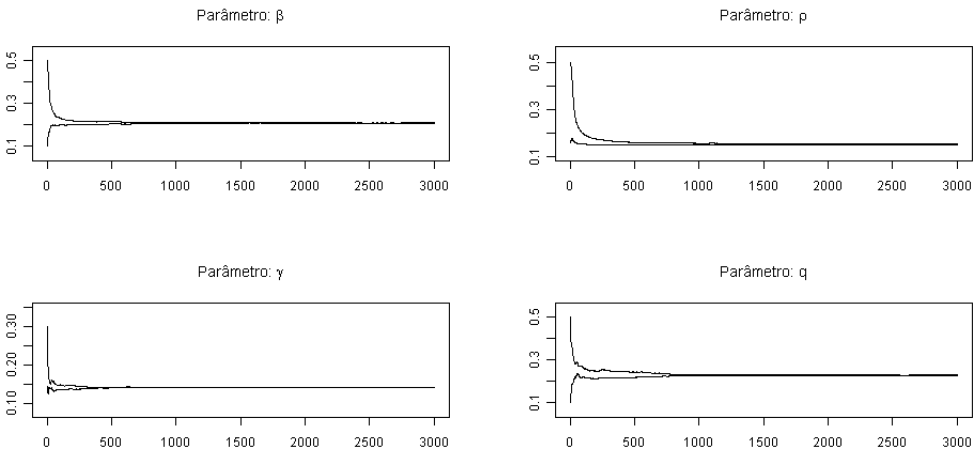


Figura 56: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

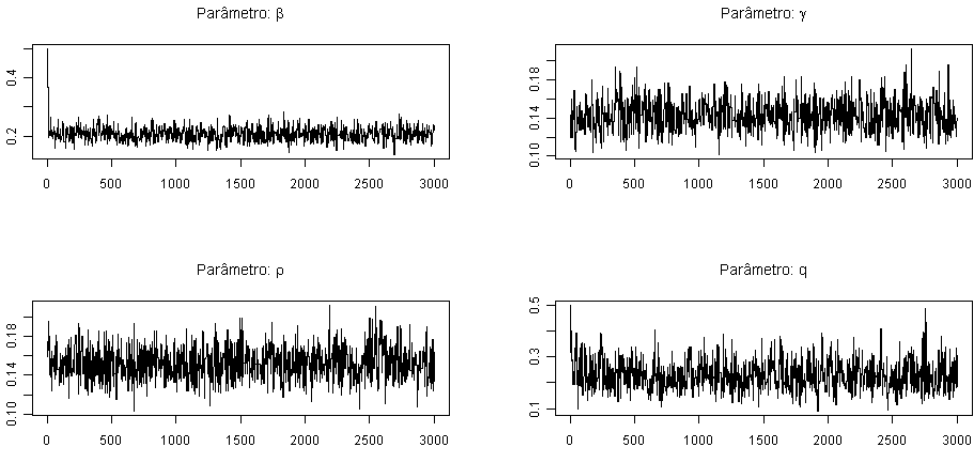


Figura 57: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

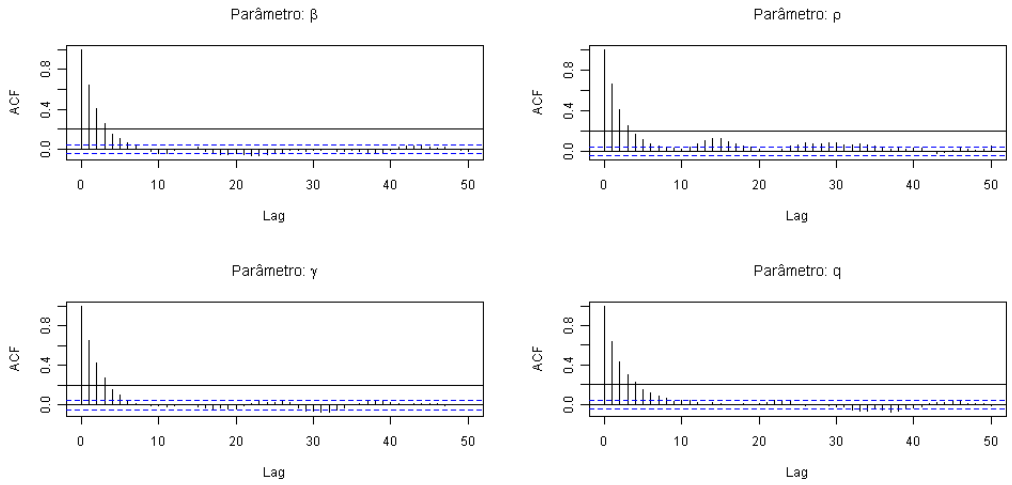


Figura 58: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

A.1.3 Distribuição a priori não centrada

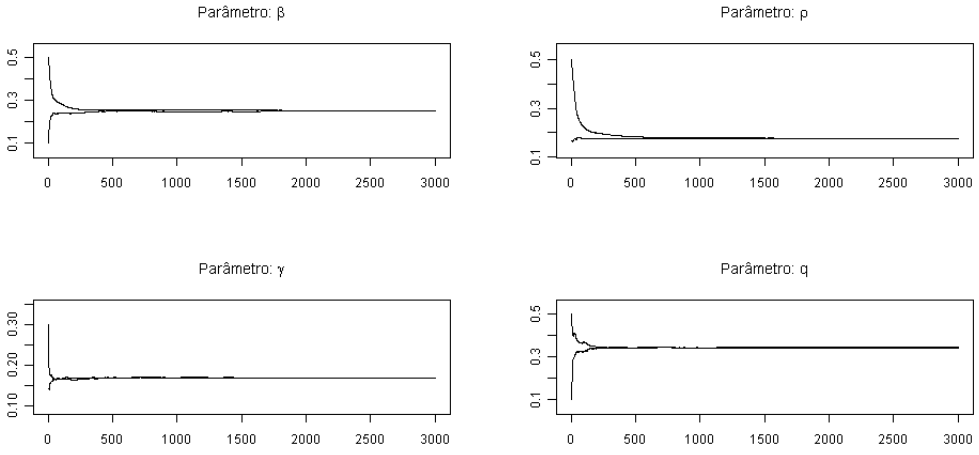


Figura 59: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não centrada com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

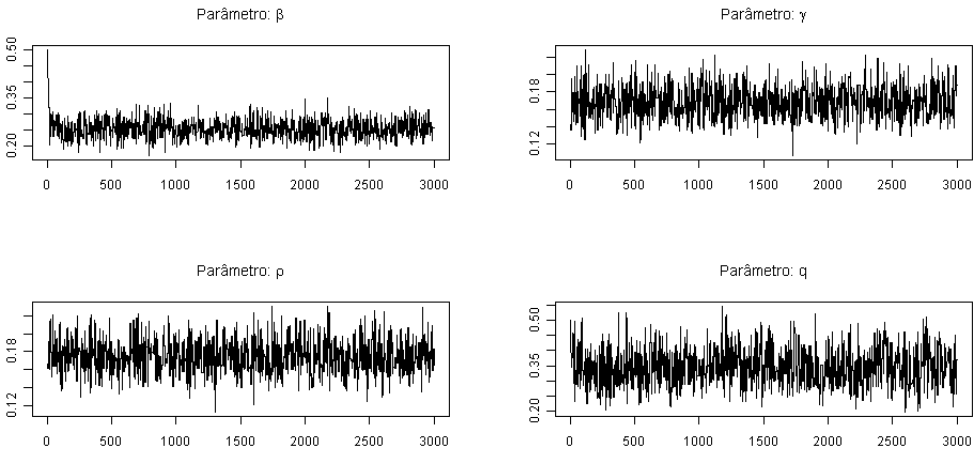


Figura 60: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não centrada com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

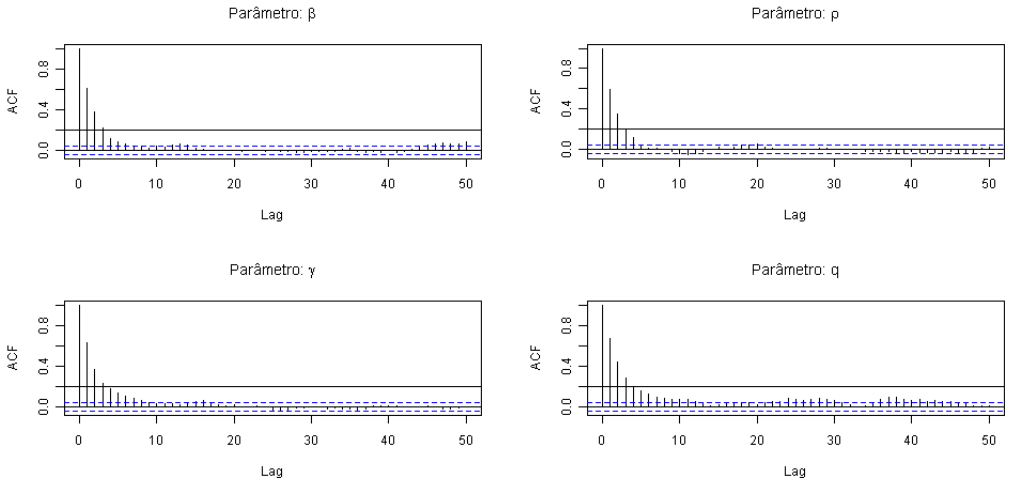


Figura 61: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não centrada com dados completos na análise de sensibilidade do modelo com taxa de remoção homogênea

A.2 B Incompleto

A.2.1 Distribuição a priori não informativa

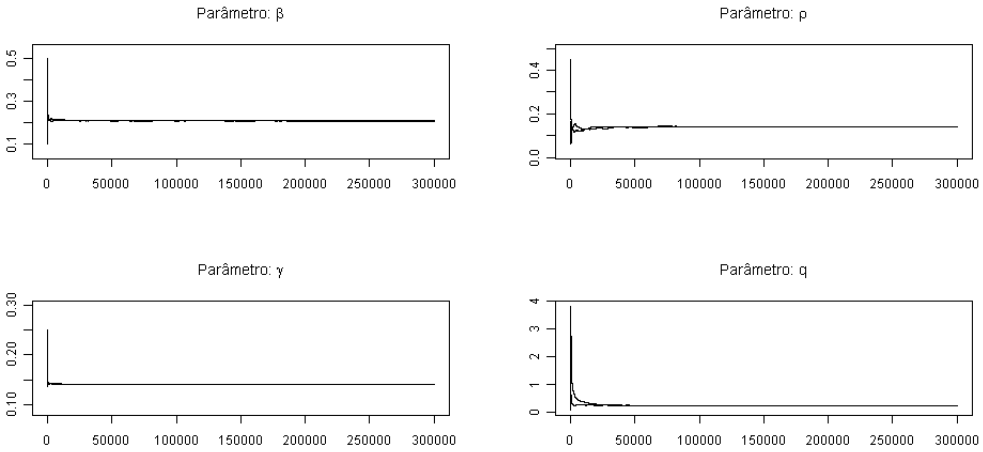


Figura 62: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

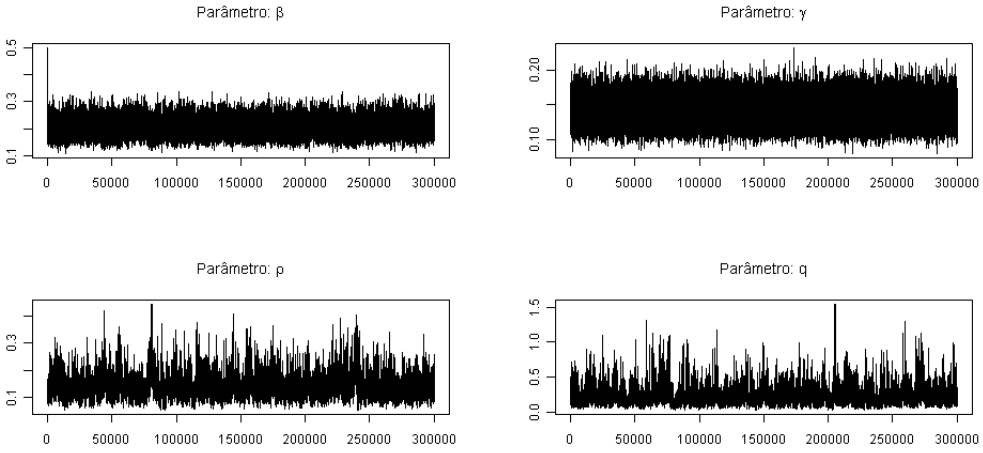


Figura 63: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

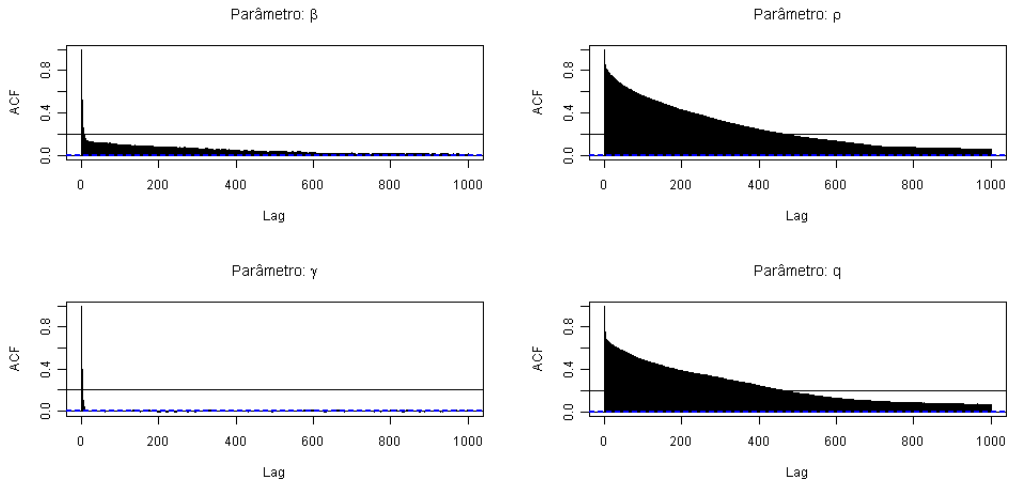


Figura 64: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

A.2.2 Distribuição a priori informativa

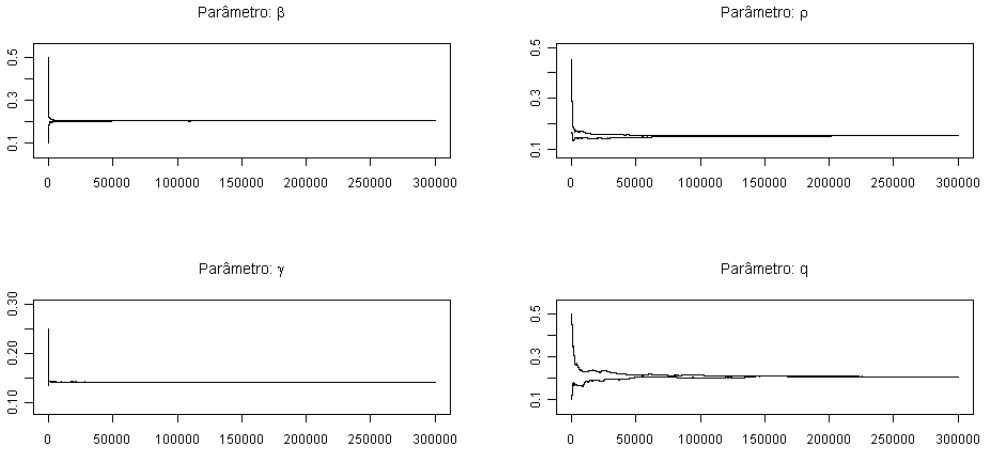


Figura 65: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

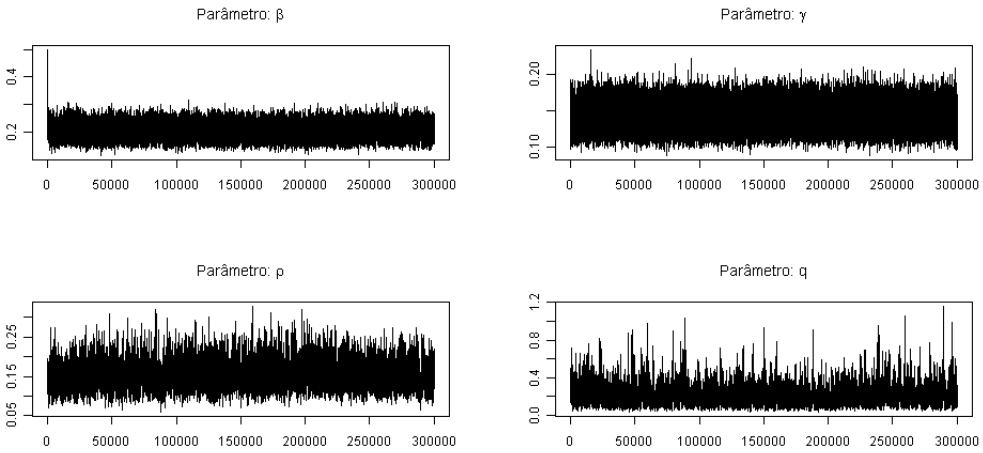


Figura 66: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

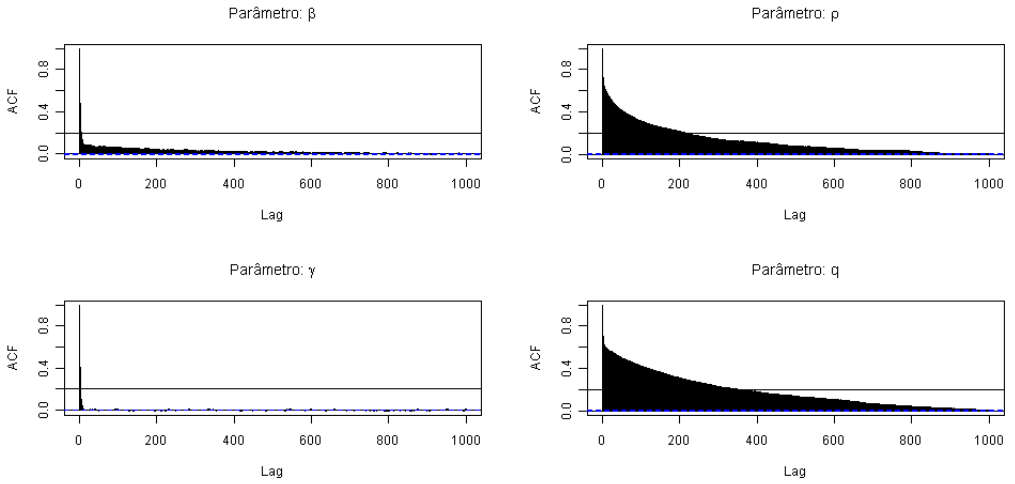


Figura 67: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori informativa com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

A.2.3 Distribuição a priori não centrada

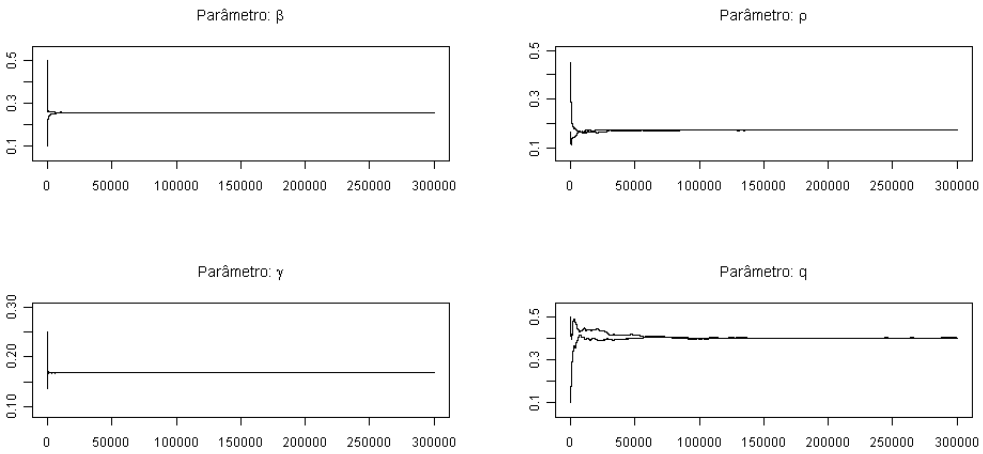


Figura 68: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não centrada com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

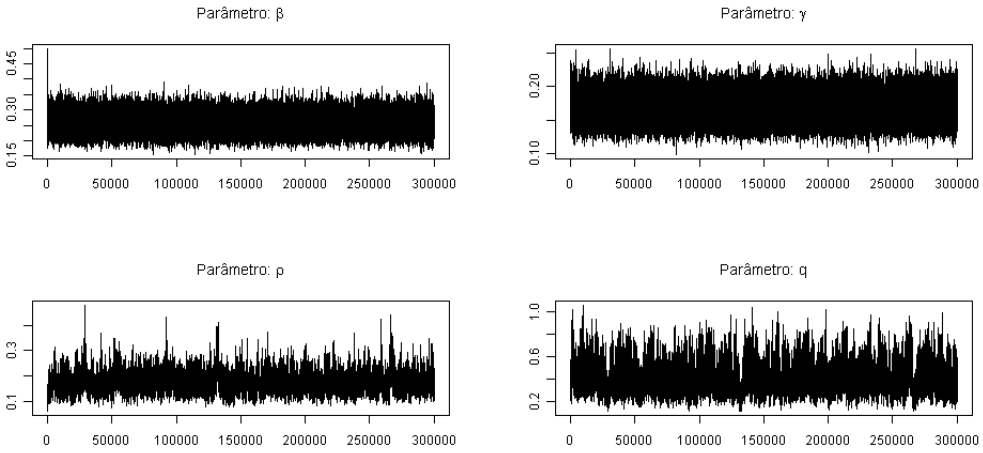


Figura 69: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não centrada com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

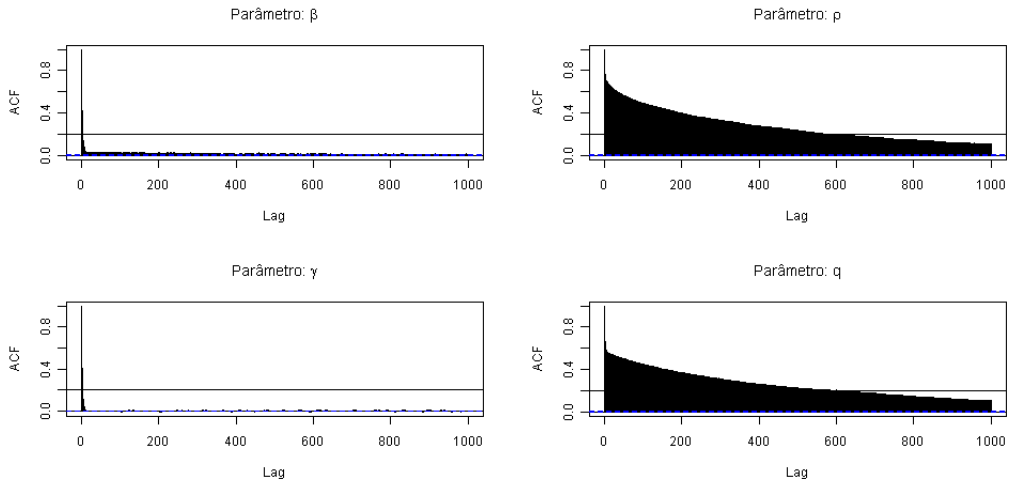


Figura 70: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não centrada com **B** incompleto na análise de sensibilidade do modelo com taxa de remoção homogênea

A.3 B, C e D incompletos

A.3.1 Distribuição a priori não informativa

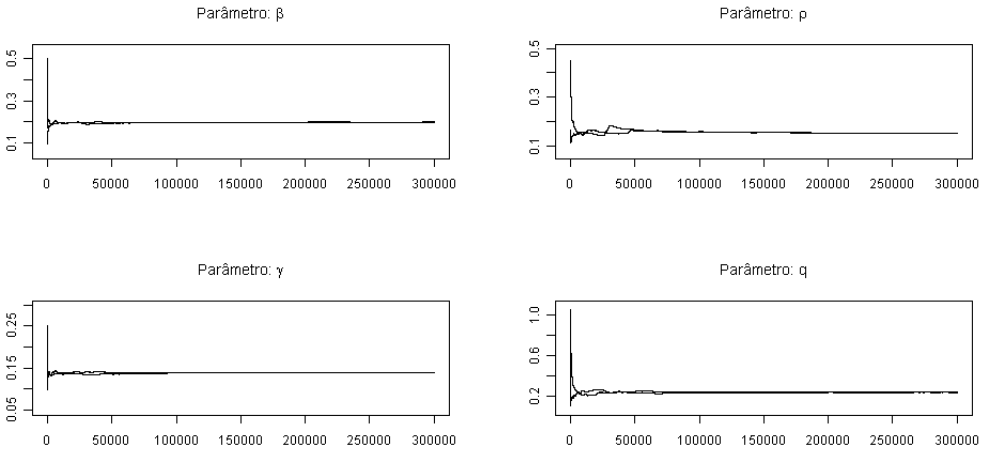


Figura 71: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

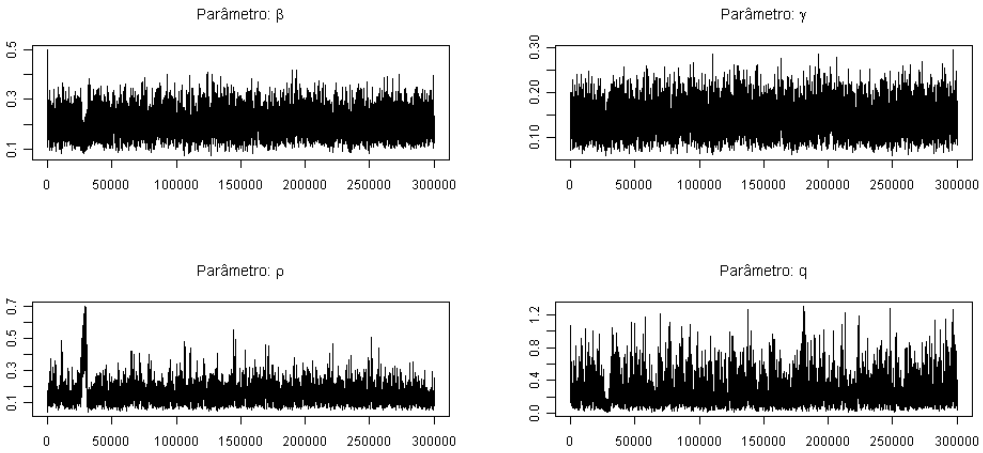


Figura 72: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

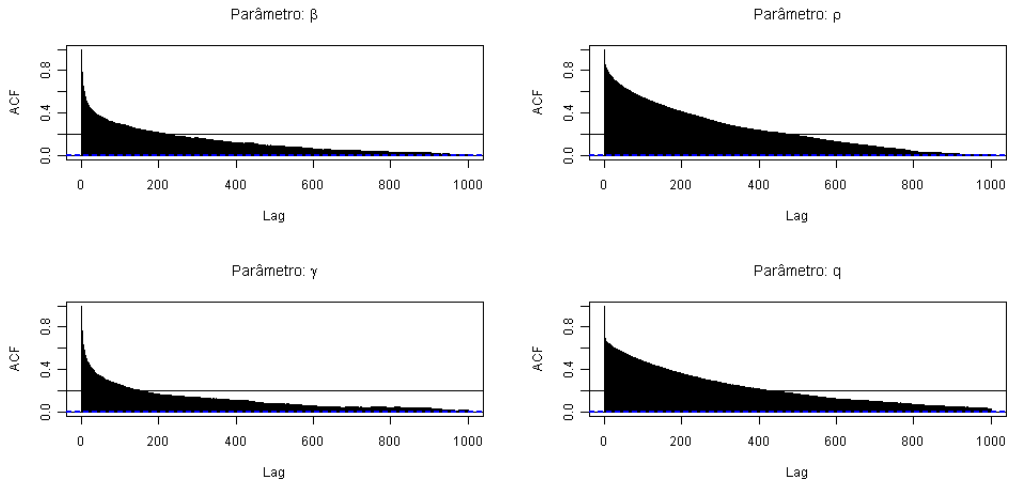


Figura 73: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

A.3.2 Distribuição a priori informativa

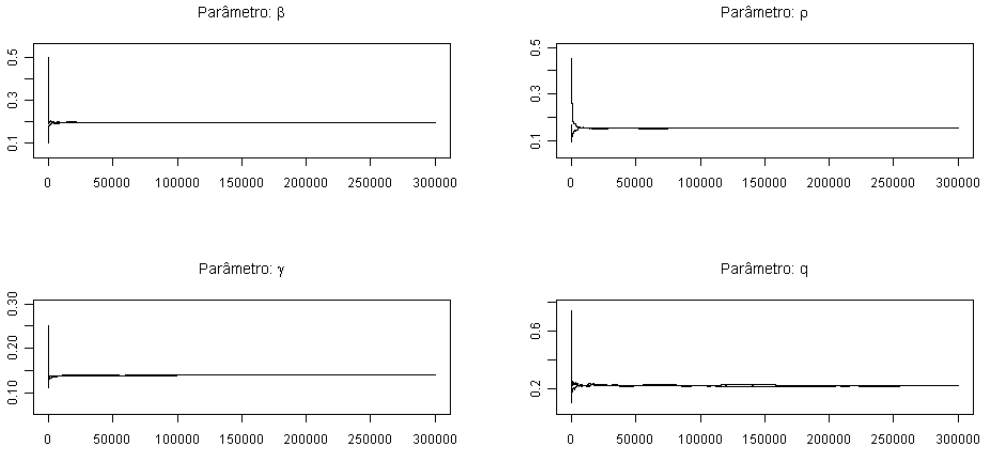


Figura 74: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

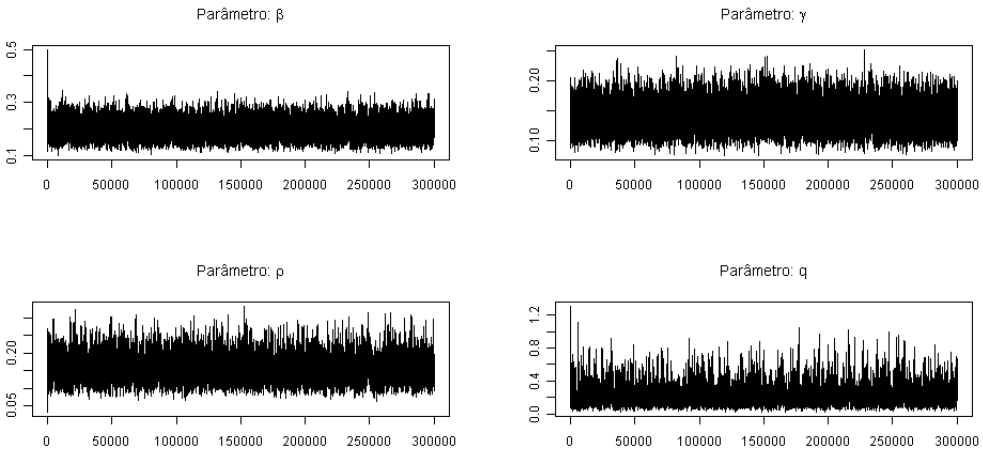


Figura 75: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

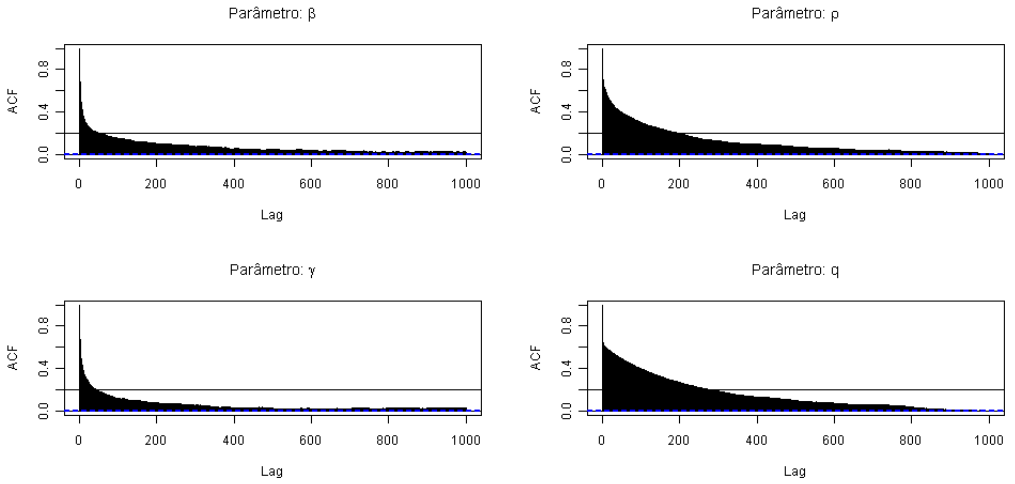


Figura 76: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

A.3.3 Distribuição a priori não centrada

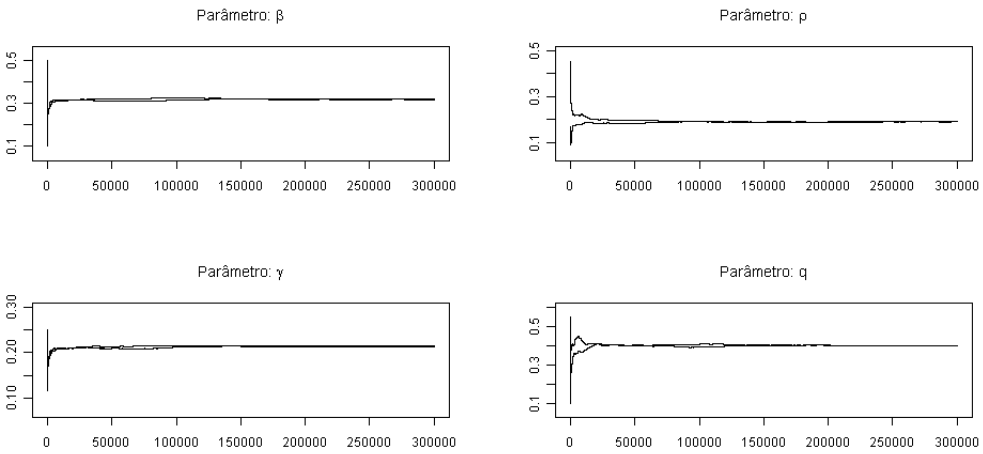


Figura 77: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não centrada com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

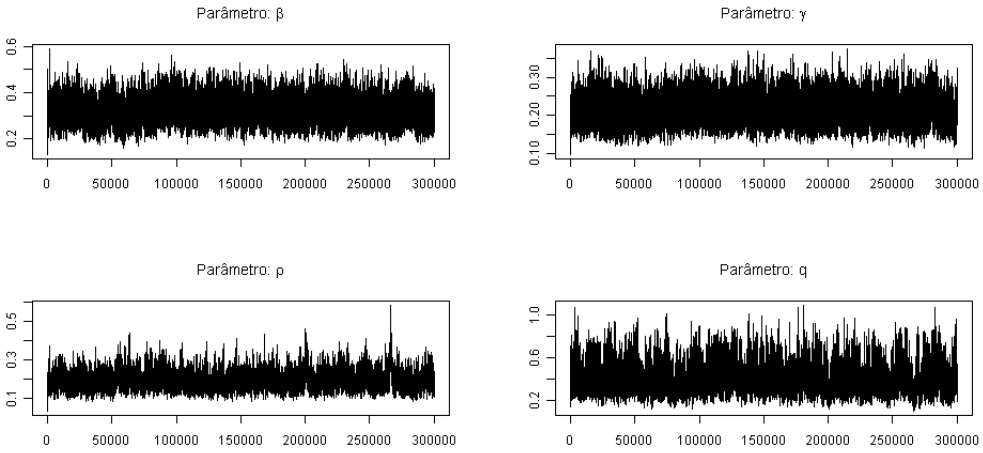


Figura 78: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não centrada com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

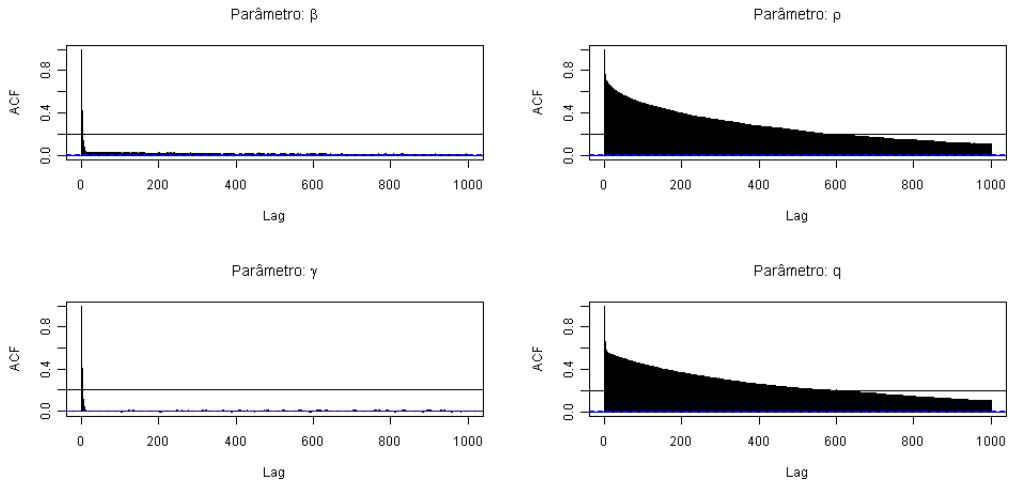


Figura 79: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não centrada com **B**, **C** e **D** incompletos na análise de sensibilidade do modelo com taxa de remoção homogênea

APÊNDICE B – Gráficos do Capítulo 4

Este apêndice contém as figuras rerente ao Capítulo 4.

B.1 Dados completos

B.1.1 Análise de sensibilidade

B.1.1.1 Distribuição a priori não informativa

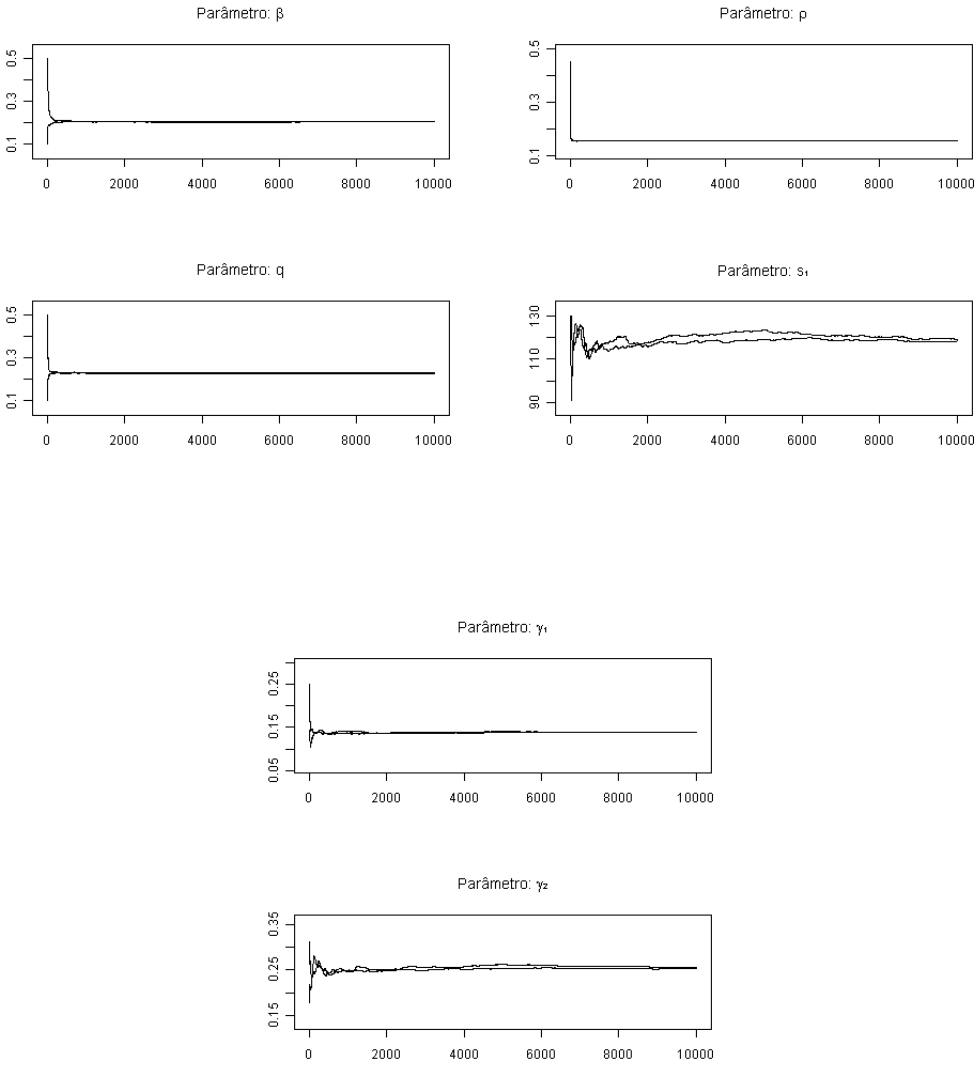


Figura 80: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

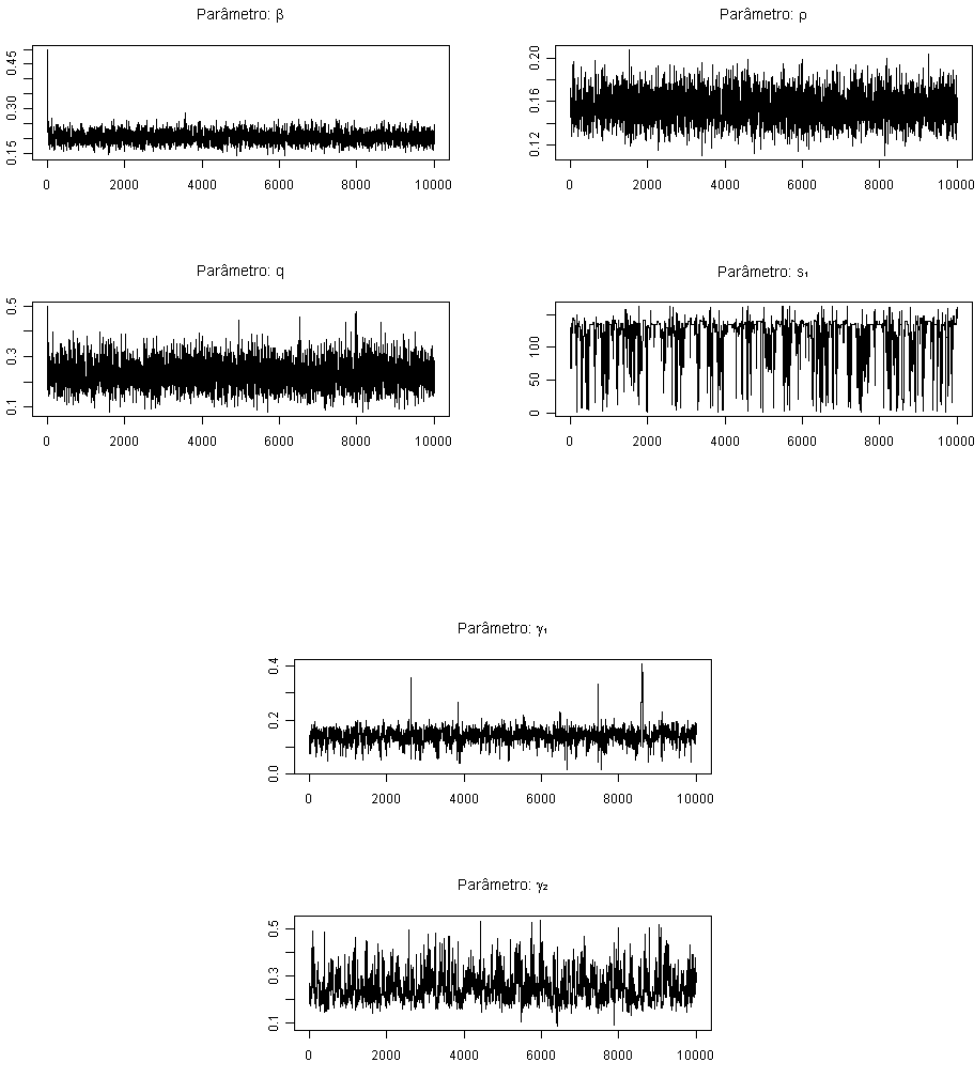


Figura 81: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

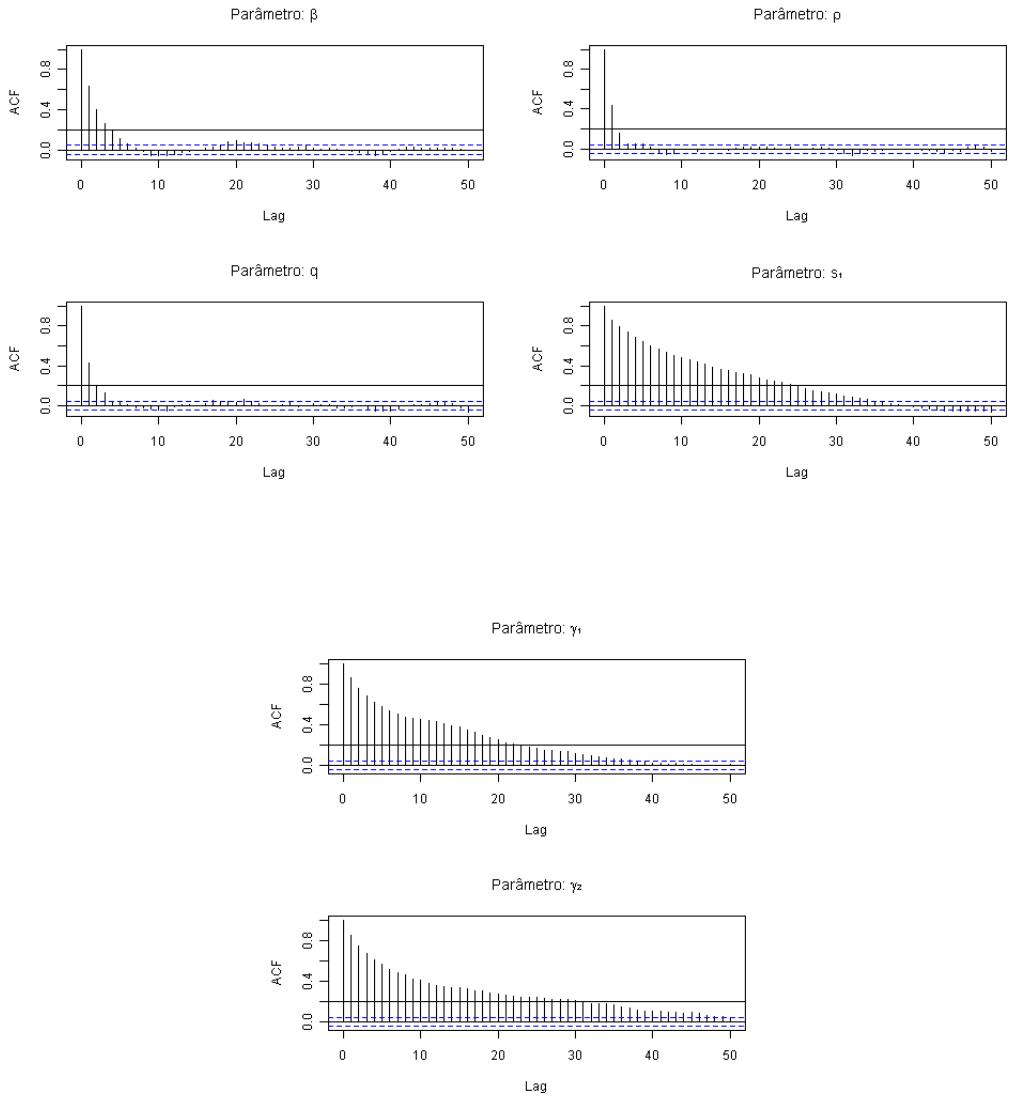


Figura 82: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

B.1.1.2 Distribuição a priori informativa

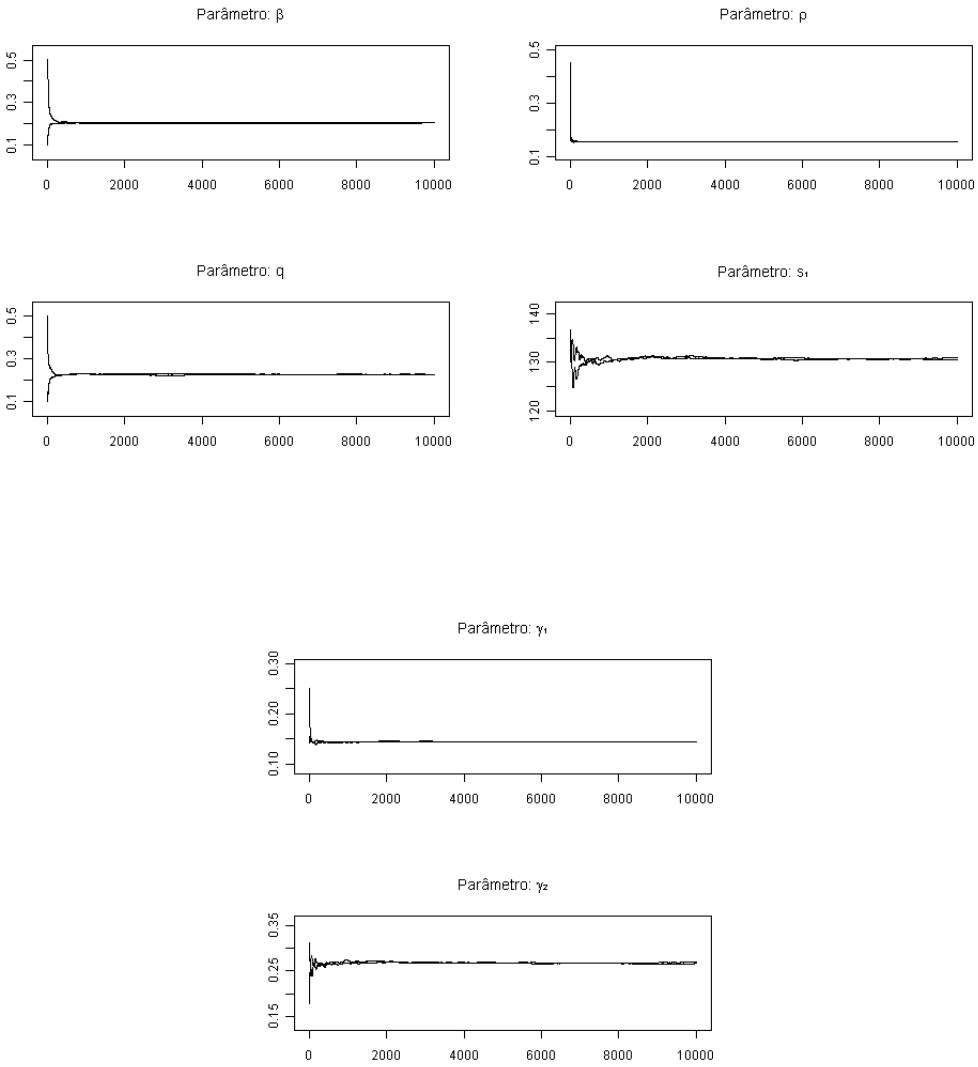


Figura 83: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

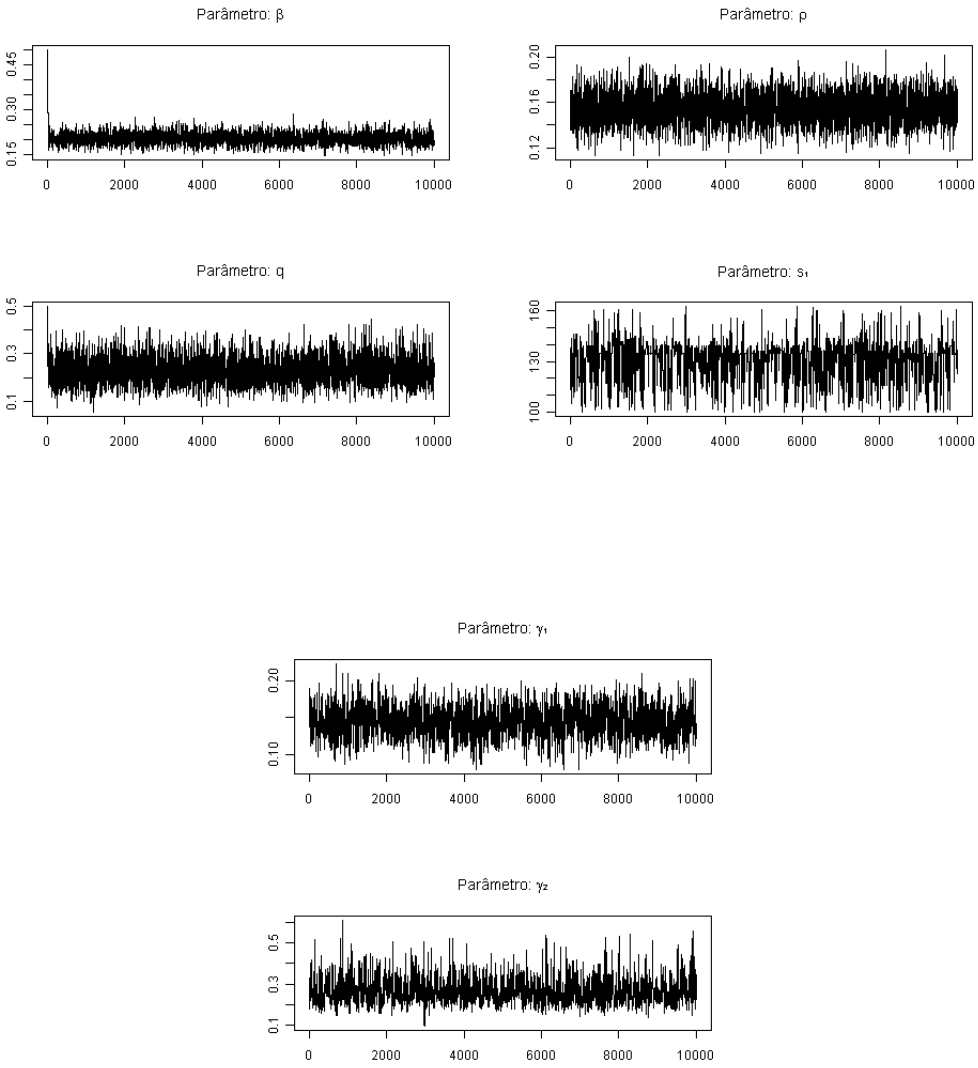


Figura 84: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

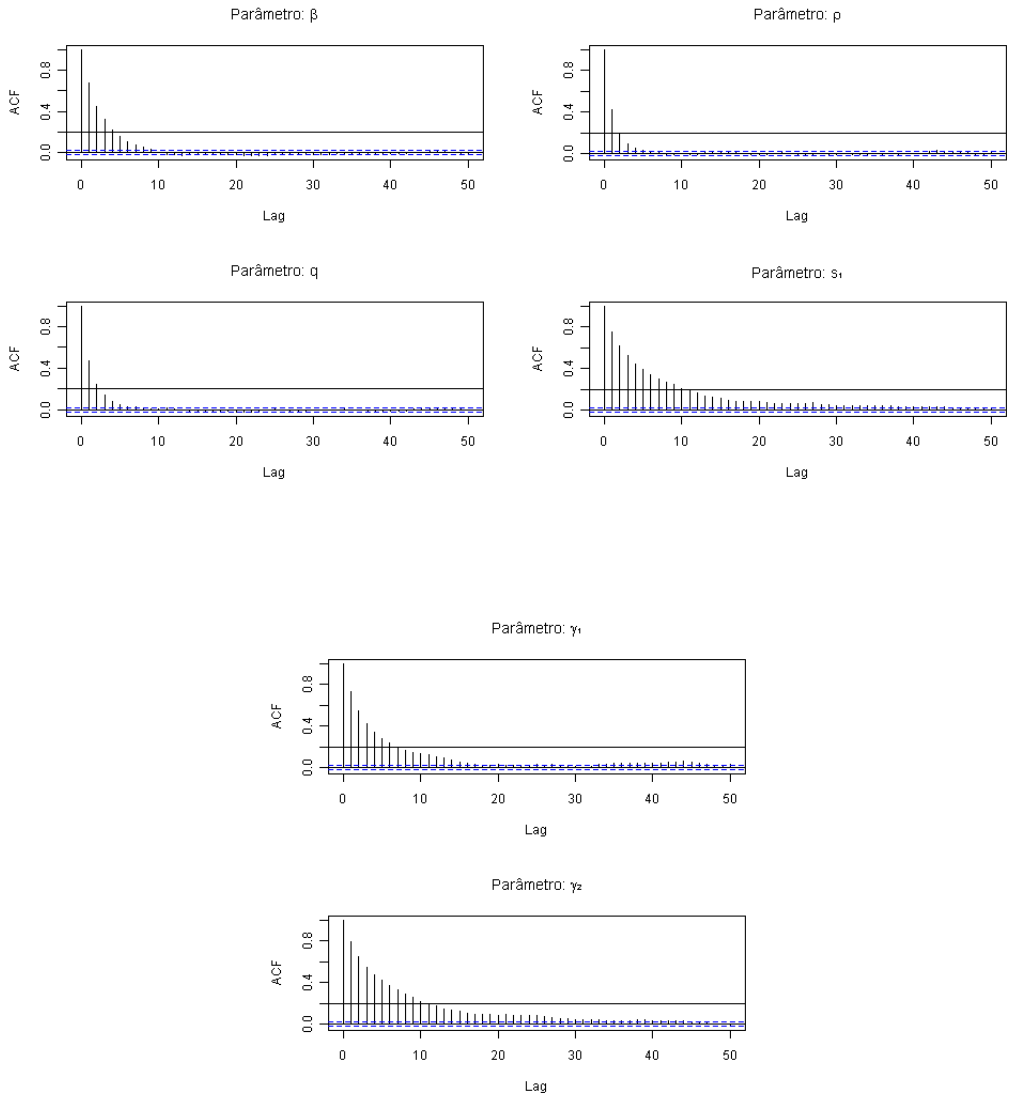


Figura 85: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori informativa com dados completos na análise de sensibilidade do modelo com taxa de remoção não homogênea

B.1.2 Erro de especificação

B.1.2.1 Modelo com taxa de remoção homogênea

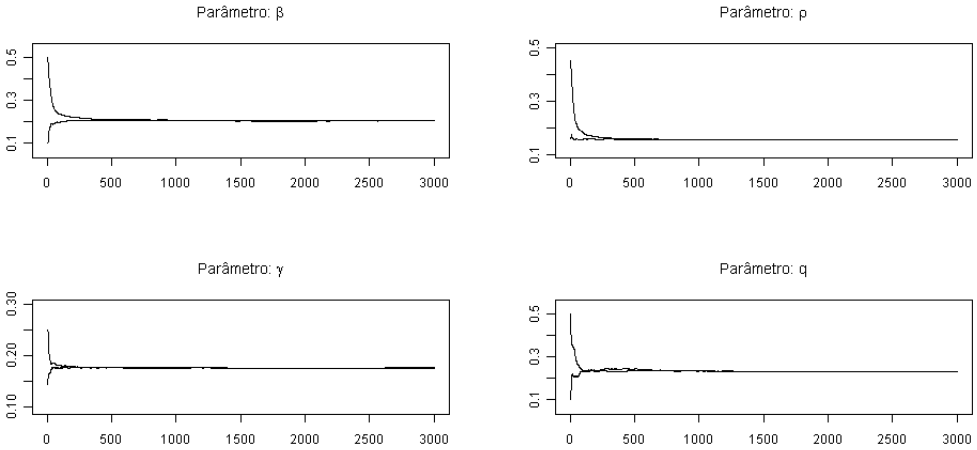


Figura 86: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção homogênea

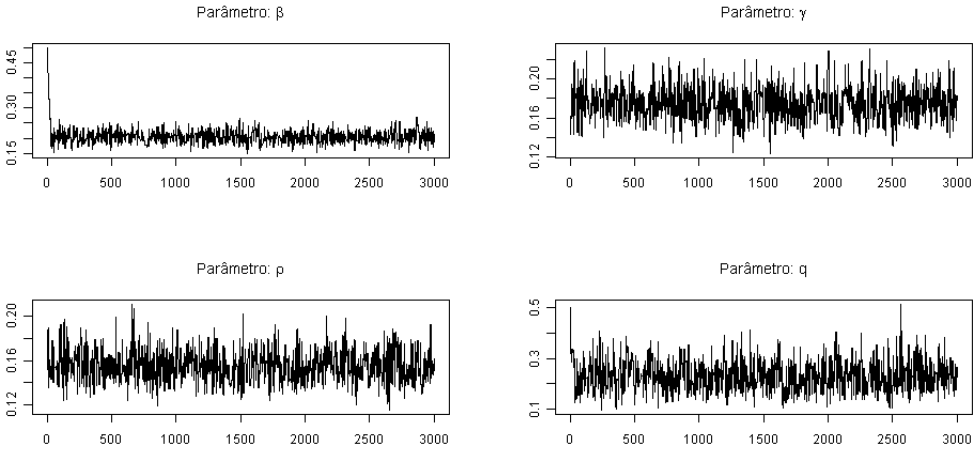


Figura 87: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção homogênea

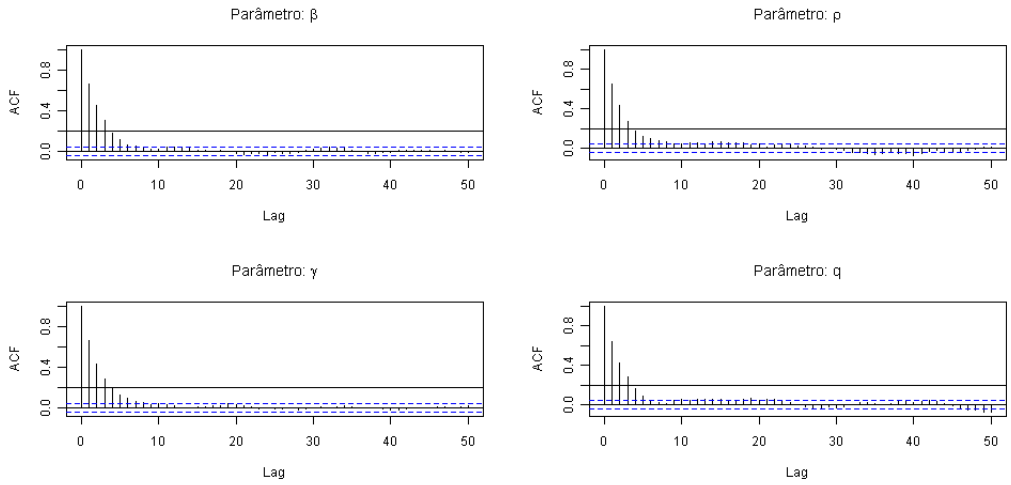


Figura 88: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção homogênea

B.1.2.2 Modelo com taxa de remoção não homogênea considerando distribuição a priori não informativa

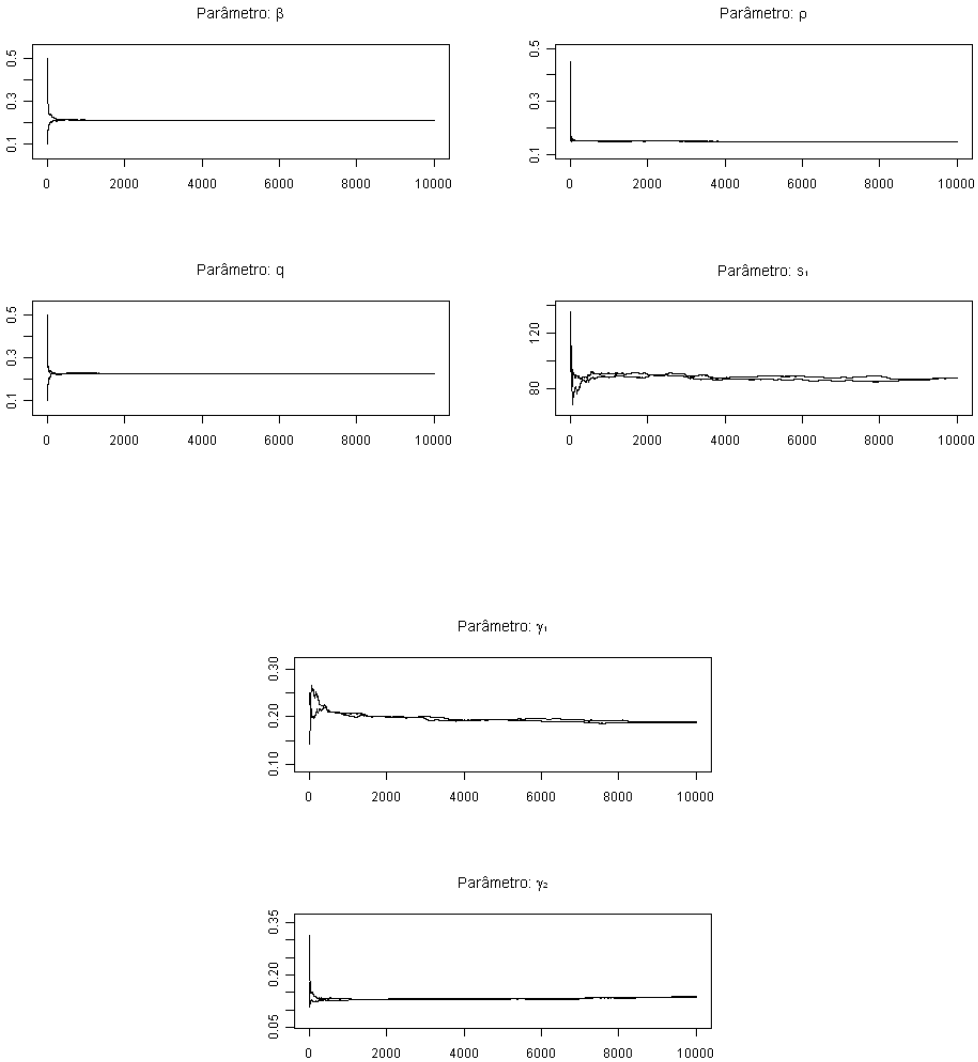


Figura 89: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

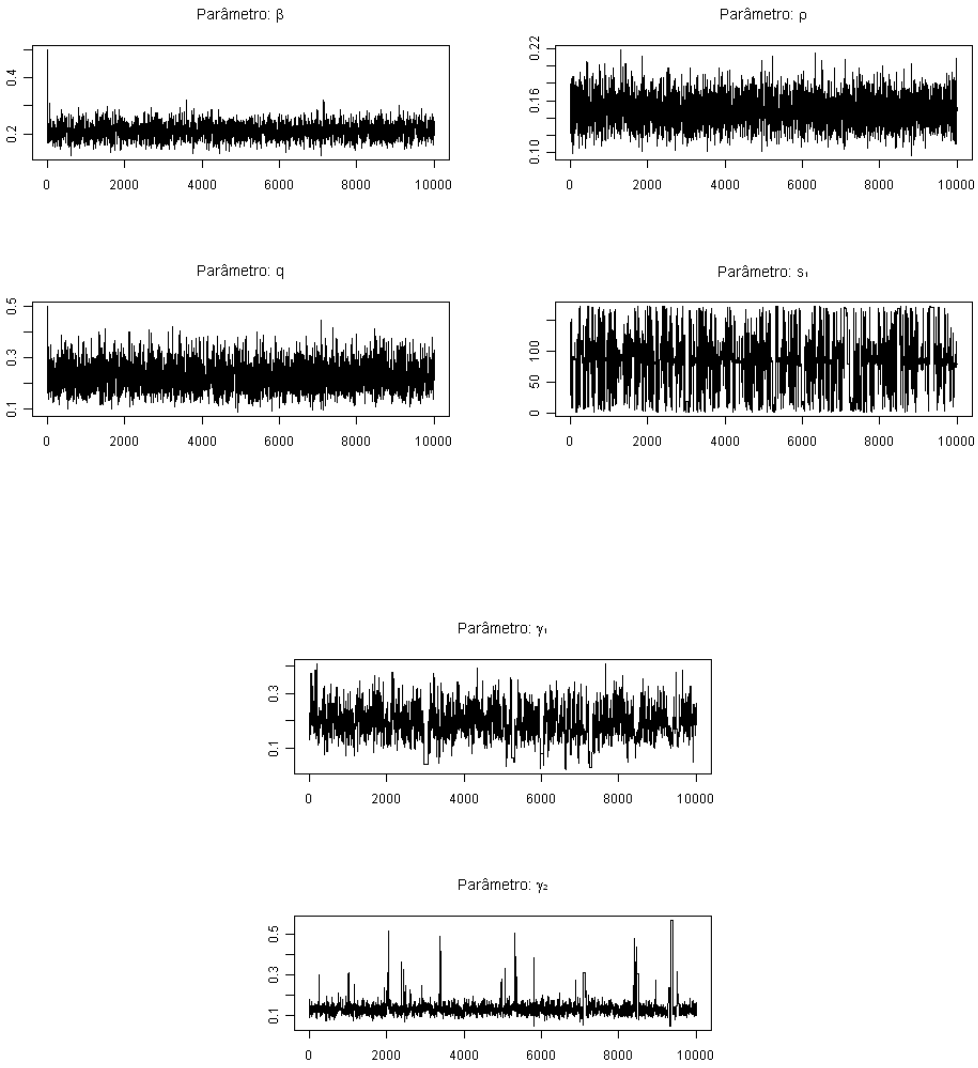


Figura 90: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

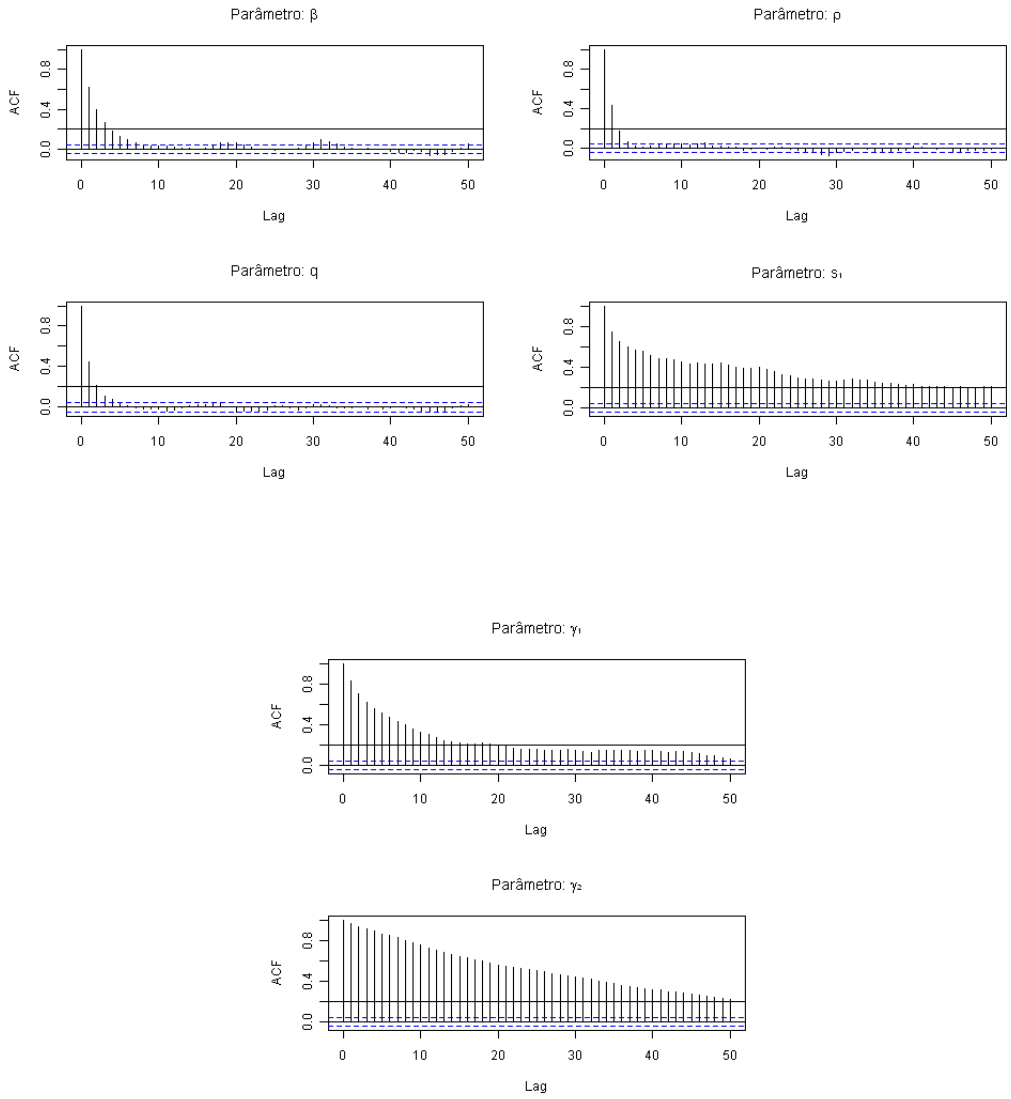


Figura 91: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

B.1.2.3 Modelo com taxa de remoção não homogênea considerando distribuição a priori informativa

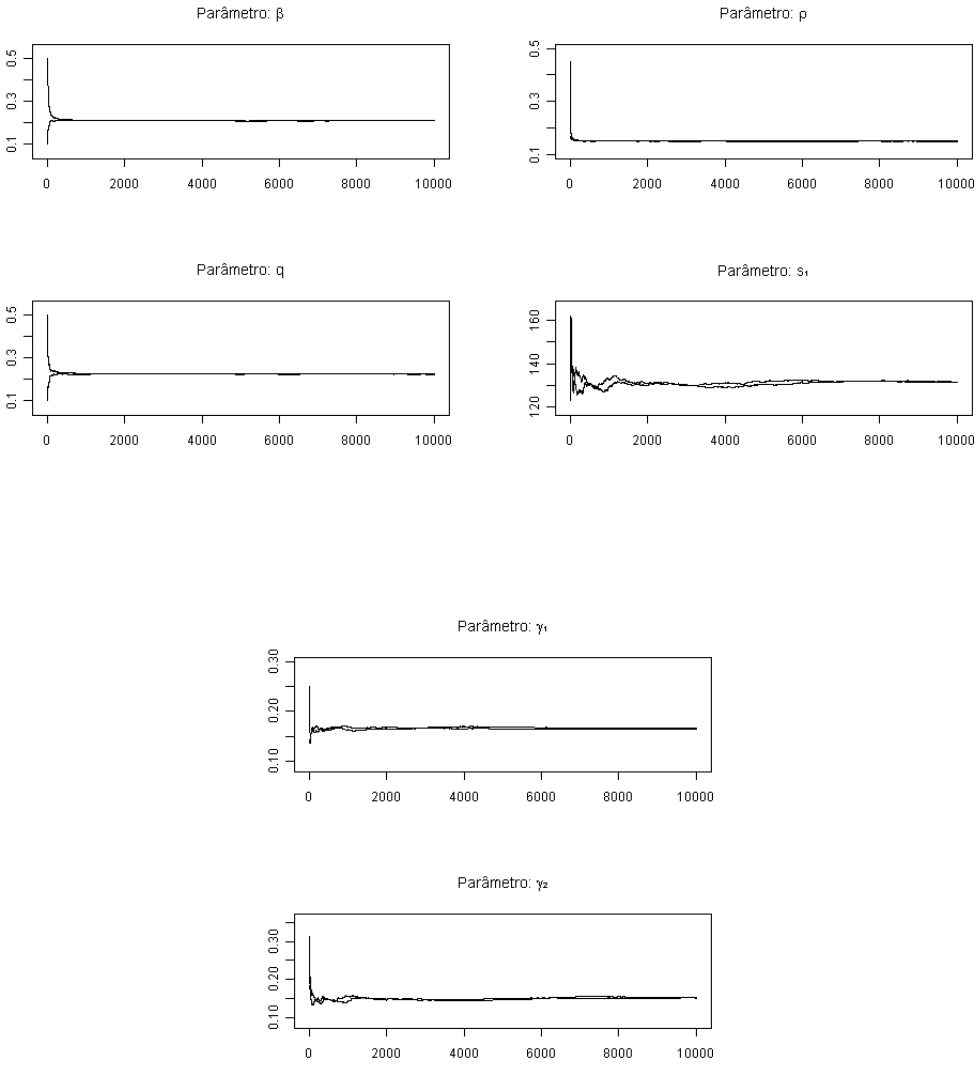


Figura 92: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

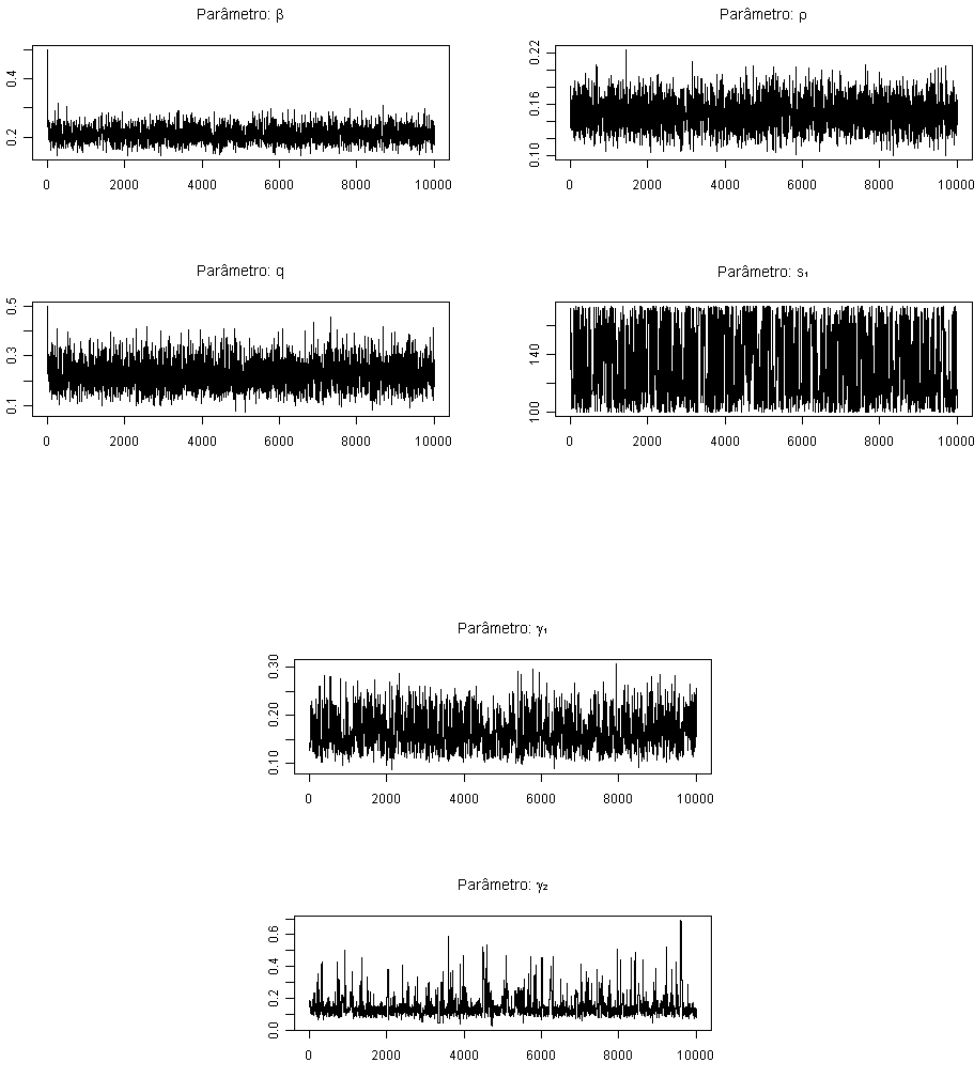


Figura 93: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

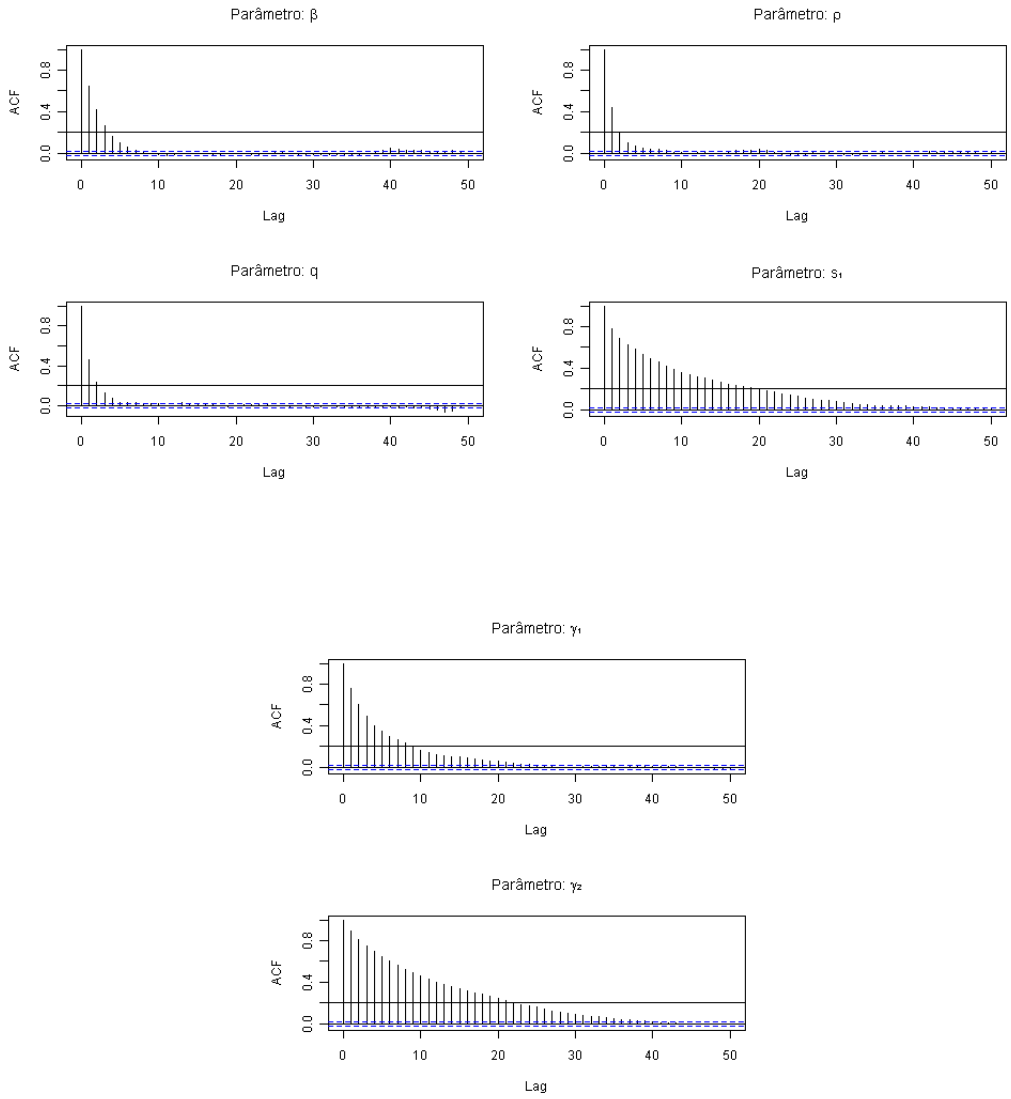


Figura 94: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori informativa com dados completos na análise de erro de especificação do modelo com taxa de remoção não homogênea

B.2 B, C, D incompletos

B.2.1 Análise de sensibilidade

B.2.1.1 Distribuição a priori não informativa

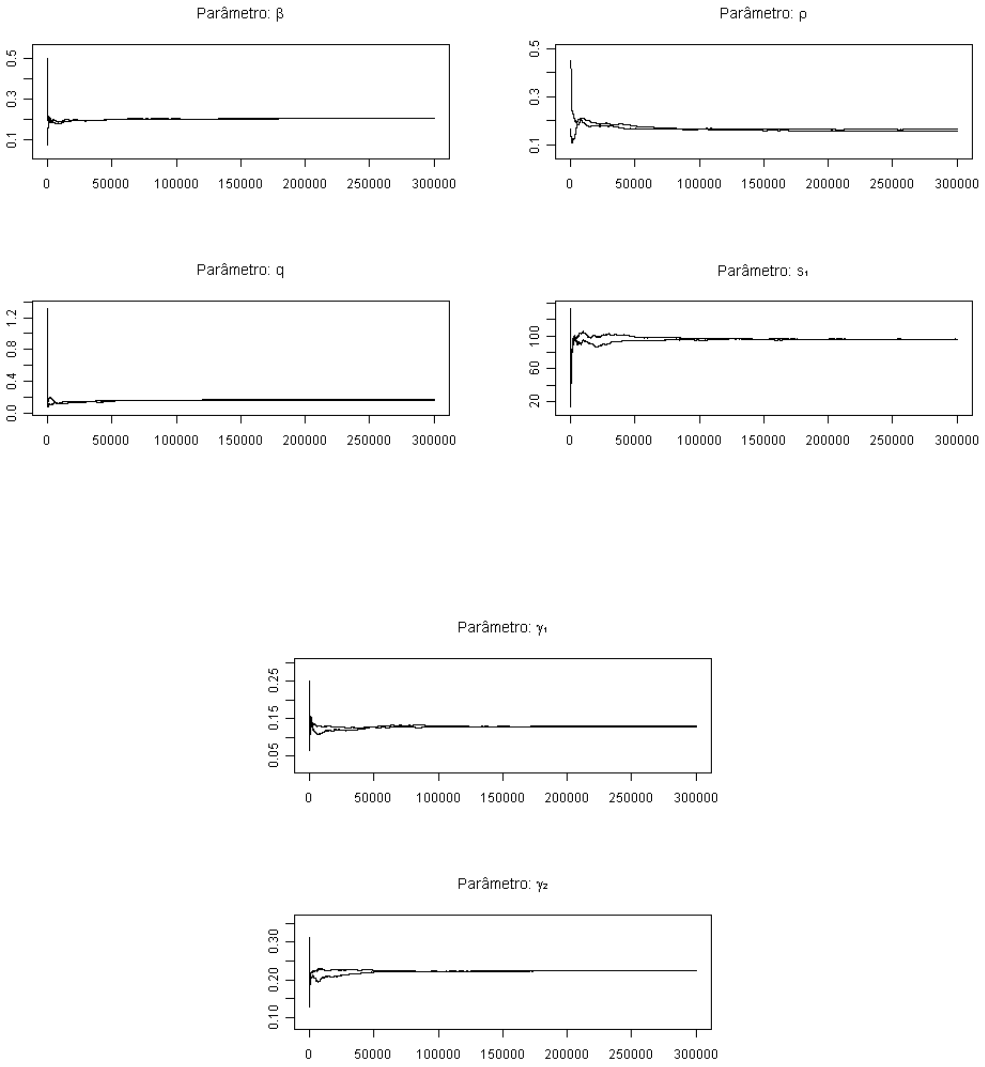


Figura 95: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

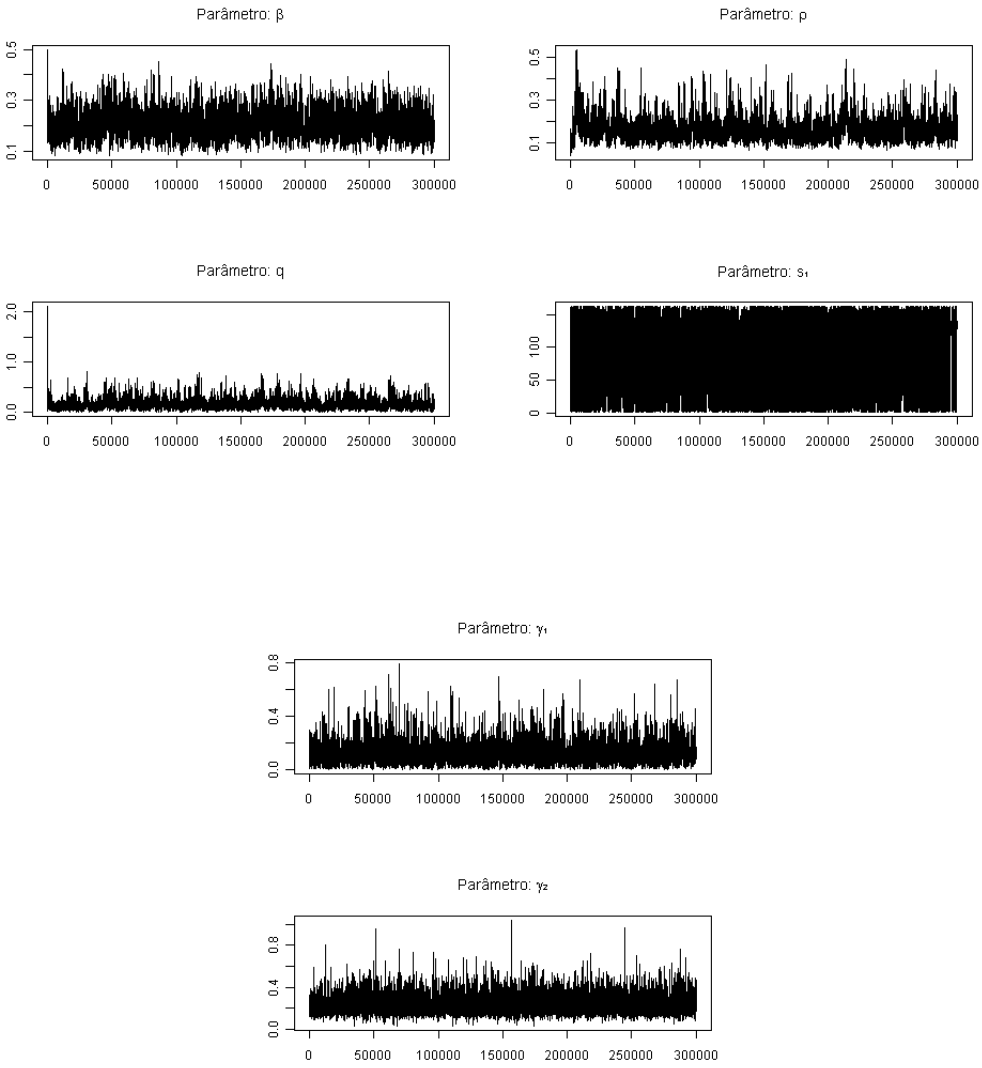


Figura 96: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

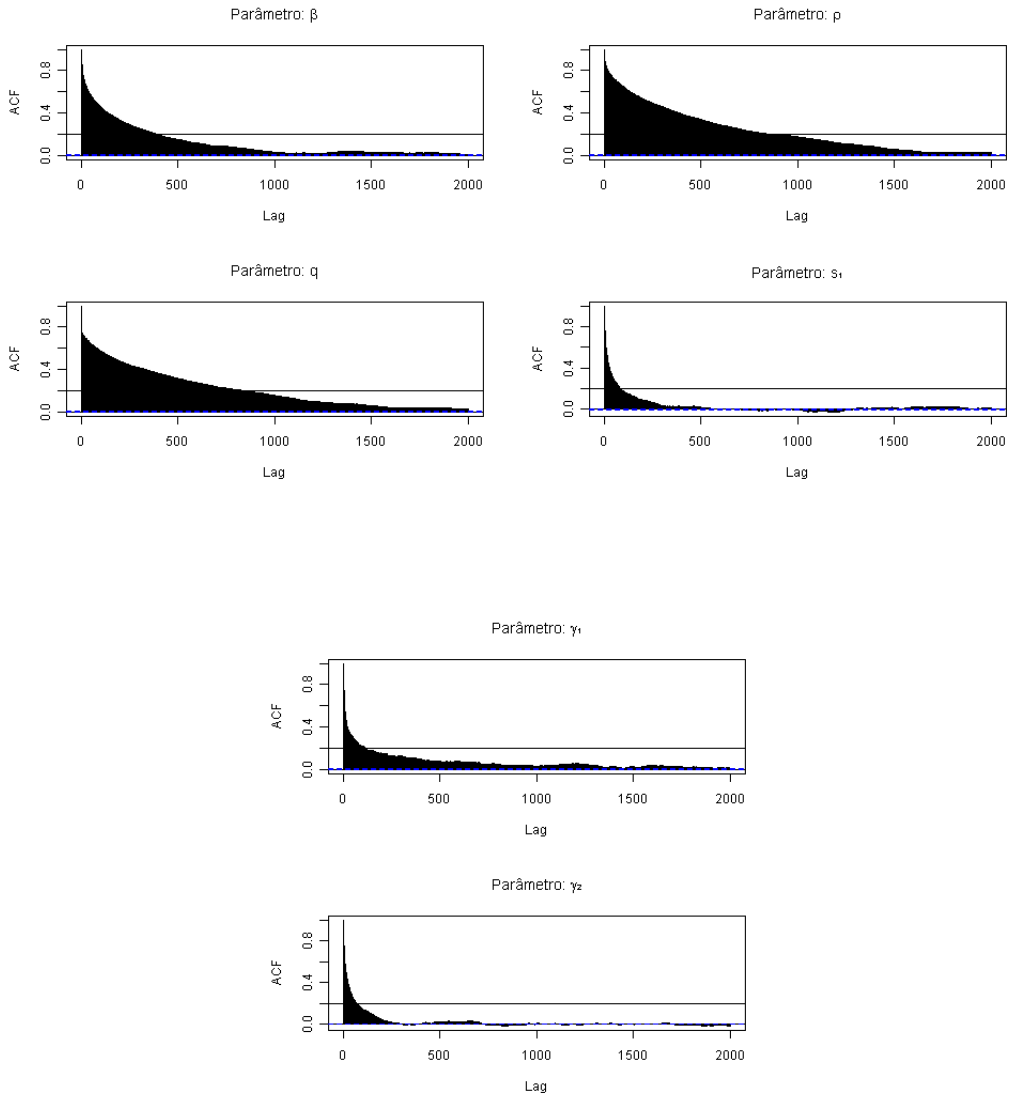


Figura 97: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

B.2.1.2 Distribuição a priori informativa

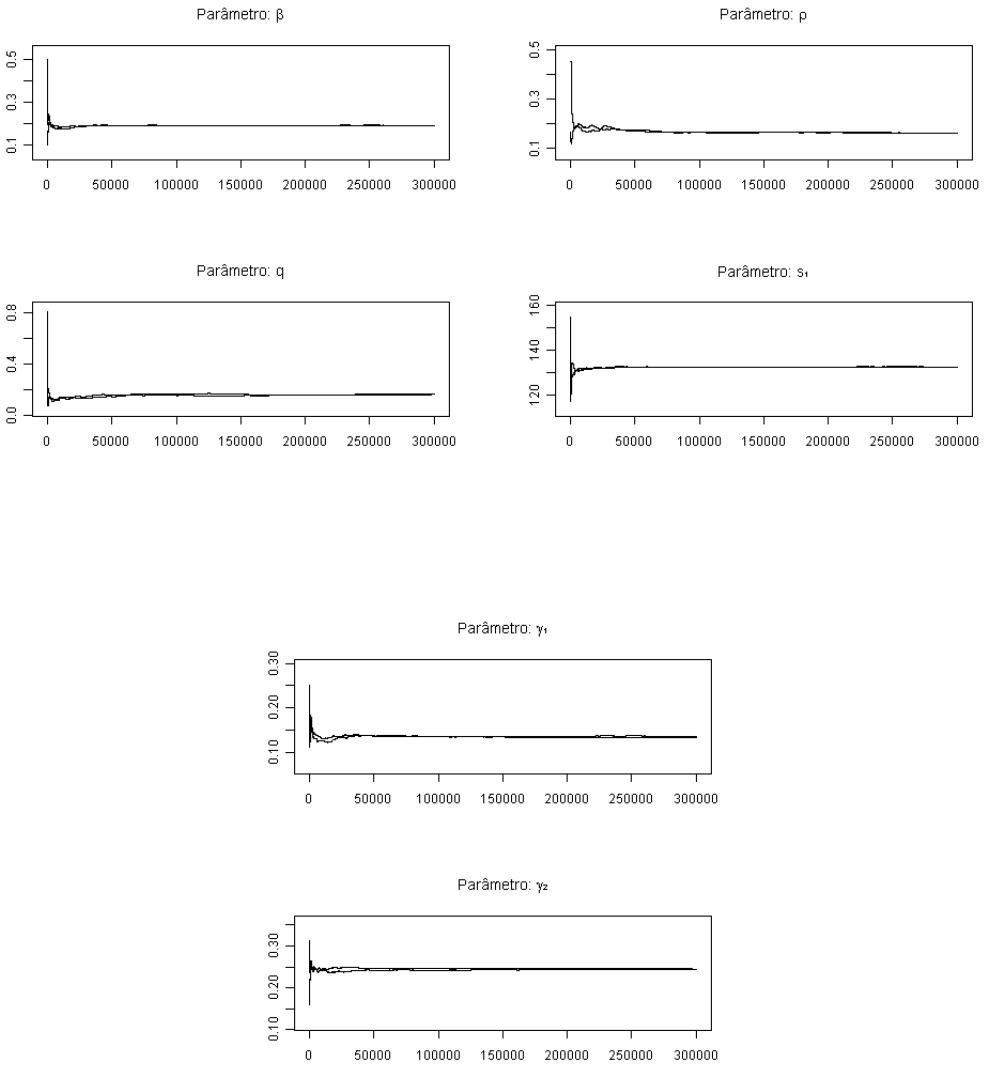


Figura 98: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

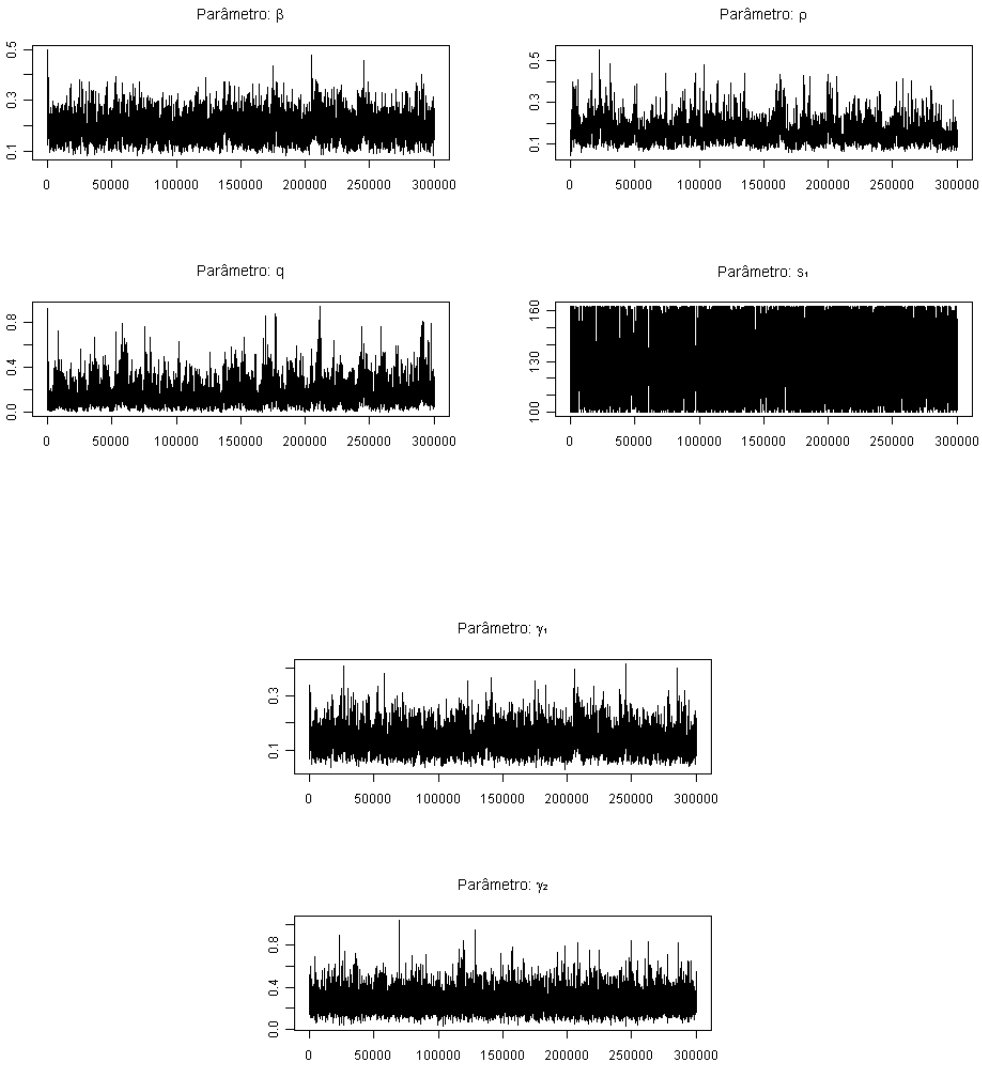


Figura 99: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

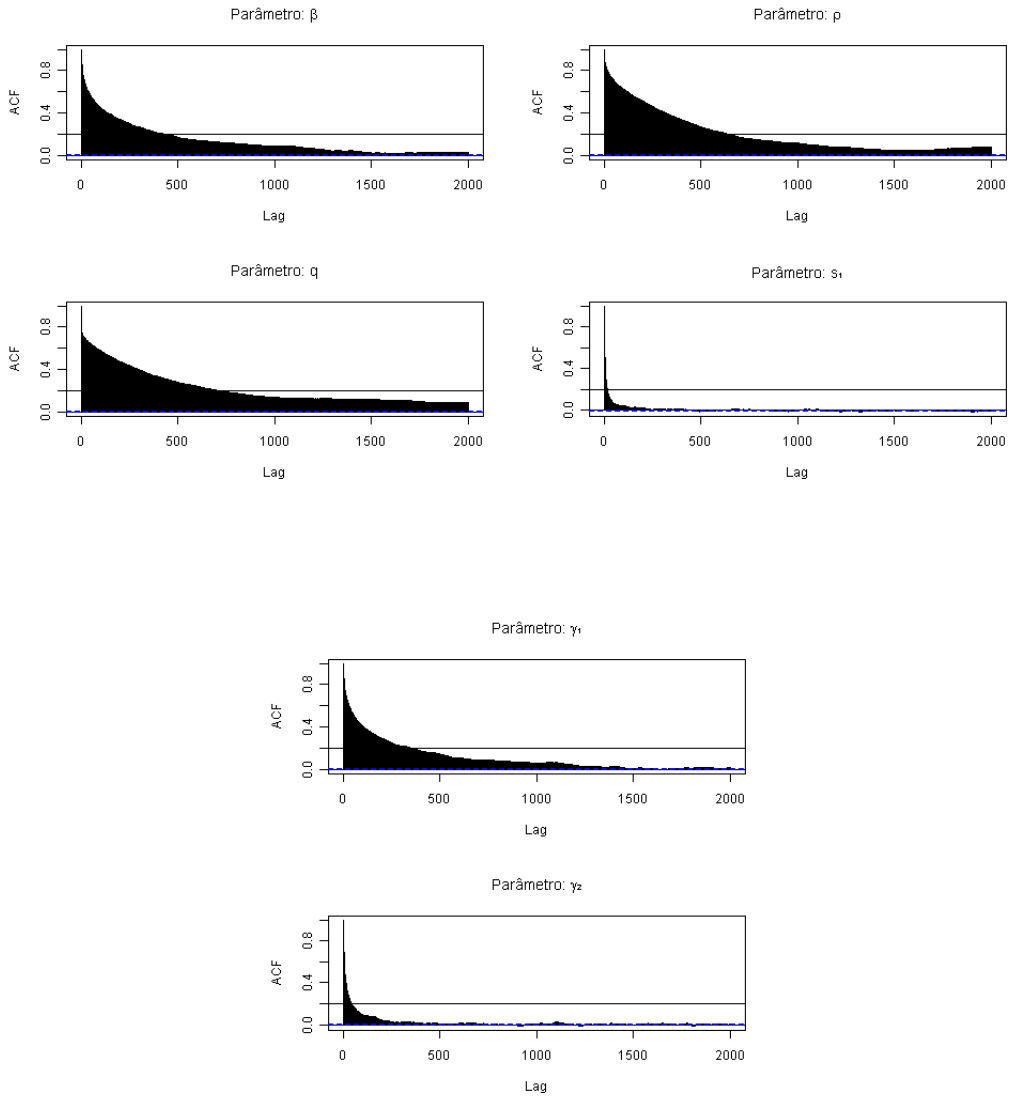


Figura 100: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori informativa com dados **B**, **C**, **D** incompletos na análise de sensibilidade do modelo com taxa de remoção não homogênea

B.2.2 Erro de especificação

B.2.2.1 Modelo com taxa de remoção homogênea

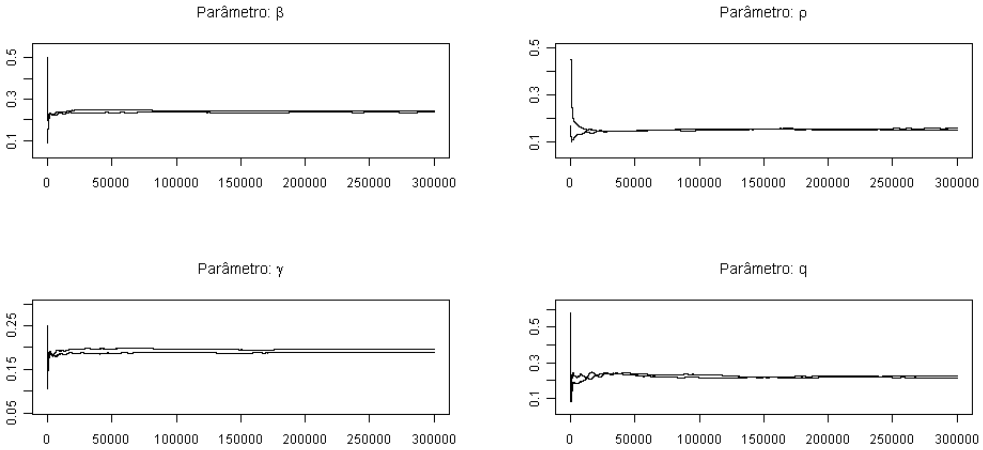


Figura 101: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção homogênea

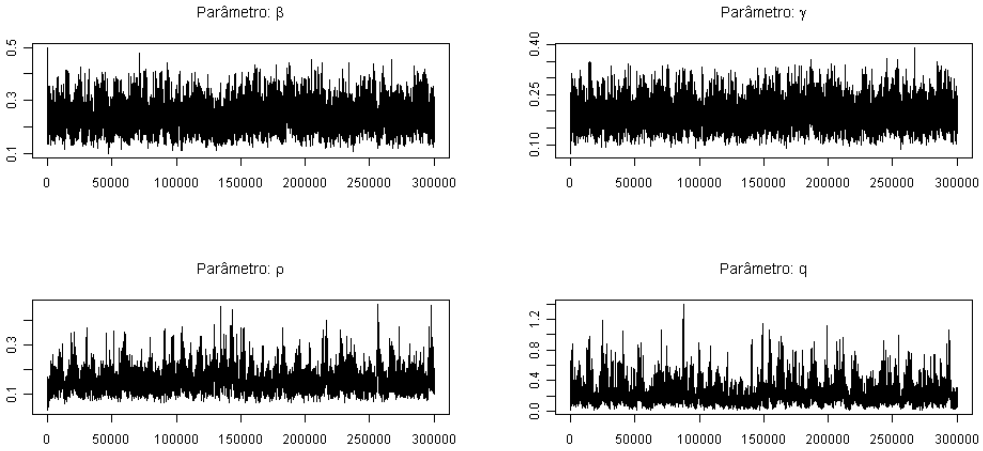


Figura 102: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção homogênea

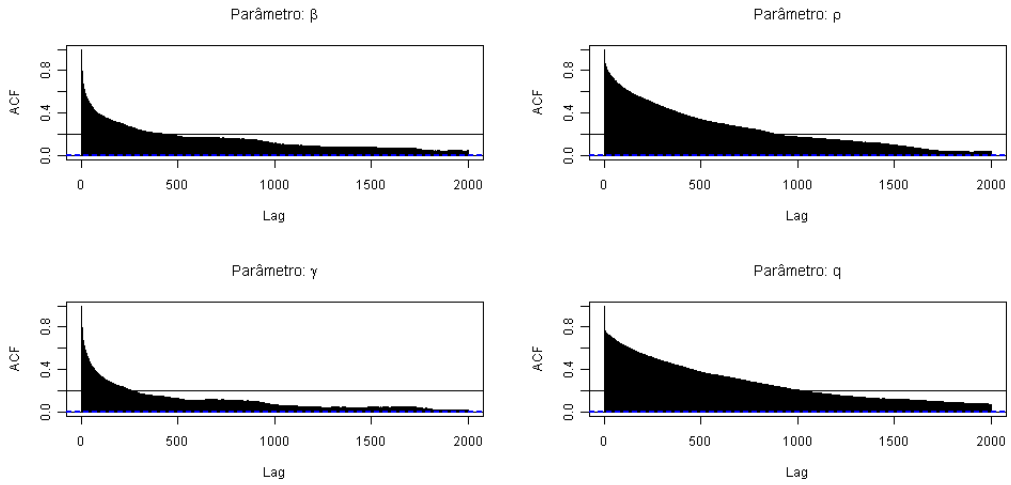


Figura 103: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção homogênea

B.2.2.2 Modelo com taxa de remoção não homogênea considerando distribuição a priori não informativa

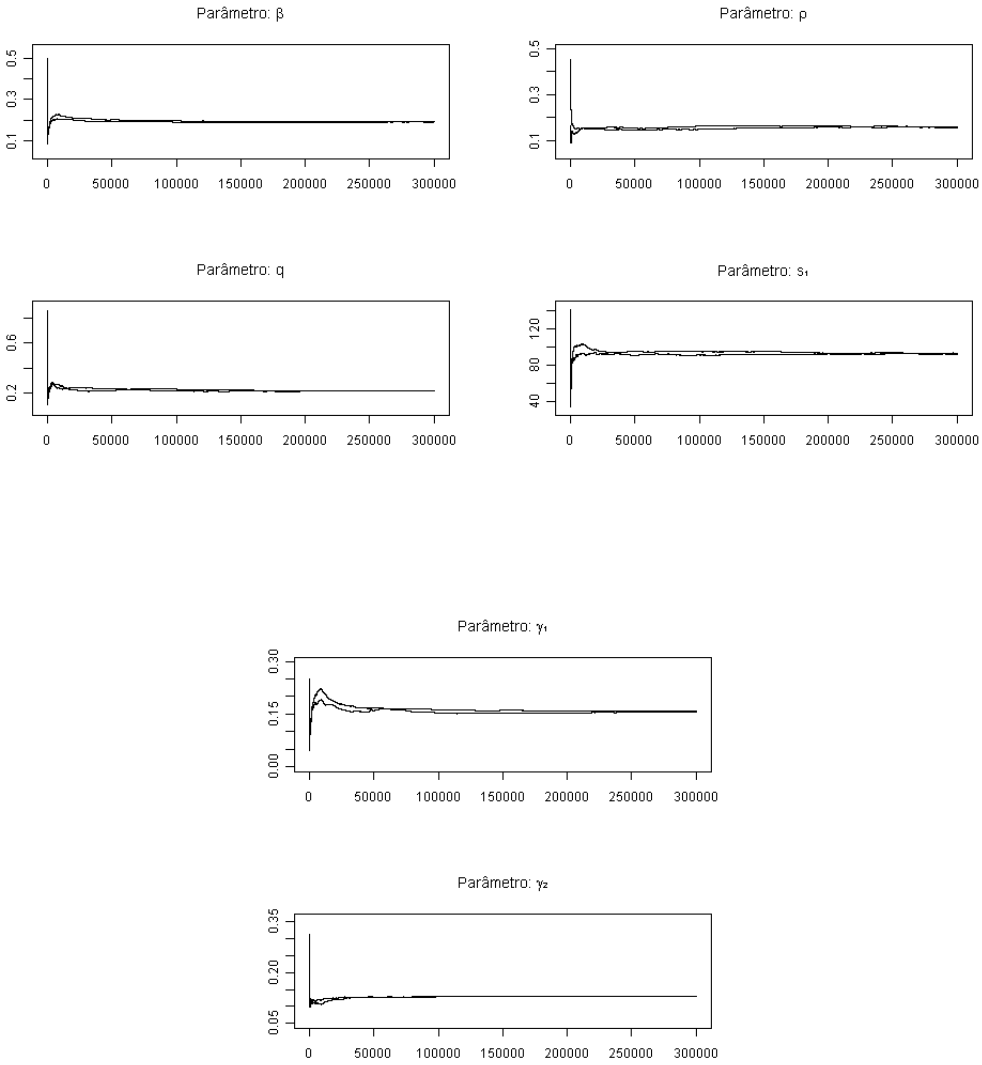


Figura 104: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

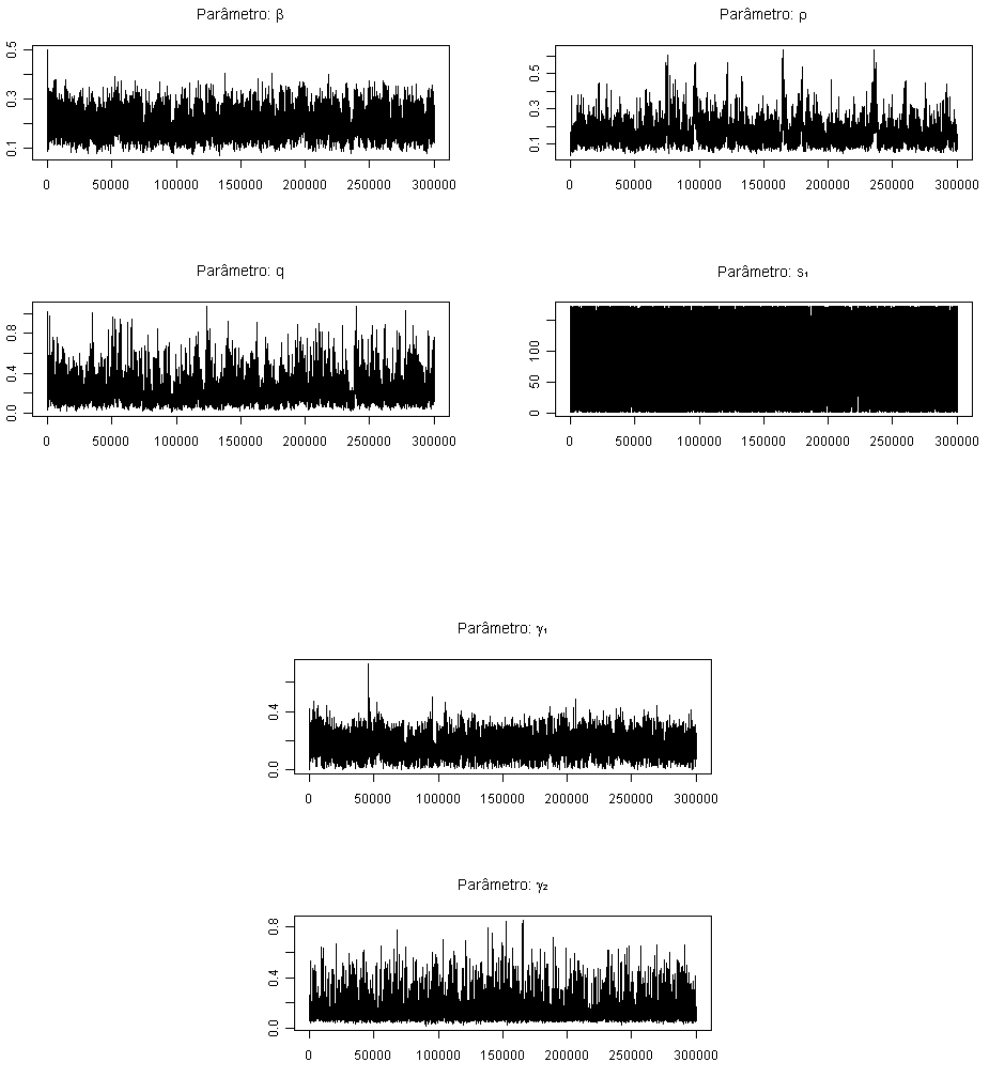


Figura 105: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

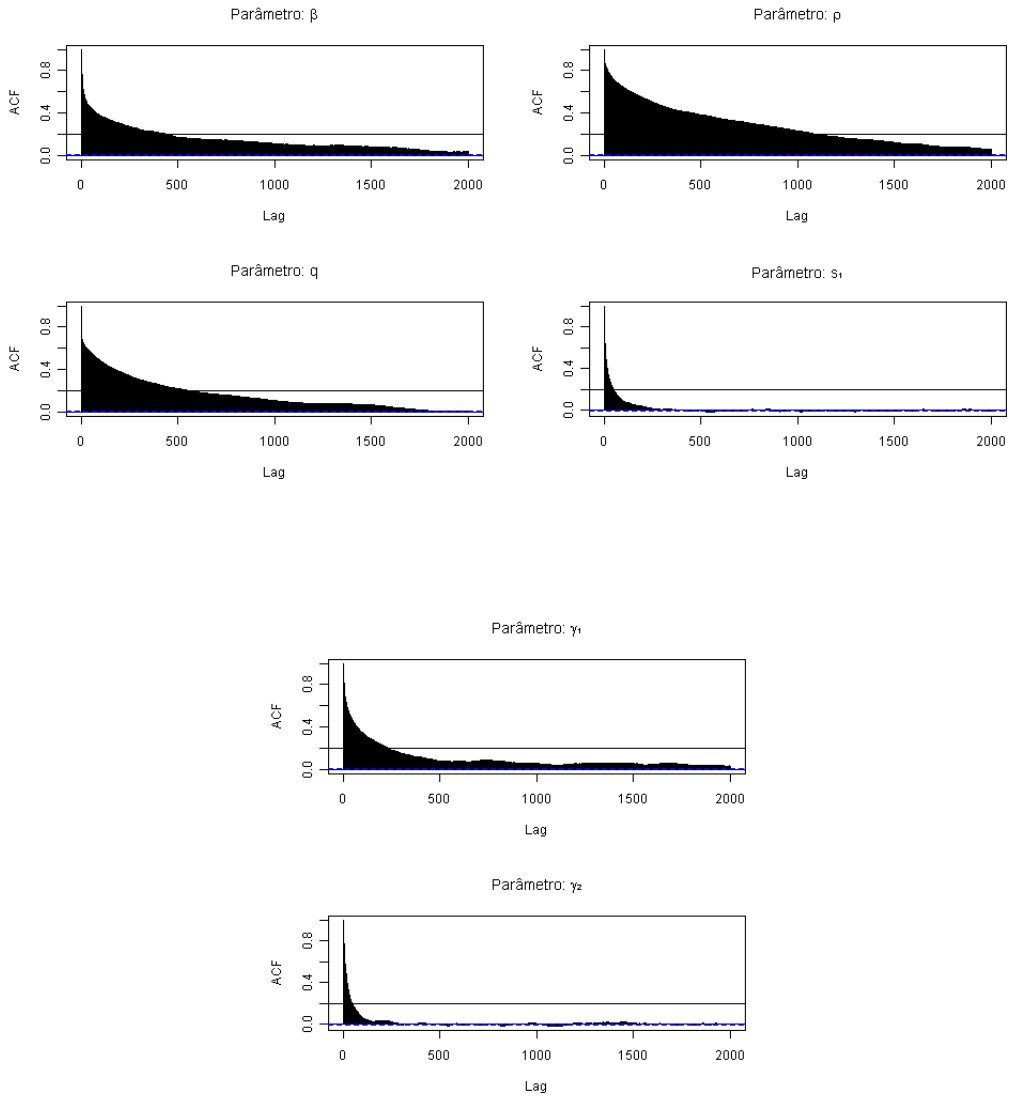


Figura 106: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

B.2.2.3 Modelo com taxa de remoção não homogênea considerando distribuição a priori informativa

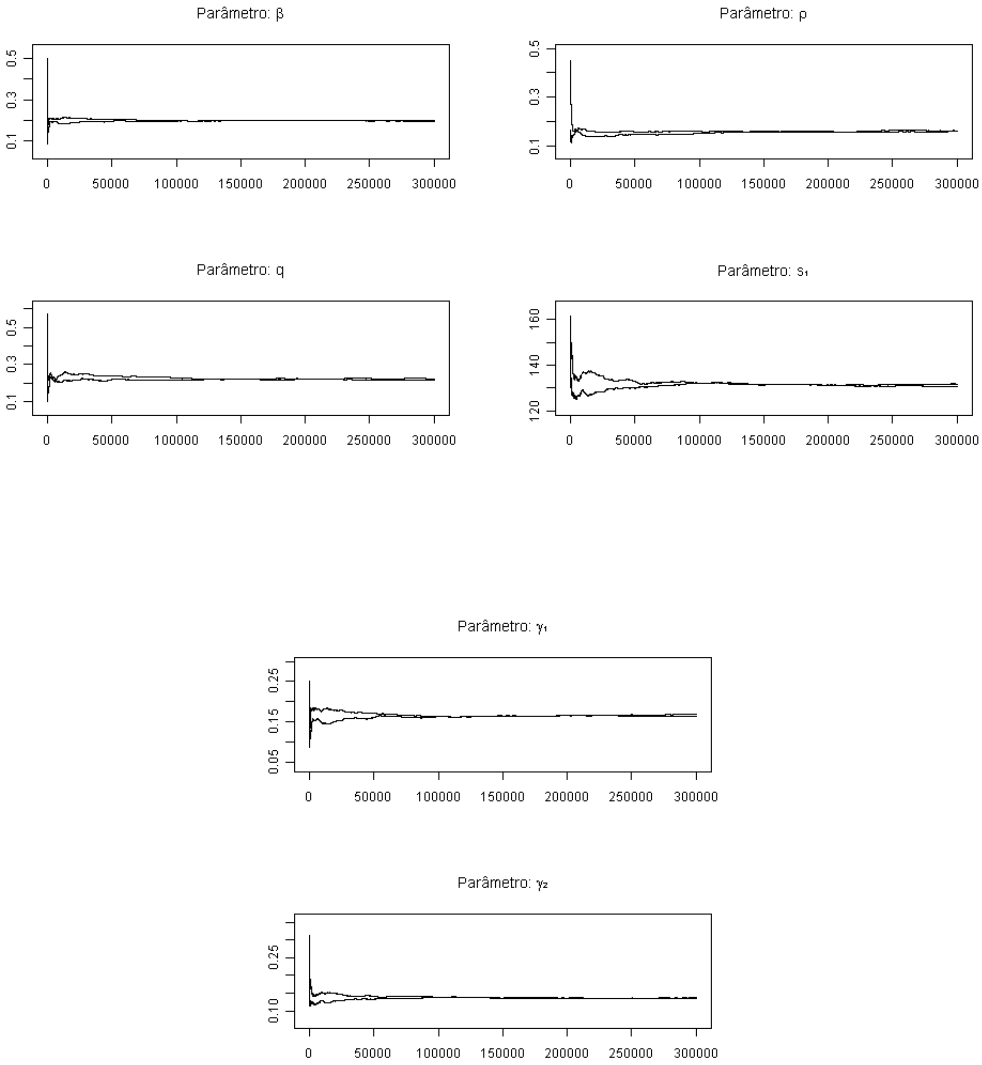


Figura 107: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

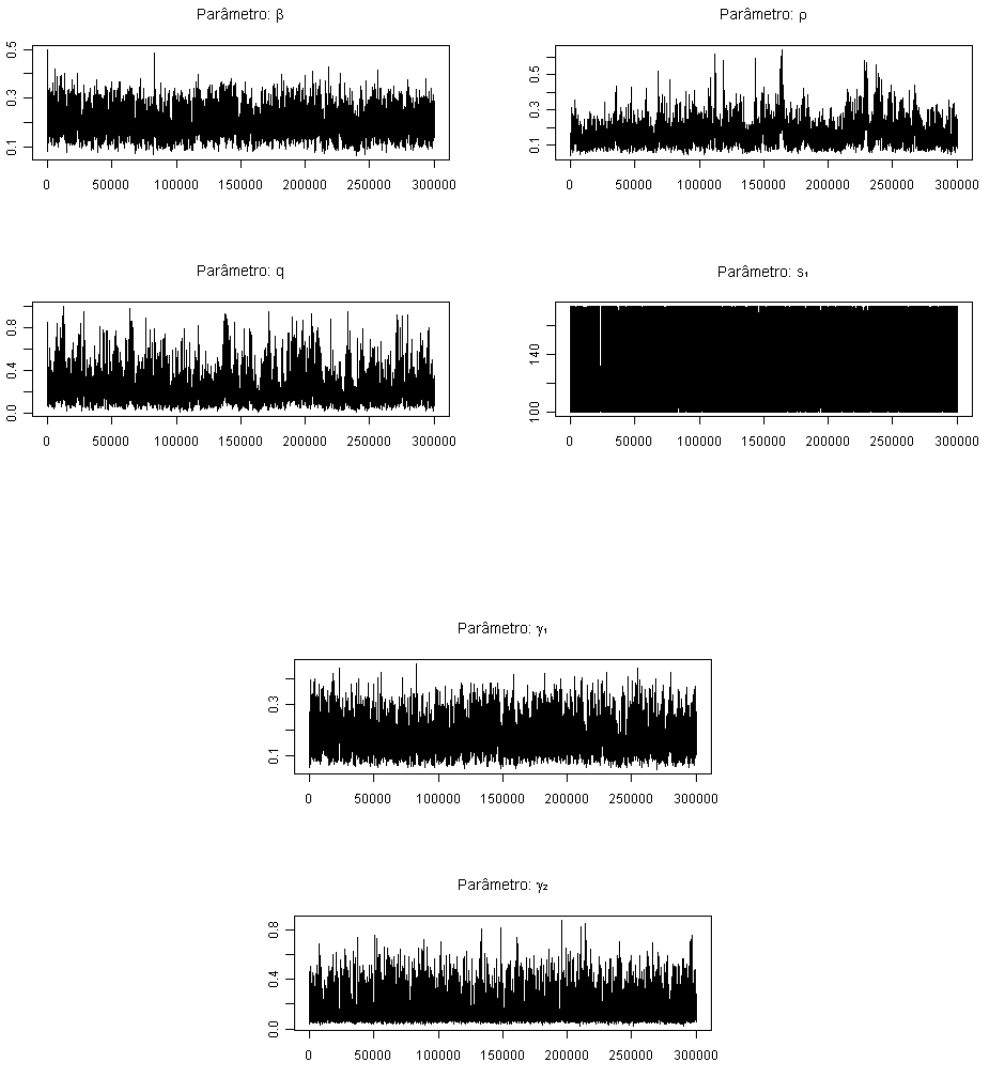


Figura 108: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

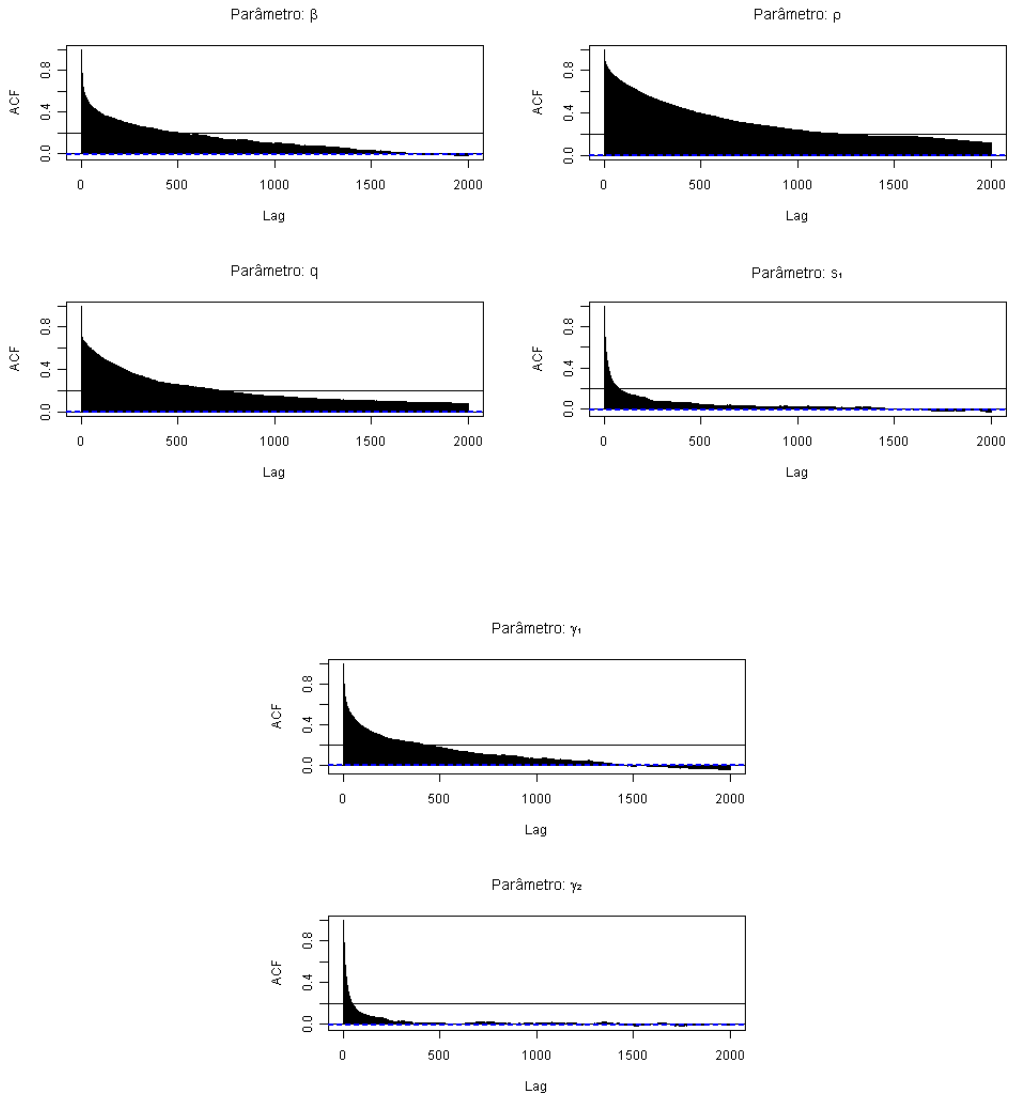


Figura 109: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori informativa com **B**, **C** e **D** incompletos na análise de erro de especificação do modelo com taxa de remoção não homogênea

APÊNDICE C – Gráficos do Capítulo 5

Este apêndice contém as figuras rerente ao Capítulo 5.

C.1 Taxa de remoção homogênea

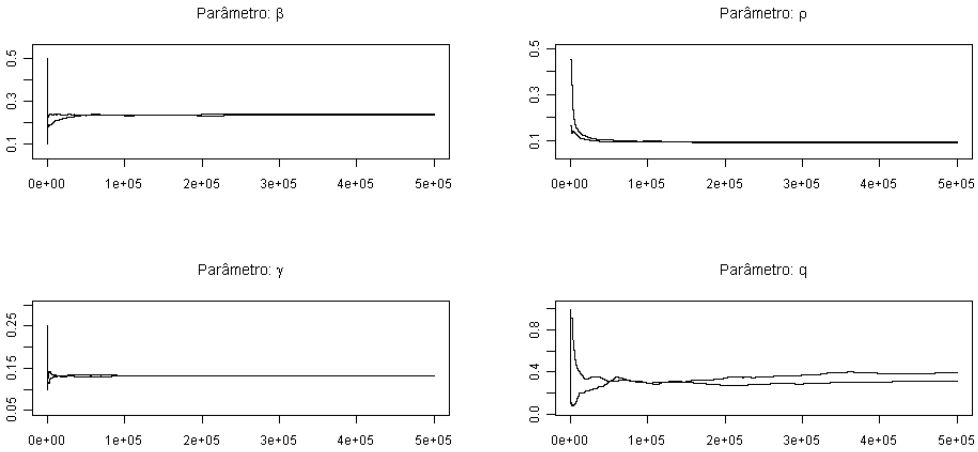


Figura 110: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção homogênea

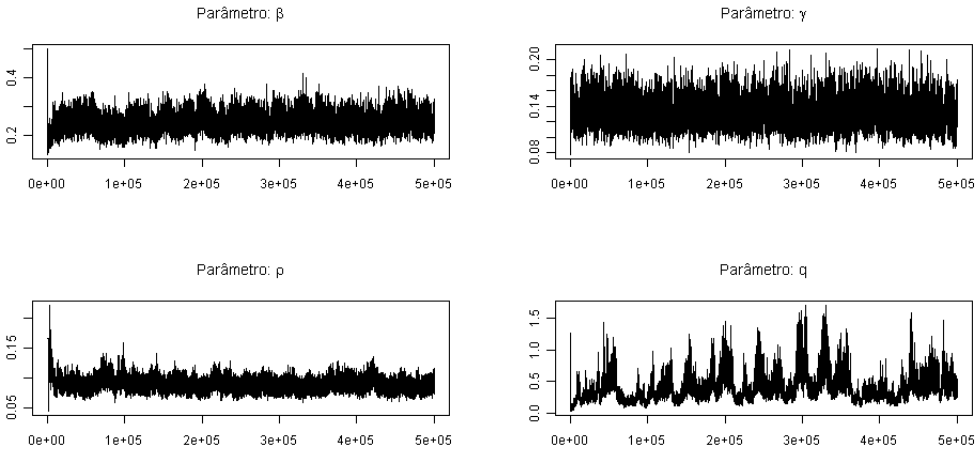


Figura 111: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção homogênea

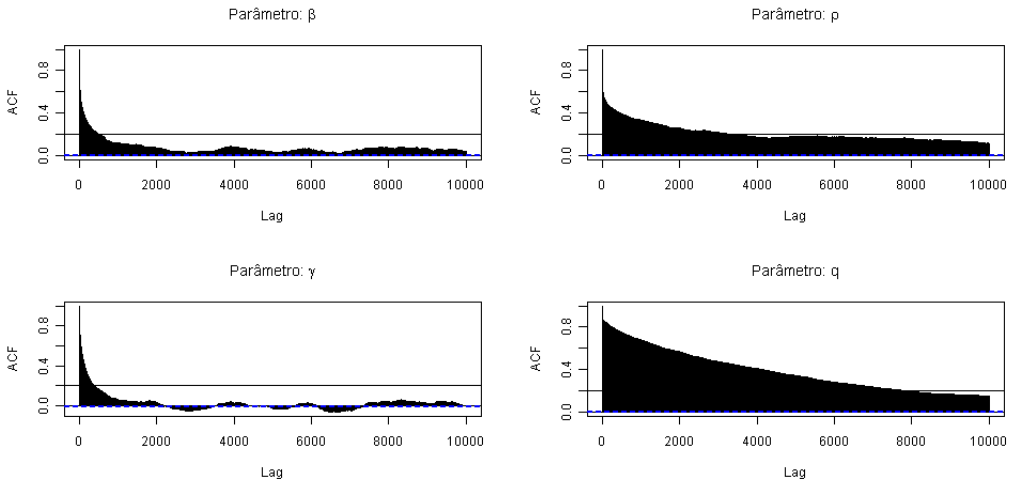


Figura 112: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa o conjunto de dados real no modelo com taxa de remoção homogênea

C.2 Taxa de remoção não homogênea

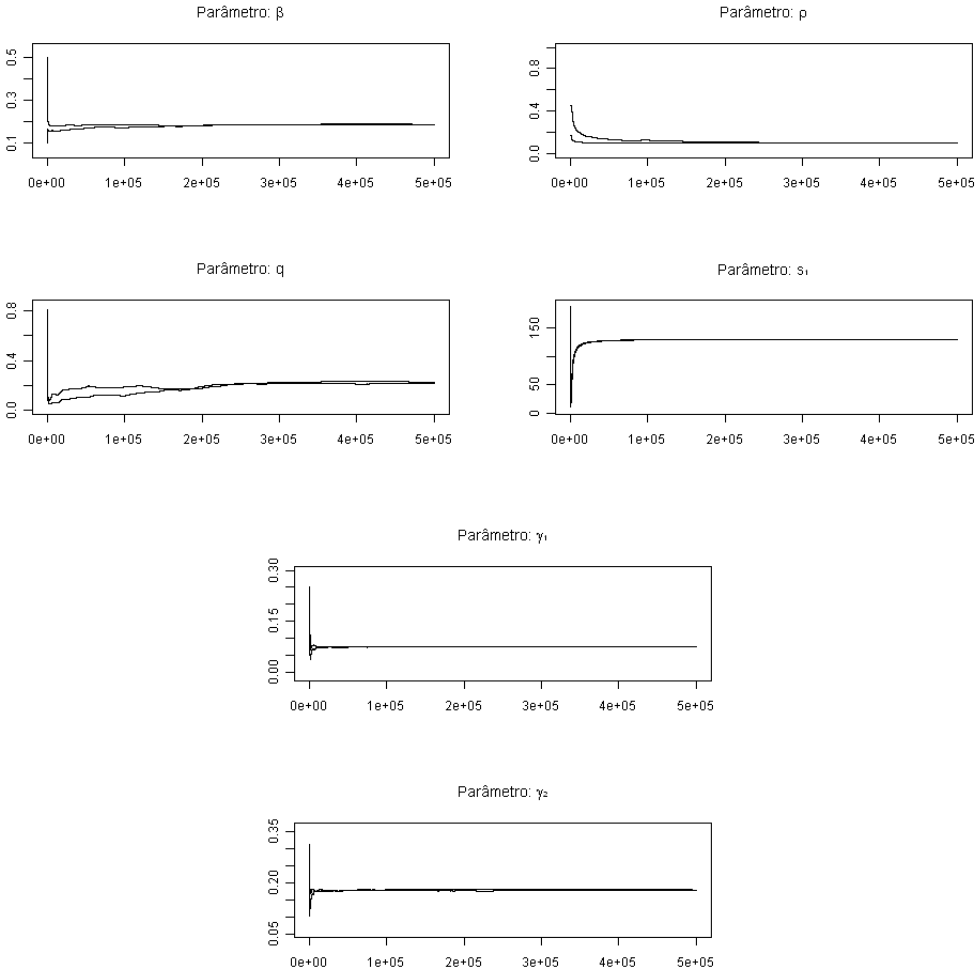


Figura 113: Médias amostrais das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção homogênea

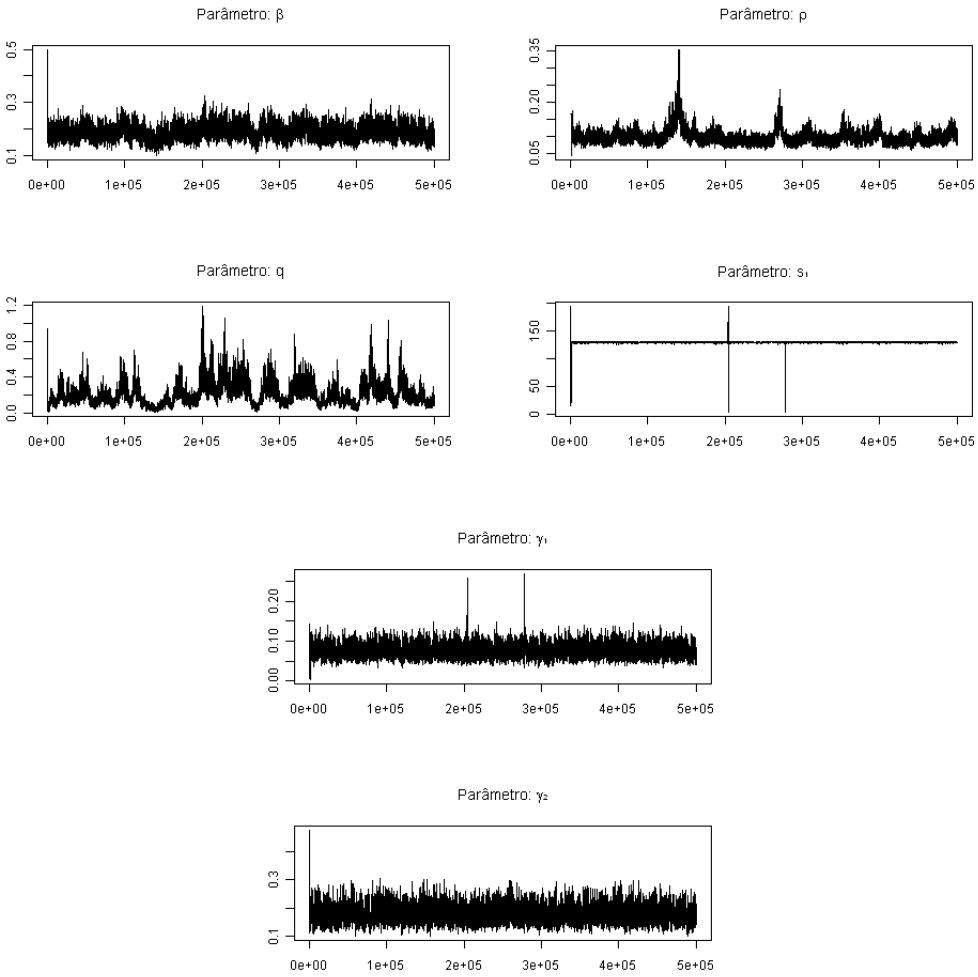


Figura 114: Históricos das amostras geradas dos parâmetros considerando a distribuição a priori não informativa com o conjunto de dados real no modelo com taxa de remoção homogênea

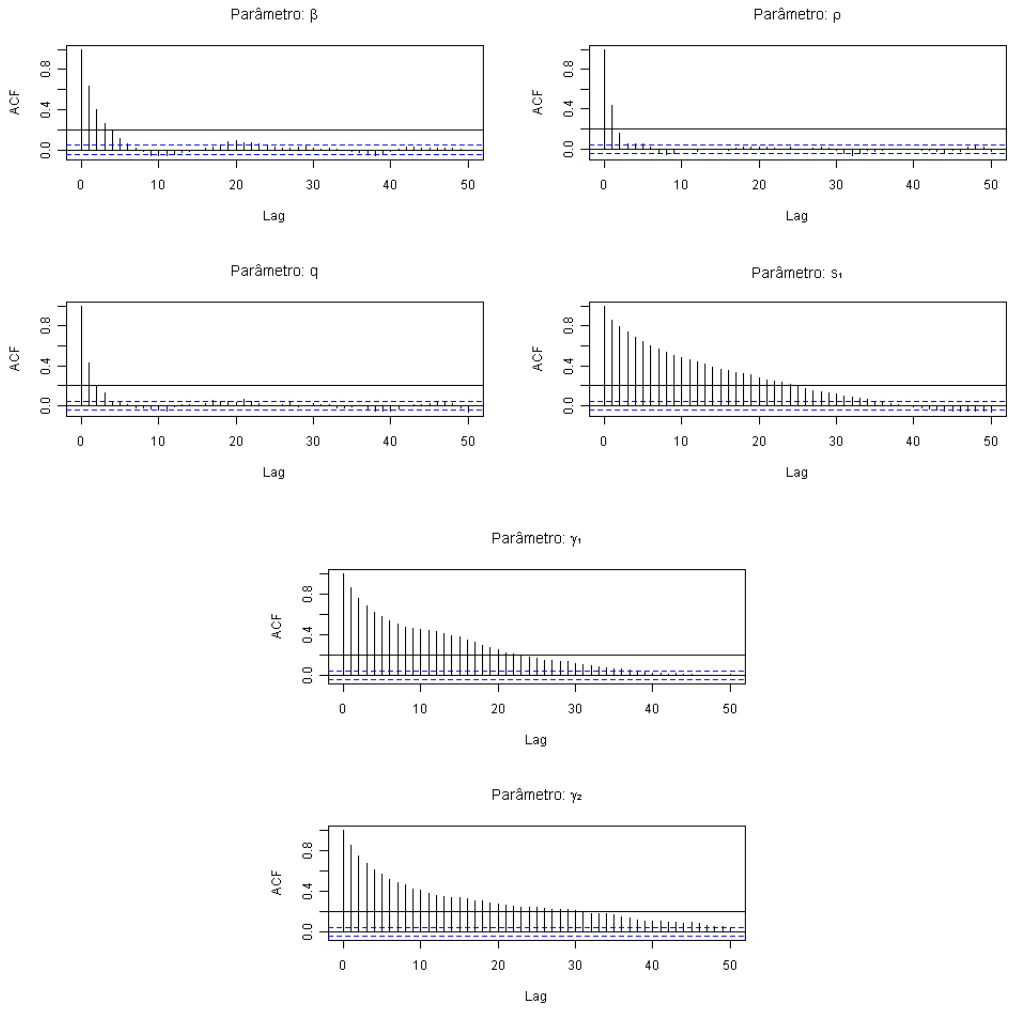


Figura 115: Autocorrelações para as amostras geradas dos parâmetros sem espaçamento amostral considerando a distribuição a priori não informativa o conjunto de dados real no modelo com taxa de remoção homogênea