

Pedro Luiz Ramos

**ASPECTOS COMPUTACIONAIS PARA
INFERÊNCIA NA DISTRIBUIÇÃO GAMA
GENERALIZADA**

Presidente Prudente, SP - Brasil

24 de fevereiro de 2014

Pedro Luiz Ramos

ASPECTOS COMPUTACIONAIS PARA INFERÊNCIA NA DISTRIBUIÇÃO GAMA GENERALIZADA

Dissertação apresentada ao Programa de Pós-graduação em Matemática Aplicada e Computacional da Universidade Estadual Paulista "Júlio de Mesquita Filho" como requisito para a obtenção do Título de Mestre em Matemática Aplicada e Computacional.

Universidade Estadual Paulista – UNESP

Faculdade de Ciências e Tecnologia

Programa de Pós-Graduação

Orientador: Prof. Dr. Jorge Alberto Achcar

Presidente Prudente, SP - Brasil

24 de fevereiro de 2014

R145a Ramos, Pedro Luiz.
Aspectos computacionais para inferência na distribuição gama
generalizada / Pedro Luiz Ramos. – Presidente Prudente: [s.n.], 2014
164 p. : il.

Orientador: Jorge Alberto Achcar
Dissertação de Mestrado – Universidade Estadual Paulista – UNESP
Faculdade de Ciências e Tecnologia
1. Distribuição Gama Generalizada. 2. Métodos Bayesianos. I. Achcar,
Jorge Alberto. II. Universidade Estadual Paulista. Faculdade de Ciências e
Tecnologia. III. Aspectos computacionais para inferência na distribuição
gama generalizada.

BANCA EXAMINADORA



PROF. DR. JORGE ALBERTO ACHCAR
ORIENTADOR



PROF. DR. FERNANDO ANTÔNIO MOALA
UNESP/FCT



PROF. DR. CARLOS APARECIDO DOS SANTOS
UEM



PEDRO LUIZ RAMOS

Presidente Prudente (SP), 24 de fevereiro de 2014.

RESULTADO: A PROVA DO

*Dedico este trabalho a minha avó Carmem, a minha mãe Rita
e à minha noiva Maysa, por tornar minha vida cada dia mais feliz.*

Agradecimentos

É difícil agradecer a todas as pessoas que, de algum modo, fizeram parte da minha vida, por isto agradeço a todos de coração.

Dedico este trabalho a minha mãe Rita e minha avó Carmem, pela determinação e luta na minha formação e do meu irmão.

À minha noiva Maysa, ofereço um agradecimento mais do que especial, por ter me aturado nos momentos de estresse e me incentivando durante todo o período de desenvolvimento deste trabalho, tornando minha vida cada dia mais feliz.

Agradeço ao meu orientador Prof. Dr. Jorge Alberto Achcar, que teve paciência e fundamental colaboração para o desenvolvimento deste trabalho e aos professores Dr. Fernando A. Moala, Dr. Josmar Mazucheli e Dr. Carlos A. Dos Santos pelas valiosas sugestões e comentários apresentados durante o exame de qualificação e de defesa.

Ao meu irmão Eduardo pela grande ajuda e conhecimento transmitido ao longo da pós-graduação e pela paciência em me ajudar em vários tópicos deste trabalho e ao meu pai Luiz por toda ajuda e correção deste trabalho.

Aos meus amigos Carlos e Frank que, ao longo dos últimos oito anos, posso considera-los como verdadeiros amigos.

Agradeço a meus colegas Marcelo, Leandro, Ana Paula, Taciana, Débora, Renato e Suelen, colegas de classe e com certeza futuros excelentes profissionais.

Agora agradeço aos funcionários da Seção de Pós-Graduação pelo auxílio prestado durante o decorrer do curso de mestrado.

Também agradeço ao Conselho de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro, possibilitando o desenvolvimento deste trabalho.

*"The Universe cannot be read until we have
learned the language in which it is written"*

Galileo Galilei

Resumo

Nesta dissertação, tem-se como objetivo desenvolver aspectos computacionais a fim de realizar inferência nos parâmetros da distribuição Gama Generalizada. A distribuição Gama Generalizada tem se mostrado um excelente modelo que leva a melhores inferências quando utilizamos dados de sobrevivência. No entanto, os Estimadores de Máxima Verossimilhança (EMV) da distribuição Gama Generalizada não têm o comportamento estável e para muitos casos não existem, apenas para amostras muito grandes. Na abordagem Bayesiana o problema permanece de tal forma que diferentes distribuições a priori podem ter grande efeito nas estimativas a posteriori de interesse. Neste trabalho, ao se utilizar a inferência Clássica, será demonstrado uma forma de simplificar as equações de verossimilhança, minimizando o problema de instabilidade do método. Considerando-se a abordagem Bayesiana, serão propostas diferentes distribuições a priori, empíricas ou não-informativas e seus efeitos serão estudados nas estimativas a posteriori. Serão propostos também diferentes métodos exploratórios, a fim de se obter bons valores iniciais para serem utilizados nos procedimentos de estimação dos parâmetros, considerando-se dados censurados (censura aleatória) sob o enfoque Clássico e Bayesiano. Por fim, serão apresentados alguns exemplos de aplicações, utilizando-se dados da literatura e dados reais.

Palavras-chaves: Distribuição Gama Generalizada, Estimador de Máxima Verossimilhança, Métodos Bayesianos, Dados de Sobrevivência.

Abstract

The most important goals of this research project are related to the study of computational aspects to get inferences for the parameters of the Generalized Gamma distribution. The Generalized Gamma distribution is a good option to get better inferences for lifetime data, but usually in applications, the maximum likelihood estimators (MLE) of the parameters of the distribution could be very unstable depending on appropriated initial values in the numerical iterative procedure, especially for small sample sizes. Under a Bayesian approach, we observe that the posterior distributions could depend heavily on the choice of a prior distribution. In this work, using the classical inference will be shown a way to simplify the likelihood equations, minimizing the problem of instability of the method. Under a Bayesian approach, we will study the effect of different empirical and non-informative prior distributions on the posterior estimates, especially for small sample sizes. Numerical methods are also proposed in this work, in order to get good initial values used to improve the obtained inferences, considering censored data (random censoring) under the Classical and Bayesian approach. Simulated and real data sets will be used to illustrate the proposed methodology.

Key-words: generalized gamma distribution, maximum likelihood estimators, Bayesian methods, lifetime data.

Lista de ilustrações

Figura 1 – Diferentes formas da função de risco.	22
Figura 2 – TTT plot de diferentes distribuições, que indicam a forma da função de risco.	24
Figura 3 – Funções densidade de probabilidade e acumulada da distribuição gama para diferentes valores de ϕ e μ	26
Figura 4 – Funções de sobrevivência e de risco da distribuição gama com diferentes valores de ϕ e μ	27
Figura 5 – Funções densidade de probabilidade e acumulada da distribuição Weibull para diferentes valores de α e β	28
Figura 6 – Funções de sobrevivência e de risco da distribuição Weibull com diferentes valores de α e β	28
Figura 7 – Funções densidade de probabilidade e acumulada da distribuição log-normal para diferentes valores de μ e σ	29
Figura 8 – Funções de sobrevivência e de risco da distribuição log-normal com diferentes valores de μ e σ	30
Figura 9 – Funções densidade de probabilidade e acumulada da distribuição Gama Generalizada para diferentes valores de ϕ , μ e α	44
Figura 10 – Funções de sobrevivência e de risco da distribuição Gama Generalizada para diferentes valores de ϕ , μ e α	46
Figura 11 – Densidades marginais a posteriori, gráficos de séries de temporais e gráficos de autocorrelação para os valores gerados de ϕ , μ e α , utilizando-se os dados de tempo de vida de 30 unidades em um experimento industrial.	88
Figura 12 – TTT-plot, função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada, utilizando-se os dados de tempo de vida de 30 unidades em um experimento industrial.	90
Figura 13 – Densidades marginais a posteriori, gráficos de séries de temporais e gráficos de autocorrelação para os valores gerados de ϕ , μ e α utilizando-se os dados de precipitação pluvial total mensal.	92
Figura 14 – TTT-plot, função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada, utilizando-se os dados de precipitação pluvial total mensal.	93
Figura 15 – Densidades marginais a posteriori, gráficos de séries de temporais e gráficos de autocorrelação para os valores gerados de ϕ , μ e α , utilizando-se os dados de tempo de vida de 30 unidades em um experimento industrial.	95

Figura 16 – TTT-plot, função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada, utilizando-se os tempos de sobrevida de 97 pacientes com câncer de pulmão. 96

Lista de tabelas

Tabela 1	– Estimativas das médias e desvios-padrão dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, através dos Métodos 1 e 2 para diferentes valores de θ e n	67
Tabela 2	– Probabilidades de cobertura com o nível de confiança de 95% dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ utilizando-se os Métodos 1 e 2.	67
Tabela 3	– Estimativas das médias e desvios-padrão dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , 20% de censura e valores iniciais obtidos através do método 1. . . .	68
Tabela 4	– Estimativas das médias e desvios-padrão dos valores iniciais $\theta^{(0)} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, obtidos através do método 3 e dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, considerando diferentes valores de θ e 20% de censura.	69
Tabela 5	– Estimativas das médias e desvios-padrão dos valores iniciais $\theta^{(0)} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, obtidos através do método 4 e dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, considerando diferentes valores de θ e 20% de censura.	69
Tabela 6	– Probabilidades de cobertura com o nível de confiança de 95% dos EMV's encontrados para 1000 amostras, utilizando-se os métodos 3 e 4, considerando-se 20% de censura e diferentes valores de θ e n	70
Tabela 7	– Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	72
Tabela 8	– Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	72
Tabela 9	– Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	73
Tabela 10	– Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$	73

Tabela 11 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\sigma_\theta^2 = (10, 10, 10)$	74
Tabela 12 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_\theta^2 = (1, 1, 1), \sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$	75
Tabela 13 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_\theta^2 = (1, 1, 1), \sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$	75
Tabela 14 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_\theta^2 = (1, 1, 1), \sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$	76
Tabela 15 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\sigma_\theta^2 = (1, 1, 1), \sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$	76
Tabela 16 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\sigma_\theta^2 = (10, 10, 10)$	77
Tabela 17 – Médias e desvios-padrão das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e considerando as distribuição a priori $\theta_i \sim \text{Uniforme}(0, 40)$ e $\theta_i \sim \text{Gamma}(0.01, 0.01)$	78
Tabela 18 – Probabilidades de cobertura com o nível de confiança de 95% das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$ e considerando as distribuição a priori $\theta_i \sim \text{Uniforme}(0, 40)$ e $\theta_i \sim \text{Gamma}(0.01, 0.01)$	78
Tabela 19 – Médias e desvios-padrão das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$, utilizando-se a distribuição a priori obtida pela Regra de Jeffreys Modificada e considerando-se diferentes valores de θ e n	79

Tabela 20 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a priori obtida pela Regra de Jeffreys Modificada.	80
Tabela 21 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_{\theta}^2 = (1, 1, 1), \sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	81
Tabela 22 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_{\theta}^2 = (1, 1, 1), \sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	82
Tabela 23 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_{\theta}^2 = (1, 1, 1), \sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	82
Tabela 24 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_{\theta}^2 = (1, 1, 1), \sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$	83
Tabela 25 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	83
Tabela 26 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_{\theta}^2 = (1, 1, 1), \sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	84
Tabela 27 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha}), \sigma_{\theta}^2 = (1, 1, 1), \sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	84

Tabela 28 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	85
Tabela 29 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$	85
Tabela 30 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$	86
Tabela 31 – Conjunto de dados de tempo de vida de 30 unidades em um experimento industrial.	88
Tabela 32 – Médias a posteriori, desvios-padrão a posteriori, erros de Monte Carlo e intervalos de credibilidade de 95% dos parâmetros ϕ, μ e α	89
Tabela 33 – Resultados dos Critérios DIC, BIC e AIC de diferentes distribuições de probabilidade, considerando-se os dados de um experimento industrial.	89
Tabela 34 – Conjunto de dados de precipitação pluvial total mensal (em mm), referente ao mês de maio, considerando um período de 56 anos (1947 a 2003).	91
Tabela 35 – Médias a posteriori, Desvios-padrão a posteriori, Erros de Monte Carlo e Intervalos de credibilidade de 95% dos parâmetros ϕ, μ e α	91
Tabela 36 – Resultados dos Critérios DIC, BIC e AIC de diferentes distribuições de probabilidade, considerando-se os dados de precipitação pluvial total mensal.	92
Tabela 37 – Os tempos de sobrevida (em dias) de 97 pacientes com câncer de pulmão (+ indica a presença de censura).	94
Tabela 38 – Medianas a posteriori, desvios-padrão a posteriori, erros de Monte Carlo e intervalos de credibilidade de 95% dos parâmetros ϕ, μ e α	94

Sumário

	Introdução	18
1	Revisão Bibliográfica	21
1.1	Análise de Sobrevivência	21
1.1.1	Dados de Sobrevivência	21
1.1.2	Gráfico TTT-plot	23
1.1.3	Censura	24
1.1.4	Alguns Modelos Paramétricos	25
1.1.4.1	Distribuição Gama	25
1.1.4.2	Distribuição Weibull	27
1.1.4.3	Distribuição Log-normal	28
1.2	Análise Clássica	30
1.2.1	Estimador de Máxima Verossimilhança	30
1.3	Métodos Bayesianos	32
1.3.1	Distribuição a Priori	32
1.3.1.1	Priori Uniforme	33
1.3.1.2	Priori Gama	33
1.3.1.3	Regra de Jeffreys	33
1.3.1.4	Priori de Jeffreys	34
1.3.1.5	Priori de Zellner (MDIP)	35
1.3.2	Distribuição a Posteriori	37
1.3.3	Crítérios de Seleção de Modelos	37
1.4	Métodos de Monte Carlo via Cadeias de Markov	39
1.4.1	Amostrador de Gibbs	39
1.4.2	Metropolis Hastings	40
1.4.3	Algoritmo Híbrido	41
1.4.4	Diagnósticos de Convergência	42
2	Distribuição Gama Generalizada	44
2.1	Média, Moda e Variância	45
2.2	Função de Sobrevivência e de Risco	45
2.3	Outros Resultados Importantes	46
2.4	Métodos para Gerar Valores da Distribuição Gama Generalizada	47
2.4.1	Observações Completas	47
2.4.2	Observações Censuradas	48
3	Análise Clássica da Distribuição Gama Generalizada	50

3.1	Estimador de Máxima Verossimilhança para Dados Completos	50
3.2	Estimador de Máxima Verossimilhança para Dados com Censura Tipo I	51
4	Análise Bayesiana da Distribuição Gama Generalizada	53
4.1	Densidades a Posteriori Considerando Observações Completas	53
4.1.1	Densidade a Posteriori com a Priori Uniforme	53
4.1.2	Densidade a Posteriori com a Priori Gama	54
4.1.3	Densidade a Posteriori com a Regra de Jeffreys	55
4.1.4	Densidade a Posteriori com a Priori de Jeffreys	55
4.1.5	Densidade a Posteriori com a Priori de Zellner	56
4.1.6	Densidade a Posteriori com a Regra de Jeffreys Modificada	57
4.2	Densidades a Posteriori Considerando Observações Censuradas	58
4.2.1	Densidade a Posteriori com a Priori Uniforme	58
4.2.2	Densidade a Posteriori com a Priori Gama	58
4.3	Algoritmo Híbrido (MCMC)	59
5	Equações Úteis para Encontrar Bons Valores Iniciais	61
5.1	Análise Clássica	61
5.2	Análise Bayesiana	62
6	Análise Numérica	65
6.1	Análise Clássica	66
6.1.1	Observações Completas	66
6.1.2	Observações Censuradas	68
6.2	Análise Bayesiana	70
6.2.1	Observações Completas	71
6.2.1.1	Prioris Empíricas	71
6.2.1.2	Prioris Não-Informativas	77
6.2.2	Observações Censuradas	80
7	Aplicações	87
7.1	Dados de um Experimento Industrial	87
7.2	Dados de Precipitação Pluvial Total Mensal	90
7.3	Dados de Sobrevida de Pacientes com Câncer de Pulmão	93
	Considerações Finais e Propostas Futuras	97
	Referências	99

Apêndices	105
APÊNDICE A – Demonstração de que as densidades a posteriori são impróprias	106
APÊNDICE B – Demonstração de que a densidade a posteriori 4.20 é própria	117
APÊNDICE C – Outros métodos para se obter valores iniciais . . .	126
APÊNDICE D – Programas Utilizados	130
APÊNDICE E – Bayesian Inference for Two-Parameter Gamma Dis- tribution Assuming Different Noninformative Pri- ors	146

Introdução

A formulação de modelos probabilísticos vem se tornando um dos desafios mais importantes a ser encarado na atualidade em análises estatísticas. Avanços nas áreas computacionais e de simulações numéricas têm possibilitado incorporar modelos cada vez mais complexos e sofisticados, resultando em modelos mais adequados, capazes de descrever problemas reais.

Muitas distribuições probabilísticas ou modelos probabilísticos têm sido utilizadas em análise de confiabilidade para modelar tempos de falha de produtos, materiais e situações clínicas. Dentre elas, podem ser citadas a distribuição Weibull, Gama, Valor Extremo, Log-normal (Lawless, 1982), Exponencial Generalizada (Gupta e Kundu, 1999, 2001), entre outras.

A distribuição Gama Generalizada de três parâmetros (ou Gama-Weibull), tem se mostrado muito flexível para modelar dados de confiabilidade, pois acomoda várias formas da função de risco, sendo estas: constante, crescente, decrescente, forma de banheira e banheira inversa.

Introduzida por Stacy (1962), distribuições relevantes podem ser obtidas a partir da distribuição Gama Generalizada (GG), como as distribuições Weibull, Gama, Log-normal e a Normal Generalizada. A distribuição Normal generalizada é também uma distribuição que inclui várias distribuições conhecidas, como as distribuições half-normal, Rayleigh e Maxwell-Boltzmann.

Aalo, et al. (2005) utilizam com sucesso este modelo para analisar a degradação do desempenho de sistemas de comunicação *wireless* (rede sem fio). Li et al. (2011) propõem utilizar a distribuição GG para se obter diferentes técnicas de processamentos em imagens do tipo SAR (*Synthetic aperture radar*). Tahai e Meyer (1999) propõem novas metodologias para analisar citações em recentes publicações e encontrar revistas com maior influência, utilizando-se a distribuição GG. Também Marani et al. (1986) utilizam este modelo para analisar dados de qualidade do ar na área de Venice, na Itália.

Outras aplicações da distribuição Gama Generalizada podem ser visualizadas em Noortwijk (2001), Dadpay et al. (2007), Chang e Kim (2011), Balakrishnan e Peng (2006), Tahai e Meyer (1999), Raju e Srinivasan, (2002), Marani et al. (1986), Santos (2010) e Ahsanullah et al. (2013).

Algumas propriedades deste modelo foram obtidas por Stacy e Mihram (1965), assim como o método de estimação dos parâmetros, baseado no método da máxima verossimilhança (abordagem Clássica). Já Hager e Bain (1970) mostram que as três

equações não-lineares obtidas por Stacy e Mihram (1965) são muito instáveis, e ao isolarem um dos parâmetros, minimizaram o problema para duas equações.

No entanto, mesmo com estes resultados, os autores mostram que as equações de verossimilhança obtidas ainda continuam muito instáveis e que as estimativas de máxima verossimilhança (EMV's) variam, dependendo dos valores iniciais utilizados no método iterativo.

Huang e Hwang (2006) utilizam o método dos momentos para se obter inferências da distribuição GG. Khodabin e Ahmadabadi (2010) comparam este método com o método da máxima verossimilhança e concluem que, em geral, os EMV's possuem melhor desempenho. No entanto, ambos os métodos são instáveis e seus valores dependem do valor inicial, escolhido nos métodos de iteração.

O problema permanece utilizando a abordagem Bayesiana de tal forma que, diferentes distribuições a priori podem ter grande efeito nas estimativas a posteriori de interesse.

Neste trabalho, ao se utilizar a inferência Clássica, será demonstrado uma maneira de simplificar as equações de verossimilhança, minimizando o problema para apenas uma equação não linear. Por meio destes resultados eliminamos a instabilidade do método e também a necessidade de se utilizar bons valores iniciais, obtendo-se facilmente boas estimativas para os parâmetros da distribuição GG.

Considerando-se a abordagem Bayesiana, serão propostas diferentes distribuições a priori, empíricas ou não-informativas e seus efeitos serão estudados nas estimativas a posteriori.

Para se obter as distribuições a priori empíricas, serão propostos métodos exploratórios a fim de se obter bons valores iniciais que, combinados com o método dos momentos, possibilitarão a obtenção dos valores dos hiperparâmetros das distribuições a priori utilizadas.

Na literatura Bayesiana, diferentes distribuições a priori não-informativas podem ser utilizadas para se realizar inferências nos parâmetros da distribuição Gama Generalizada. Dentre elas podemos citar as prioris: de Jeffreys (Jeffreys, 1967), MDIP (Zellner, 1977, 1984) e a obtida pela Regra de Jeffreys (Box e Tiao, 1973). No entanto será demonstrado que tais distribuições a priori levam a distribuições a posteriori impróprias e não devem ser utilizadas em uma análise Bayesiana. Uma modificação na priori obtida através da Regra de Jeffreys foi realizada e demonstramos que, utilizando tal distribuição a priori, obtemos uma densidade a posteriori própria.

Em análise de confiabilidade, ao analisar dados de falha, certamente haverá unidades na amostra que podem não ter falhado, isto é, o exato tempo de falha das unidades não é conhecido. Estes tipos de observações são chamados de dados censurados (Hamada et al.,

2008).

Neste trabalho, também serão propostos diferentes métodos exploratórios a fim de se obter bons valores iniciais para os procedimentos de estimação dos parâmetros, considerando-se dados censurados (censura aleatória), sob o enfoque Clássico e Bayesiano. Esses resultados são de grande interesse prático, pois possibilitarão o uso da distribuição Gama Generalizada em diversas áreas de aplicação como medicina, engenharia e climatologia.

Por fim, iremos apresentar alguns exemplos de aplicações, utilizando-se dados da literatura e dados reais. Para isto, iremos considerar a distribuição Gama Generalizada para analisar tais conjuntos de dados e iremos comparar os resultados obtidos com outros modelos (se possível), como as distribuições Weibull, Gama e Log-normal, através dos testes AIC (Akaike, 1974), BIC (Schwarz, 1978) e DIC (Spiegelhalter et al., 2002).

A fim de elucidar o trabalho proposto, as informações se encontram organizadas da seguinte forma: no Capítulo 1, será feita uma revisão bibliográfica, apresentando alguns conceitos básicos de análise de sobrevivência, inferência Clássica, inferência Bayesiana e em especial sobre métodos MCMC (Monte Carlo via Cadeias de Markov); no Capítulo 2, serão apresentadas algumas propriedades da distribuição Gama Generalizada; nos Capítulos 3 e 4 serão apresentados os métodos de estimação dos parâmetros da distribuição GG, considerando-se observações completas e censuradas, sob o enfoque Clássico e Bayesiano e o algoritmo MCMC, utilizado para se obter as estimativas a posteriori (abordagem Bayesiana); no Capítulo 5 apresentaremos os métodos exploratórios para se obter bons valores iniciais; no Capítulo 6 apresentamos algumas conclusões baseadas em simulações numéricas; no Capítulo 7 serão realizadas algumas aplicações utilizando-se dados da literatura e dados reais; por fim, no Capítulo 8 apresentamos algumas conclusões e perspectivas futuras.

Alguns dos programas desenvolvidos no *software* R e no OPENBUGS, podem ser visualizados nos Apêndices, os demais programas podem ser obtidos diretamente com o autor desta dissertação.

1 Revisão Bibliográfica

1.1 Análise de Sobrevivência

O rápido avanço no desenvolvimento de novas tecnologias tem aumentado intensamente a competição entre inúmeras empresas de tecnologia, visando-se produzir produtos mais confiáveis e de excelente qualidade. Consumidores em geral, esperam comprar produtos confiáveis, em que com alta probabilidade executem suas devidas funções, durante um grande período de tempo.

Para verificar se determinados produtos ou componentes irão realizar suas devidas funções, durante um período específico de tempo, colocamos estes sob teste, estimando-se características relacionadas aos seus tempos de vida, como o tempo médio de vida ou a probabilidade deste componente durar ao menos quatro anos. Esta área de estudo é denominada análise de sobrevivência ou análise de confiabilidade.

A análise de sobrevivência pode ser definida, de forma ampla, como a análise do tempo até a ocorrência de um dado evento (Colosimo e Giolo, 2006). Este tempo, denominado de tempo de falha, pode ser atribuído à falha de um componente, ou até a morte de um paciente. Uma característica muito importante da análise de sobrevivência é a presença de observações incompletas ou parciais. Estas observações, denominadas censuras, referem-se a situações em que, por alguma razão, não observamos a ocorrência do evento de interesse até o final do experimento.

Apresentaremos a seguir, alguns conceitos importantes utilizados em análise de sobrevivência.

1.1.1 Dados de Sobrevivência

Em análise de sobrevivência as observações (ou respostas) são caracterizadas pelos tempos de falhas, ou também pela presença de censura. Estas observações têm como escala de medida o tempo até a ocorrência do evento de interesse. Desta forma a variável aleatória (v.a) T possuirá valores reais não-negativos. A variável aleatória $T > 0$ pode ser expressa através de funções matemáticas. Dentre elas, iremos utilizar a função densidade de probabilidade, $f(t)$, a acumulada, $F(t)$, a de sobrevivência, $S(t)$ e a função de risco, $h(t)$. A função densidade de probabilidade de uma v.a T positiva, com a probabilidade de

um componente experimentar o evento de interesse no tempo t , é dada por

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t}, \quad f(t) \geq 0. \quad (1.1)$$

A função de sobrevivência de uma v.a T positiva, com a probabilidade de um componente não falhar (ou do componente sobreviver) até certo tempo t , é obtida através das seguintes relações (veja, Lawless, 1982),

$$\begin{aligned} S(t) &= P[T > t] = 1 - P[T \leq t] \\ &= 1 - \int_0^t f(t)dt = 1 - F(t), \quad 0 < S(t) < 1, \end{aligned} \quad (1.2)$$

sendo $F(t)$ a probabilidade de um componente experimentar o evento de interesse até o tempo t . A $F(t)$ é também conhecida como função acumulada (f.d.a.) da v.a T .

A função de risco $h(t)$, conhecida também como taxa de falha instantânea, informa a probabilidade de falha instantânea de um componente dado que sobreviveu até o tempo t . Esta pode ser obtida através dos seguintes resultados :

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \quad f(t) \geq 0. \quad (1.3)$$

Muito útil para descrever a distribuição do tempo de vida de pacientes (ou componentes), a função de risco caracteriza a forma em que a taxa de falha instantânea muda através do tempo. A Figura 1 mostra diferentes formas que a função de risco pode assumir, dentre elas: constante, decrescente, crescente, forma de banheira e inversa forma de banheira.

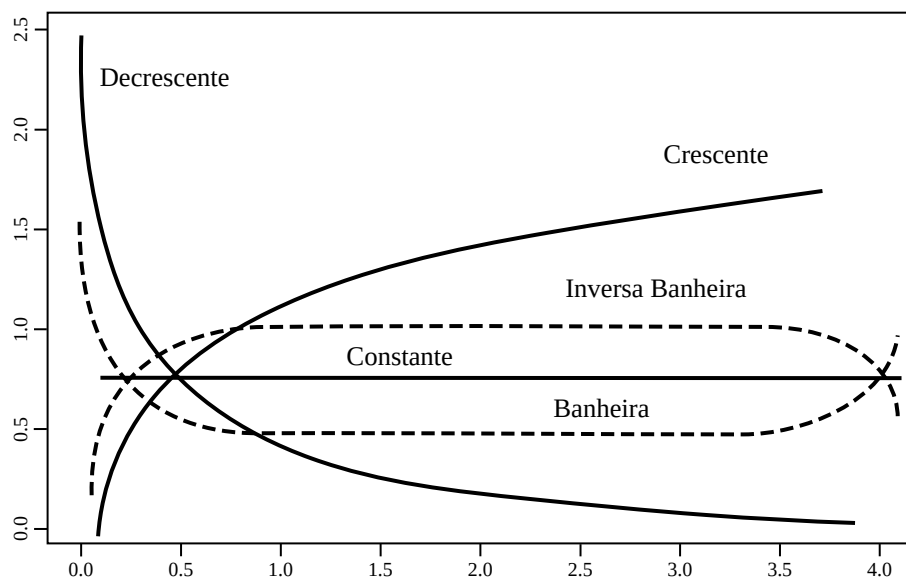


Figura 1 – Diferentes formas da função de risco.

Algumas relações matemáticas importantes podem ser obtidas através das funções anteriores, dentre elas:

$$h(t) = \frac{S'(t)}{S(t)} = \frac{\partial}{\partial t} \log(S(t)),$$

$$S(t) = \exp \left[- \int_0^t h(t) dt \right],$$

$$f(t) = h(t)S(t).$$

Desta forma, o conhecimento de uma função, por exemplo $h(t)$, implica no conhecimento das outras funções, $f(t)$, $F(t)$, e $S(t)$.

1.1.2 Gráfico TTT-plot

Barlow e Campo (1975) propuseram uma técnica gráfica para verificar o comportamento da função de risco, chamada de TTT-plot (tempo total em teste). O gráfico é construído com o plot consecutivo das quantidades $[r/n, G(r/n)]$, sendo $G(r/n)$ a função dada por

$$G(r/n) = \frac{\sum_{i=1}^r t_i + (n-r)t_{(r)}}{\sum_{i=1}^n t_i}, \quad (1.4)$$

onde $r = 1, \dots, n$, $i = 1, \dots, n$ sendo $t_{(i)}$ as estatística de ordem da amostra.

Utilizando-se a curva TTT em escala empírica, pode-se detectar o tipo de função de risco que os dados de tempo de vida possuem. Uma reta diagonal indica que os dados possuem uma função de risco contante. Se a curva for convexa, o risco é decrescente. Já a côncava indica que o risco é crescente. A forma convexa e depois côncava indica que a função tem risco em forma de banheira e se for côncava e convexa, tem a forma inversa à de banheira. A Figura 2 seguir mostra de forma mais clara, como verificar o comportamento da função de risco, utilizando-se o TTT-plot.

Algumas transformações TTT podem ser utilizadas em diferentes aplicações. Nair et al. (2008) mostram algumas destas transformações, aplicadas em análise de sobrevivência.

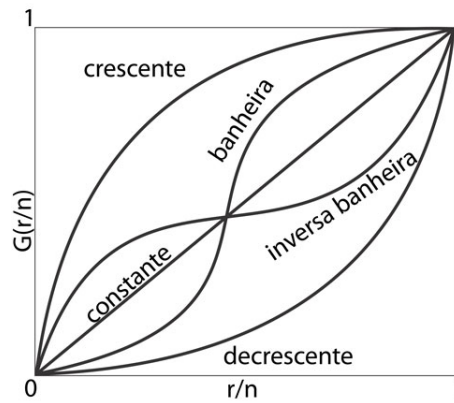


Figura 2 – TTT plot de diferentes distribuições, que indicam a forma da função de risco.

1.1.3 Censura

Em estudos que envolvem uma resposta temporal, uma característica muito observada é a longa duração até as ocorrências dos eventos de interesse. Pelo fato destas ocorrências possuírem longa duração, em muitos casos, é inviável esperar até o final do experimento.

Uma característica muito importante da análise de sobrevivência passa a ser a presença de observações incompletas ou parciais, denominadas censuras (veja, Hamada et al., 2008).

É importante salientar que, mesmo incompletas ou parcialmente conhecidas, estas observações não podem ser descartadas, pois fornecem informações relevantes sobre o tempo de vida dos componentes em estudo. A omissão destas observações também pode acarretar em conclusões viciadas ao calcularmos estatísticas de interesse.

Os dados de sobrevivência podem ser representados pelo par (t_i, δ_i) , sendo t_i o tempo de falha ou censura do i -ésimo componente e δ_i a variável indicadora de censura, isto é

$$\delta_i = \begin{cases} 1, & \text{se } t_i \text{ é uma observação completa} \\ 0, & \text{se } t_i \text{ é uma observação censurada} \end{cases}, \quad i = 1, \dots, n.$$

Os mecanismos de censura são diferenciados. Dentre eles iremos citar quatro tipos de censura muito utilizadas (Lawless, 1982).

- **Censura tipo I:** Ocorre quando removemos os componentes que não falharam em um tempo pré-estabelecido, ou seja, o experimento acaba, após fixarmos um tempo limite. Assim o tempo de falha de todos os componentes, que não puderam ser observados completamente, passam a ser considerados censurados.

- **Censura tipo II:** Acontece quando o experimento é encerrado após a ocorrência de falha em um número fixado de componentes. Desta forma, ao se realizar um experimento com n componentes, o pesquisador pode encerrar o experimento após r componentes falharem, possuindo no final $n - r$ observações censuradas.
- **Censura aleatória:** Ocorre quando o componente em estudo é eliminado do experimento por ter ocorrido outro evento, que não seja a falha de interesse. Este tipo de censura ocorre também com muita frequência em pesquisas médicas, quando um paciente é retirado do experimento, sem que tenha ocorrido a falha, ou que, por alguma razão, venha a morrer por uma causa diferente da estudada.
- **Censura Intervalar:** Acontece quando não se sabe exatamente o tempo em que o evento de interesse ocorreu, tendo apenas como informação um intervalo de tempo especificado.

1.1.4 Alguns Modelos Paramétricos

Diferentes métodos podem ser utilizados para se analisar dados de sobrevivência. Para tais análises, empregamos comumente abordagens paramétricas e não paramétricas, (veja, Meeker e Escobar, 1998).

Utilizando-se a abordagem paramétrica, podemos citar vários modelos probabilísticos que possuem destaque em análise de sobrevivência. Em particular, nesta seção, serão apresentados três modelos que descrevem adequadamente várias situações reais. Estas distribuições são conhecidas como: Gama, Weibull e Log-normal.

1.1.4.1 Distribuição Gama

A distribuição gama é comumente utilizada para se resolver problemas em análise de confiabilidade, previsões do tempo, para modelar seguros de crédito, entre outros.

Seja T uma variável aleatória, com uma distribuição de probabilidade gama denotado por $\text{Gama}(\phi, \mu)$, cuja função de densidade de probabilidade é dada por

$$f(t|\phi, \mu) = \frac{\mu^\phi}{\Gamma(\phi)} t^{\phi-1} \exp(-\mu t), \quad (1.5)$$

para todo $t > 0$, $\phi > 0$ e $\mu > 0$, sendo $\Gamma(\phi)$ a função gama definida como

$$\Gamma(\phi) = \int_0^\infty e^{-x} x^{\phi-1} dx. \quad (1.6)$$

Podemos ver na Figura 3 diferentes formas para a f.d.p. e f.d.a. da distribuição gama.

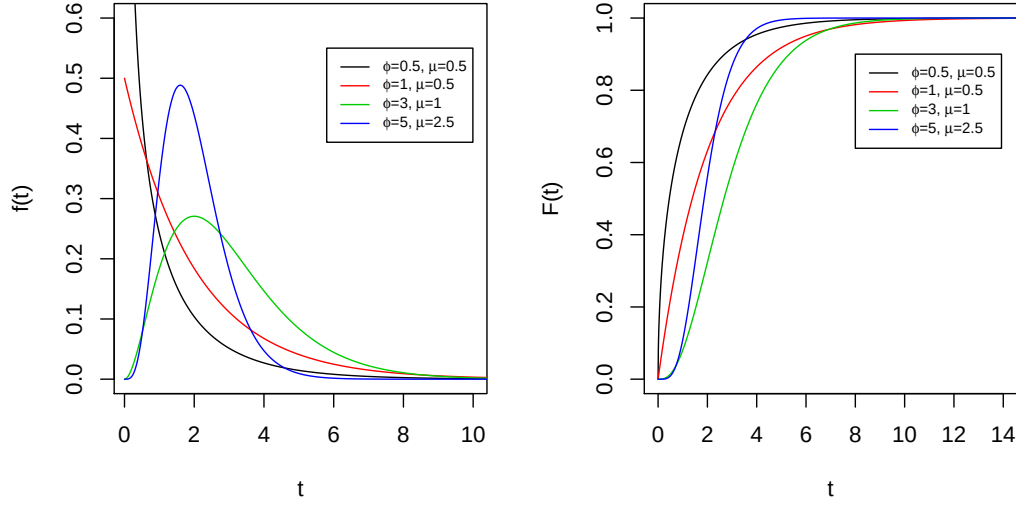


Figura 3 – Funções densidade de probabilidade e acumulada da distribuição gama para diferentes valores de ϕ e μ .

A média e a variância da distribuição gama 1.5 são dadas respectivamente por

$$E(T) = \frac{\phi}{\mu} \text{ e } Var(T) = \frac{\phi}{\mu^2}. \quad (1.7)$$

As funções de sobrevivência e de risco são dadas respectivamente por

$$S(t|\phi, \mu) = \frac{\Gamma[\mu t, \phi]}{\Gamma(\phi)} \text{ e } h(t|\phi, \mu) = \frac{\mu^\phi t^{\phi-1} \exp(-\mu t)}{\Gamma[\mu t, \phi]}, \quad (1.8)$$

sendo $\Gamma[y, x] = \int_x^\infty w^{y-1} e^{-w} dw$ denominada de gama incompleta superior.

Observa-se na Figura 4 diferentes formas para a função de sobrevivência e a função de risco da distribuição gama.

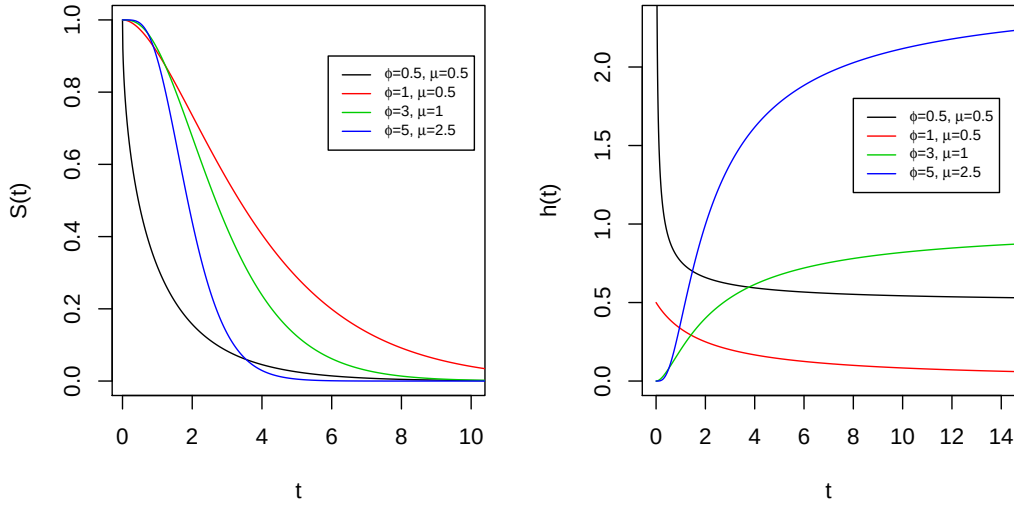


Figura 4 – Funções de sobrevivência e de risco da distribuição gama com diferentes valores de ϕ e μ .

1.1.4.2 Distribuição Weibull

Nomeada em homenagem a Waloddi Weibull, a distribuição Weibull é uma das principais distribuições de probabilidade contínuas, sendo utilizada rotineiramente em diversas áreas de pesquisa.

Se T segue uma distribuição Weibull denotada por $\text{Weibull}(\alpha, \beta)$, então sua f.d.p. é dada por

$$f(t|\alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{t}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{t}{\beta}\right)^{\alpha}\right), \quad (1.9)$$

para todo $t > 0$, $\alpha > 0$ e $\beta > 0$.

Podemos ver na Figura 5 diferentes formas para a f.d.p. e f.d.a. da distribuição Weibull.

A média e a variância da distribuição Weibull (1.9) são dadas respectivamente por

$$E(T) = \beta \Gamma\left(1 + \frac{1}{\alpha}\right) \text{ e } Var(T) = \beta^2 \left(\Gamma\left(1 + \frac{1}{\alpha}\right) - \Gamma\left(1 + \frac{1}{\alpha}\right)^2\right). \quad (1.10)$$

As funções de sobrevivência e de risco são dadas respectivamente por

$$S(t|\alpha, \beta) = \exp\left(-\left(\frac{t}{\beta}\right)^{\alpha}\right) \text{ e } h(t|\alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{t}{\beta}\right)^{\alpha-1}. \quad (1.11)$$

Observa-se na Figura 6 diferentes formas para a função de sobrevivência e a função de risco da distribuição Weibull.

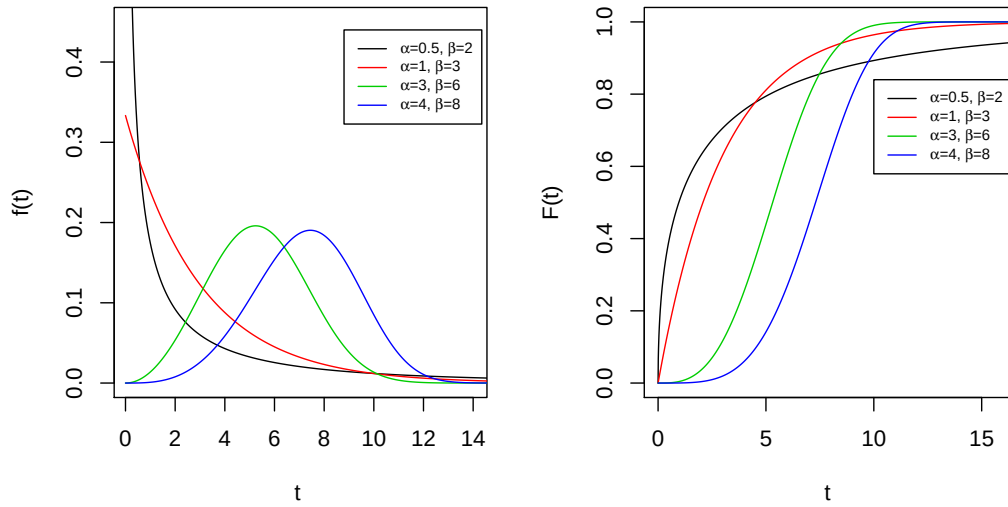


Figura 5 – Funções densidade de probabilidade e acumulada da distribuição Weibull para diferentes valores de α e β .

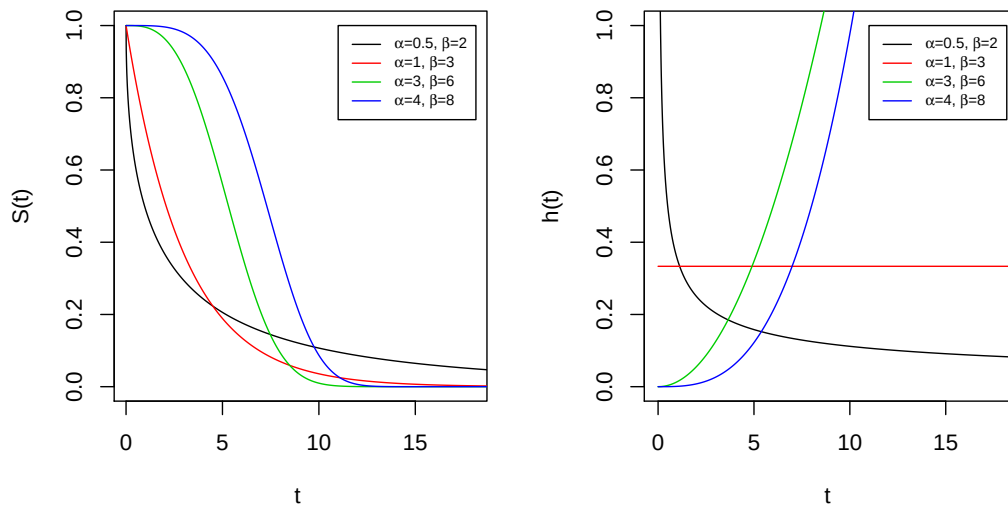


Figura 6 – Funções de sobrevivência e de risco da distribuição Weibull com diferentes valores de α e β .

1.1.4.3 Distribuição Log-normal

A distribuição log-normal tem sido utilizada como modelo em diversas aplicações em engenharia, medicina entre outras. Se Y possui uma distribuição normal com média μ e variância σ^2 e $T = \exp(Y)$, então dizemos que T possui uma distribuição log-normal e

sua f.d.p. é dada por

$$f(t|\mu, \sigma) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma t} \exp\left(-\frac{1}{2}\left(\frac{\log(t) - \mu}{\sigma}\right)^2\right), \quad (1.12)$$

para todo $t > 0$, $\mu > 0$ e $\sigma > 0$.

Podemos ver na Figura 7 diferentes formas para a f.d.p. e f.d.a. da distribuição Log-normal.

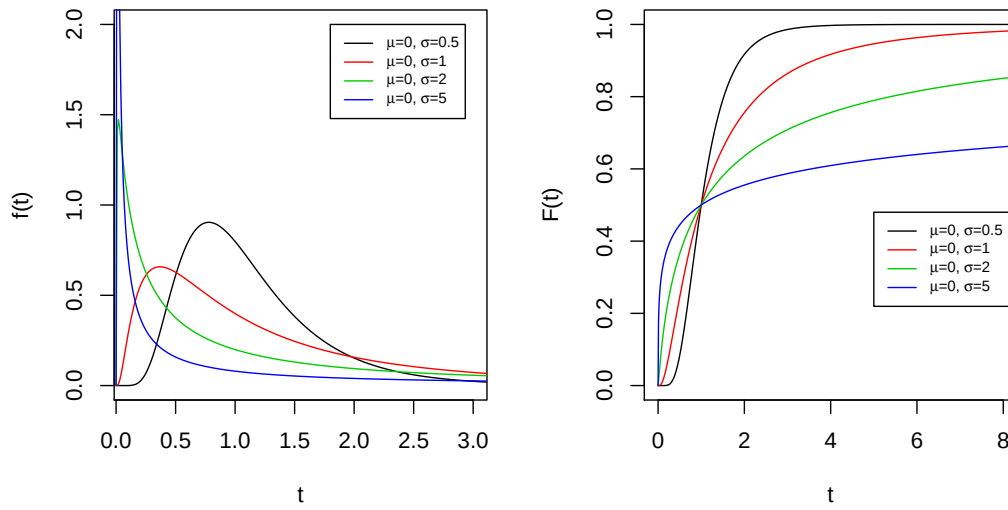


Figura 7 – Funções densidade de probabilidade e acumulada da distribuição log-normal para diferentes valores de μ e σ .

A média e variância da distribuição log-normal (1.12) são dadas respectivamente por

$$E(T) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \quad \text{e} \quad Var(T) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}. \quad (1.13)$$

As funções de sobrevivência e de risco são dadas respectivamente por

$$S(t|\alpha, \beta) = 1 - \Phi\left(\frac{\log(t)}{\sigma}\right) \quad \text{e} \quad h(t|\alpha, \beta) = \frac{1}{(2\pi)^{\frac{1}{2}}\sigma t \left(1 - \Phi\left(\frac{\log(t)}{\sigma}\right)\right)} \exp\left(-\frac{1}{2}\left(\frac{\log(t) - \mu}{\sigma}\right)^2\right), \quad (1.14)$$

sendo $\Phi(x) = \int_{-\infty}^0 \frac{1}{(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{u^2}{2}\right) du$.

Observa-se na Figura 8 diferentes formas para a função de sobrevivência e a função de risco da distribuição Log-normal.

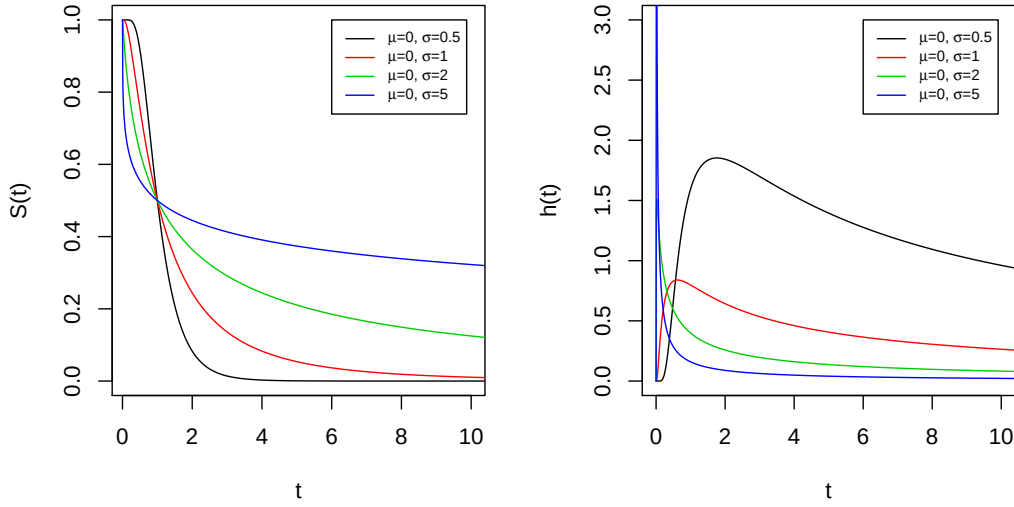


Figura 8 – Funções de sobrevivência e de risco da distribuição log-normal com diferentes valores de μ e σ .

1.2 Análise Clássica

Ao se realizar uma análise estatística, empregam-se, comumente, dois paradigmas alternativos: o clássico (ou frequentista) e o Bayesiano.

Na Análise Clássica o parâmetro θ desconhecido é tratado como uma constante fixa. Uma forma de se avaliar esse parâmetro é encontrar um estimador pontual para θ . Sendo $\mathbf{x} = (x_1, x_2, \dots, x_n)$ um vetor de n observações geradas por um modelo probabilístico $f(\mathbf{x}|\theta)$, é então escolhido algum princípio tal como estimador de máxima verossimilhança, mínimos quadrados, método dos momentos, entre outros.

1.2.1 Estimador de Máxima Verossimilhança

Dentre os métodos estatísticos de inferência Clássica, o método de máxima verossimilhança é preferível devido às suas melhores propriedades assintóticas. No método de máxima verossimilhança os estimadores são obtidos a partir da maximização da função de verossimilhança (ver, por exemplo, Casella e Berger, 2002).

A função de verossimilhança de θ , correspondente à amostra aleatória observada, é dada por

$$L(\theta, \mathbf{t}) = \prod_{i=1}^n f(t_i|\theta), \quad (1.15)$$

que resume a informação a respeito de $\theta = (\theta_1, \dots, \theta_k)$, contida nos dados amostrais x .

Na presença de censura, altera-se a função de verossimilhança do modelo probabilístico. No caso em que não há censura presente na amostra a função de verossimilhança é o produto das densidades completas (1.15). Na presença de censura do tipo I, ou censuras à direita, não se sabe exatamente quando a falha ocorre, se o intervalo do tempo de falha está (t, ∞) , a contribuição na função de verossimilhança também estará neste intervalo então $P[T > t] = S(t)$. A função de verossimilhança para dados com censura tipo I será escrita na forma

$$L(\boldsymbol{\theta}, \mathbf{t}) = \prod_{i=1}^r f(t_i|\boldsymbol{\theta}) \prod_{i=r+1}^n S(t_i|\boldsymbol{\theta}), \quad (1.16)$$

ou pode ser escrita também como

$$L(\boldsymbol{\theta}, \mathbf{t}) = \prod_{i=1}^n f(t_i|\boldsymbol{\theta})^{\delta_i} S(t_i|\boldsymbol{\theta})^{1-\delta_i} = \prod_{i=1}^n h(t_i|\boldsymbol{\theta})^{\delta_i} S(t_i|\boldsymbol{\theta}), \quad (1.17)$$

sendo δ a variável indicadora da censura. Para dados completos a variável $\delta = 1$, já para dados censurados a variável assume $\delta = 0$.

Para um modelo com número de parâmetros k , se a função de verossimilhança for diferenciável em θ_i , as equações de verossimilhança são obtidas através do sistema

$$\frac{\partial}{\partial \theta_i} \log(L(\boldsymbol{\theta}, \mathbf{t})) = 0, i = 1, 2, \dots, k. \quad (1.18)$$

As soluções deste sistema (1.18) fornecem os estimadores de máxima verossimilhança. Em muitos casos, métodos numéricos como os de Newton-Raphson são necessários para se encontrar tais soluções.

O algoritmo de Newton-Raphson é representado através das atualizações sucessivas de

$$\boldsymbol{\theta}^{(n+1)} = \boldsymbol{\theta}^{(n)} - [J(\boldsymbol{\theta}^{(n)})]^{-1} G(\boldsymbol{\theta}^{(n)}), \quad (1.19)$$

sendo $G(\boldsymbol{\theta})$, o vetor contendo as derivadas parciais de $\log(L(\boldsymbol{\theta}, \mathbf{t}))$ e $J(\boldsymbol{\theta})$ a matriz Jacobiana de $G(\boldsymbol{\theta})$.

Os estimadores de máxima verossimilhança de $\boldsymbol{\theta}$ são viciados para amostras pequenas. No caso de grandes amostras, tornam-se não-viciados e assintoticamente eficientes. Tais estimadores têm distribuição conjunta assintoticamente normal dada por

$$(\hat{\boldsymbol{\theta}}) \sim N_k[(\boldsymbol{\theta}), I^{-1}(\boldsymbol{\theta})] \text{ para } n \rightarrow \infty, \quad (1.20)$$

onde $I(\boldsymbol{\theta})$ é a matriz de Informação de Fisher resultante de

$$I(\boldsymbol{\theta}) = \begin{bmatrix} I_{11}(\boldsymbol{\theta}) & \dots & I_{1k}(\boldsymbol{\theta}) \\ \vdots & \ddots & \vdots \\ I_{1k}(\boldsymbol{\theta}) & \dots & I_{kk}(\boldsymbol{\theta}) \end{bmatrix},$$

sendo $I_{ij}(\boldsymbol{\theta})$, a quantidade denominada Informação de Fisher de $\boldsymbol{\theta}$ em i e j dada por

$$I_{ij}(\boldsymbol{\theta}) = E \left[\left(\frac{\partial}{\partial \theta_i \partial \theta_j} \log(f(\mathbf{x}|\boldsymbol{\theta})) \right)^2 \right], i, j = 1, 2, \dots, k. \quad (1.21)$$

Quando o tamanho amostral é grande, pode-se construir intervalos de confiança aproximados para os parâmetros individuais θ_i , com coeficiente de confiança $100(1 - \gamma)\%$, através das distribuições marginais dadas por

$$(\hat{\theta}_i) \sim N[(\theta), I_{ii}^{-1}(\boldsymbol{\theta})] \text{ para } n \rightarrow \infty, \quad (1.22)$$

Com estes resultados é possível construir intervalos de confiança individuais aproximados.

O estimador de máxima verossimilhança possui uma propriedade muito importante conhecida como *propriedade de invariância* (ver Casella e Berger, 2002, p. 319). Isto é, suponha que exista pelo menos uma transformação um a um que leva θ a $\tau(\theta)$, utilizando-se a propriedade de invariância do estimador de máxima verossimilhança, pode-se afirmar que se $\hat{\theta}$ é um EMV de θ , portanto, para qualquer função de $\tau(\theta)$, o EMV de $\tau(\theta)$ é $\tau(\hat{\theta})$.

1.3 Métodos Bayesianos

O uso dos métodos Bayesianos tem desempenhado um papel fundamental na análise de dados. Grande parte se deve avanço nos métodos computacionais e às boas inferências obtidas sem a necessidade de propriedades assintóticas.

Na Inferência Bayesiana, o parâmetro $\boldsymbol{\theta}$ é tratado como uma variável aleatória e a distribuição que representa o conhecimento sobre $\boldsymbol{\theta}$ (a quantidade desconhecida) é denominada distribuição a priori, e definida por $\pi(\boldsymbol{\theta})$. Esta distribuição expressa o conhecimento ou incerteza sobre $\boldsymbol{\theta}$, antes da obtenção dos dados amostrais $\mathbf{x}$. A função de verossimilhança resume toda informação a respeito de $\boldsymbol{\theta}$, contida nos dados amostrais $\mathbf{x}$.

Observado $\mathbf{x}$, uma forma natural de se combinar as informações resultantes da distribuição a priori e da função de verossimilhança, pode ser feita envolvendo-se probabilidades condicionais por meio do Teorema de Bayes, resultando em uma distribuição condicional de $\boldsymbol{\theta}$ dado $\mathbf{x}$, conhecida como distribuição a posteriori.

1.3.1 Distribuição a Priori

Segundo Jaynes (1980), o desenvolvimento de uma teoria geral para a construção de distribuições a priori que quantifiquem adequadamente a informação a priori sobre o parâmetro de interesse, é um tópico prioritário a ser pesquisado na Inferência Bayesiana.

As distribuições a priori podem ser classificadas de várias maneiras. Uma classificação muito comum é denominá-las de informativas, não-informativas ou empíricas, sendo que vários métodos têm sido propostos na literatura para se obter tais distribuições.

Um método para se obter a distribuição a priori pode ser subjetivo, quando as prioris resultantes incorporam a opinião pessoal do pesquisador, ou de um especialista a respeito dos parâmetros de interesse, opinião esta proveniente de sua experiência ou de sua intuição. Esta distribuição a priori é também denominada priori informativa.

O método pode ser objetivo quando nenhuma ou pouca informação a priori é disponível, sendo então a distribuição a priori obtida denominada como priori não-informativa. O grande número de prioris não-informativas pode causar dificuldades em sua escolha, principalmente quando estas prioris não produzem resultados similares.

Por fim o método empírico é um procedimento de inferência em que a distribuição a priori é obtida utilizando-se os dados iniciais. Esta distribuição a priori é também denominada de distribuição a priori empírica.

1.3.1.1 Priori Uniforme

Uma distribuição a priori, que expressa nenhuma ou pouca informação sobre θ , pode ser obtida a partir de uma priori Uniforme (Bayes, 1763 e Laplace 1774), desta forma não se favorece nenhum particular valor de θ . A distribuição a priori conjunta de θ é dada por

$$\pi_U(\theta) \propto cte . \quad (1.23)$$

1.3.1.2 Priori Gama

Outra distribuição a priori muito utilizada na literatura é construída através do produto de distribuições de probabilidades independentes. Desta forma, é fácil propor distribuições a priori para cada θ_i que sejam independentes entre si e seu produto seja a distribuição a priori conjunta de θ . Para distribuições de probabilidade que possuem parâmetros reais não-negativos, a distribuição a priori conjunta de $\theta = (\theta_1, \dots, \theta_k)$ pode ser obtida através do produto de distribuições gama, resultando em

$$\pi_G(\theta) \propto \prod_{i=1}^k \theta_i^{a_i-1} \exp(-b_i \theta_i), \quad (1.24)$$

sendo a_i e $b_i, i = 1, \dots, k$ hiperparâmetros conhecidos. Vale salientar que, dependendo dos valores destes hiperparâmetros, a distribuição a priori será informativa ou não informativa.

1.3.1.3 Regra de Jeffreys

Jeffreys considerou vários cenários para a construção de distribuições a priori não-informativas (veja Box e Tiao, 1973). Imaginando que o espaço paramétrico fosse

um intervalo limitado ou também um intervalo $(-\infty, \infty)$, Jeffreys argumenta que, nestes casos, a distribuição a priori seja considerada uma constante. Considerando os casos em que o intervalo é $(0, \infty)$, comumente associado ao desvio padrão σ , Jeffreys propôs utilizar como priori $\pi_\sigma = \frac{1}{\sigma}$. A principal justificativa para esta escolha é que ela é invariante sob transformações do parâmetro. Generalizando estes resultados, se $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)$ está contido em intervalos limitados ou $(-\infty, \infty)$ e $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ está contido em intervalos $(0, \infty)$ então a distribuição a priori não-informativa obtida pela regra de Jeffreys é dada por

$$\pi_R(\boldsymbol{\theta}, \boldsymbol{\mu}) \propto \prod_{i=1}^k \frac{1}{\theta_i}. \quad (1.25)$$

1.3.1.4 Priori de Jeffreys

Em um artigo publicado em 1946, Jeffreys propôs uma "regra geral" para a construção de distribuições a priori não-informativas. Esta regra tem sido muito utilizada na literatura, pois é invariante a transformações um a um. Esta distribuição a priori será denominada de priori de Jeffreys, sendo definida a partir da raiz quadrada do determinante da matriz de informação de Fisher $I(\boldsymbol{\theta})$, portanto

$$\pi_J(\boldsymbol{\theta}) \propto \sqrt{\det I(\boldsymbol{\theta})}. \quad (1.26)$$

Em muitos casos há uma grande dificuldade em se calcular (1.21), esta equação pode ser reescrita na seguinte forma

$$I_{ij}(\boldsymbol{\theta}) = E \left[\left(\frac{\partial}{\partial \theta_i \partial \theta_j} \log(f(\mathbf{x}|\boldsymbol{\theta})) \right)^2 \right] = -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(f(\mathbf{x}|\boldsymbol{\theta})) \right], \quad i, j = 1, 2, \dots, k. \quad (1.27)$$

Uma propriedade importante estabelece que para uma amostra aleatória $X_1, \dots, X_n$ de uma v.a. X com f.d.p. $f(\mathbf{x}|\boldsymbol{\theta})$, a informação total de Fisher de $\boldsymbol{\theta}$, correspondente à amostra observada, é a soma das informações de Fisher das n observações da amostra, ou seja, sendo (1.15) a função de verossimilhança tem-se que

$$\begin{aligned} -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(L(\boldsymbol{\theta}; \mathbf{x})) \right] &= -E \left[\sum_{i=1}^n \frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(f(x_i|\boldsymbol{\theta})) \right] \\ &= \sum_{i=1}^n E \left[-\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log(f(x_i|\boldsymbol{\theta})) \right] \\ &= n I_{ij}(\boldsymbol{\theta}), \quad i, j = 1, 2, \dots, k. \end{aligned} \quad (1.28)$$

Exemplo 1.3.1. Seja X uma v.a. com distribuição de probabilidade Gamma(ϕ, μ) a priori de Jeffreys deste modelo é calculada através dos seguintes resultados:

$$\log(f(x|\phi, \mu)) = \phi \log(\mu) + (\phi - 1) \log(x) - \mu x - \log(\Gamma(\phi)). \quad (1.29)$$

Utiliza-se o logaritmo natural da distribuição gama (1.29) para calcular os elementos da matriz de Informação de Fisher, dados a seguir:

$$\begin{aligned} I_{11}(\phi, \mu) &= E \left[-\frac{\partial^2}{\partial \phi^2} \log(f(x|\phi, \mu)) \right] = E[-(-\psi'(\phi))] = \psi'(\phi); \\ I_{12}(\phi, \mu) &= E \left[-\frac{\partial^2}{\partial \phi \partial \mu} \log(f(x|\phi, \mu)) \right] = E \left[-\frac{1}{\mu} \right] = -\frac{1}{\mu}; \\ I_{22}(\phi, \mu) &= E \left[-\frac{\partial^2}{\partial \mu^2} \log(f(x|\phi, \mu)) \right] = E \left[-\left(-\frac{\phi}{\mu^2} \right) \right] = \frac{\phi}{\mu^2}. \end{aligned} \quad (1.30)$$

Utilizando o resultado (1.28) a matriz de Informação de Fisher para $X_1, \dots, X_n$ é dada por

$$I(\phi, \mu) = n \begin{bmatrix} \psi'(\phi) & \frac{1}{\mu} \\ \frac{1}{\mu} & \frac{\phi}{\mu^2} \end{bmatrix}, \quad (1.31)$$

Desta forma utilizando a equação (1.26) temos que a priori de Jeffreys da distribuição Gama (1.5) é

$$\pi_J(\phi, \mu) \propto \frac{\sqrt{\phi \psi'(\phi) - 1}}{\mu}. \quad (1.32)$$

1.3.1.5 Priori de Zellner (MDIP)

A idéia básica da priori de Zellner (1977, 1984), conhecida como "*Maximal Data Information Prior*", é maximizar a informação proveniente dos dados em relação à informação conhecida a priori sobre os parâmetros. Zellner utiliza a medida

$$H_\theta(X) = \int_a^b f(x|\theta) \log(f(x|\theta)) dx. \quad (1.33)$$

Shannon (1948) define $-H(X)$ como entropia negativa da v.a X .

A "Informação Passada" a priori sobre o parâmetro θ , antes da realização do experimento, é dada por

$$H(\theta) = \int_a^b \pi(\theta) \log(\pi(\theta)) d\theta, \quad (1.34)$$

sendo interpretada como a quantidade de informação, sobre o parâmetro θ fornecida pela variável aleatória θ .

Suponha-se que $\theta = \theta_0$ é um valor observado para o parâmetro, então a quantidade de informação fornecida pelos dados é dada por

$$H_{\theta_0}(X) = \int_{R_x} f(x|\theta_0) \log(f(x|\theta_0)) dx. \quad (1.35)$$

A "Informação Presente" é obtida tomando-se a média das medidas $H_\theta(X)$ obtidas para cada valor θ fixado. Obtendo-se em média o quanto de informação a variável aleatória

X (os dados) informa sobre os parâmetros θ , sendo medida por

$$E[H_\theta(X)] = \int_a^b f(x|\theta) \log(f(x|\theta)) d\theta. \quad (1.36)$$

O ganho na informação fornecida pelos dados através do experimento sobre a informação a priori do parâmetro é:

$$G = \int_a^b H_\theta(X) \pi(\theta) d\theta - \int_a^b \pi(\theta) \log(\pi(\theta)) d\theta. \quad (1.37)$$

Deseja-se obter uma priori $\pi(\theta)$ que maximize G , tornando maior possível o ganho na informação fornecida pelos dados sobre a informação a priori sobre o parâmetro. G é maximizada selecionando a $\pi(\theta)$ sujeita a condição $\int_a^b \pi(\theta) d\theta = 1$.

A solução é uma distribuição a priori própria, dada por

$$\pi(\theta) = k \exp(H_\theta(X)), \quad a \leq \theta \leq b, \quad (1.38)$$

sendo $k^{-1} = \int_a^b \exp(H_\theta(X)) d\theta$. uma constante normalizadora.

A priori proposta por Zellner em (1.38) possui propriedades limitadas de invariância, sendo invariante apenas sob transformações lineares de X e θ , e não para toda transformação diferenciável um a um.

Exemplo 1.3.2. Seja X uma v.a. com distribuição de probabilidade Gama(ϕ, μ), a priori de Zellner é calculada através dos seguintes resultados:

Primeiramente deve-se encontrar $H_\theta(X)$ da distribuição gama por meio de

$$H_\theta(X) = \int_0^\infty \log\left(\frac{\mu^\phi}{\Gamma(\phi)} x^{\phi-1} \exp(-\mu x)\right) f(x|\phi, \mu) dx. \quad (1.39)$$

Após algumas manipulações algébricas, obtemos que

$$H_\theta(X) = \phi \log(\mu) + (\phi - 1)E[\log(X)] - \mu E[X] - \log(\Gamma(\phi)). \quad (1.40)$$

É fácil mostrar que $E[\log(X)] = \psi(\phi) - \log(\mu)$ e utilizando o resultado (1.7), obtemos

$$H_\theta(X) = \log(\mu) + (\phi - 1)\psi(\phi) - \phi - \log(\Gamma(\phi)). \quad (1.41)$$

Concluindo, a priori de Zellner (1.38) da distribuição gama (1.5) é dada por

$$\pi_Z(\phi, \mu) \propto \frac{\mu}{\Gamma(\phi)} \exp((\phi - 1)\psi(\phi) - \phi). \quad (1.42)$$

1.3.2 Distribuição a Posteriori

Considerando-se uma amostra aleatória $x_1, \dots, x_n$ de uma distribuição, cuja a densidade é $f(\mathbf{x}|\boldsymbol{\theta})$, a função de verossimilhança $L(\boldsymbol{\theta}|\mathbf{x})$ resume toda informação a respeito de $\boldsymbol{\theta}$, contida nos dados amostrais $\mathbf{x}$. Sendo $\pi(\boldsymbol{\theta})$ uma distribuição a priori conjunta de $\boldsymbol{\theta}$, utiliza-se o Teorema de Bayes para combinar ambas as fontes de informação (distribuição a priori e a função de verossimilhança) encontrando assim a distribuição a posteriori de $\boldsymbol{\theta}$, dado $\mathbf{x}$, denominada $p(\boldsymbol{\theta}|\mathbf{x})$. Esta relação pode ser descrita pelas igualdades a seguir:

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{p(\boldsymbol{\theta}, \mathbf{x})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int p(\boldsymbol{\theta}, \mathbf{x})d\boldsymbol{\theta}} = \frac{L(\boldsymbol{\theta}|\mathbf{x})\pi(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta}|\mathbf{x})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \quad (1.43)$$

Note que $1/p(\mathbf{x})$ não depende de $\boldsymbol{\theta}$, sendo uma constante normalizadora para $p(\boldsymbol{\theta}|\mathbf{x})$. Uma forma muito útil e simplificada para representar $p(\boldsymbol{\theta}|\mathbf{x})$, em problemas que envolvem estimação de parâmetros, é dada por

$$p(\boldsymbol{\theta}|\mathbf{x}) \propto L(\boldsymbol{\theta}|\mathbf{x})\pi(\boldsymbol{\theta}), \quad (1.44)$$

visto que

$$p(\mathbf{x}) = E_{\boldsymbol{\theta}}[p(\mathbf{x}|\boldsymbol{\theta})] = \int L(\boldsymbol{\theta}|\mathbf{x})\pi(\boldsymbol{\theta})d\boldsymbol{\theta} = c^{-1}$$

é apenas uma constante normalizadora. Em outras situações como seleção de modelos, esta constante possui um papel crucial.

A distribuição a posteriori de um parâmetro θ contém toda informação probabilística deste parâmetro. Existem várias formas de resumir esta informação e esta etapa geralmente envolve a avaliação de probabilidades, esperanças de funções particulares do parâmetro θ ou distribuições a posteriori marginais, quando θ for multivariado. Desta forma, o problema geral da inferência Bayesiana consiste em calcular tais valores esperados segundo a distribuição a posteriori de $\boldsymbol{\theta}$.

Vale salientar que a habilidade de integrar funções, muitas vezes complexas e multidimensionais, é de extrema importância na inferência Bayesiana. Soluções exatas somente são possíveis se estas integrais puderem ser calculadas analiticamente, caso contrário, deve-se utilizar aproximações numéricas.

1.3.3 Critérios de Seleção de Modelos

Em problemas que envolvem incerteza, utilizar critérios que auxiliem na comparação e seleção de modelos paramétricos é de fundamental importância em uma análise de dados. Neste tópico serão discutidos três diferentes critérios de comparação de modelos: o AIC, o BIC e o DIC.

Akaike (1974) propõe um método baseado na medida de Informação de Kullback-Leibler. Seja k o número de parâmetros a serem estimados, n o número de observações de $\mathbf{x}$ e $\hat{\boldsymbol{\theta}}$ uma estimativa de $\boldsymbol{\theta}$, o critério de informação de Akaike (AIC) é obtido por

$$AIC = -2\log(L(\hat{\boldsymbol{\theta}}; \mathbf{x})) + 2k . \quad (1.45)$$

Dado um conjunto de modelos candidatos para $\mathbf{x}$, ajustados os dados, o preferido será o que fornecer o menor AIC. Além de selecionar um ótimo ajuste, o critério penaliza a adição de parâmetros, desencorajando *overfitting*, ou seja, a seleção de um modelo extremamente complexo e com muitos parâmetros que tenham um pobre desempenho preditivo.

Outro critério muito utilizado é o Critério de Informação Bayesiana (BIC), desenvolvido por Schwarz (1978). Este critério possui estrutura semelhante ao AIC, mas é adotado sobre o paradigma Bayesiano. O BIC é obtido através de resultados assintóticos e da suposição de que os dados pertencem à família exponencial, a equação é dada por

$$BIC = -2\log(L(\hat{\boldsymbol{\theta}}; \mathbf{x})) + k \log(n) . \quad (1.46)$$

Semelhante ao AIC, o critério penaliza a adição de parâmetros. Dos modelos candidatos, o adequado será o que possuir o menor BIC. Vale salientar que, utilizando-se a equação (1.46), a contribuição da distribuição a priori não é considerada.

O Critério de Informação da Deviance (DIC) é uma generalização dos critérios AIC e BIC. Proposto por Spiegelhalter et al. (2002), este critério é de grande utilidade em problemas de seleção de modelos Bayesianos nos quais, amostras da distribuição a posteriori dos parâmetros de interesse são obtidas utilizando-se os métodos MCMC.

O deviance é definido por

$$D(\boldsymbol{\theta}) = -2\log(L(\boldsymbol{\theta}; \mathbf{x})) + C, \quad (1.47)$$

sendo C uma constante que anula-se durante os cálculos e portanto não necessita ser conhecida para comparação de modelos. O DIC é obtido por

$$DIC = \bar{D} + p_D, \quad (1.48)$$

em que $\bar{D} = E_{\boldsymbol{\theta}}[D(\boldsymbol{\theta})]$ é o deviance a posteriori que mede a qualidade do ajuste, $p_D = \bar{D} - D(\hat{\boldsymbol{\theta}})$ é o número efetivo de parâmetros do modelo e $D(\hat{\boldsymbol{\theta}})$ é a deviance obtida utilizando-se a média a posteriori de $\boldsymbol{\theta}$.

Semelhante aos métodos AIC e BIC, dos modelos candidatos a $\mathbf{x}$, o melhor será o que possuir o menor DIC.

1.4 Métodos de Monte Carlo via Cadeias de Markov

O avanço recente no uso de métodos bayesianos em diversas áreas pode ser atribuído, em grande parte, ao avanço dos Métodos de Monte Carlo via Cadeias de Markov (MCMC).

Suponha que o objetivo seja gerar uma amostra de uma distribuição a posteriori não regular $p(\boldsymbol{\theta}|\mathbf{x})$, $\boldsymbol{\theta} \in \mathbb{R}^k$, porém não pode-se fazê-lo diretamente. Entretanto, é possível construir uma cadeia de Markov homogênea, em que a distribuição de equilíbrio convirja para a distribuição de interesse $p(\boldsymbol{\theta}|\mathbf{x})$. Após obter muitos valores simulados desta cadeia, podemos utilizá-los como base para sumarizar as características de interesse de $p(\boldsymbol{\theta}|\mathbf{x})$.

Existem vários métodos para se construir estas cadeias de Markov, cuja distribuição de interesse é a distribuição a posteriori $p(\boldsymbol{\theta}|\mathbf{x})$. Serão discutidos dois métodos: o Amostrador de Gibbs, proposto por Geman e Geman (1984), sendo inicialmente aplicado com sucesso na inferência Bayesiana por Gelfand e Smith (1990); além do algoritmo de Metropolis-Hastings, originalmente desenvolvido por Metropolis et al., (1953) e generalizado por Hastings (1970).

1.4.1 Amostrador de Gibbs

Genericamente conhecido como amostrador de Gibbs, este método tem desempenhado papel essencial na inferência Bayesiana. Desenvolvido inicialmente por Geman e Geman (1984) para modelar problemas de processamento de imagens, este método tem como característica principal gerar valores de uma distribuição, sem a necessidade de empregar diretamente sua função densidade de probabilidade.

Gelfand e Smith (1990) foram os primeiros a mostrar com sucesso que o método proposto por Geman e Geman (1984) poderia ser aplicado para diferentes distribuições a posteriori. Desde então, este método tem sido muito utilizado para resolver vários problemas na inferência Bayesiana.

O amostrador de Gibbs é um método MCMC, onde o núcleo de transição é formado pelas distribuições condicionais completas. Considere que a distribuição conjunta dos parâmetros de interesse é $p(\boldsymbol{\theta}|\mathbf{x})$, sendo $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$ e que as distribuições condicionais completas $p_i(\theta_i|\theta_1, \dots, \theta_{(i-1)}, \theta_{(i+1)}, \dots, \theta_k, \mathbf{x})$, para $i = 1, \dots, k$ são conhecidas e podem ser amostradas facilmente.

Desta forma, quando não se consegue trabalhar diretamente com $p(\theta_1, \theta_2, \dots, \theta_k|\mathbf{x})$, mas é possível gerar valores das condicionais completas p_i , o amostrador de Gibbs proporciona um ótimo esquema de amostragem da distribuição conjunta, via amostragem iterativa das distribuições condicionais completas.

O método é descrito pelos passos a seguir:

1. Defina os valores iniciais $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})'$ e inicie o contador $j = 1$;
2. Obtenha novo valor de $\theta^{(j)} = (\theta_1^{(j)}, \dots, \theta_k^{(j)})'$ através da geração sucessiva dos valores

$$\begin{aligned}\theta_1^{(j)} &\sim p_1(\theta_1 | \theta_2^{(j-1)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)}, \mathbf{x}), \\ \theta_2^{(j)} &\sim p_2(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)}, \mathbf{x}), \\ &\vdots \\ \theta_k^{(j)} &\sim p_k(\theta_k | \theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_{k-1}^{(j)}, \mathbf{x}),\end{aligned}\tag{1.49}$$

3. Mude o contador j para $j + 1$, volte ao passo 2 e repita este procedimento n vezes;

Conforme o número de iterações aumenta, a cadeia $\theta^{(j)}$ se aproxima da distribuição de equilíbrio e assume-se convergência (aproximada). Sob certas condições de regularidade (veja Geman e Geman, 1984), $\theta_i^{(j)}$ pode ser considerada (aproximadamente) uma observação simulada da distribuição $p_i(\theta_i)$, $i = 1, \dots, k$. Desta forma, repetindo-se o procedimento n vezes, e obtida a convergência, os valores $\theta^{(j)}$, $j = 1, \dots, n$ resultantes, formam uma amostra de tamanho n da distribuição a posteriori conjunta.

1.4.2 Metropolis Hastings

Quando as distribuições condicionais não são identificadas facilmente, possuindo uma forma padrão (Normal, Gama, etc), a incapacidade de gerar amostras diretamente destas distribuições limita o uso do amostrador de Gibbs. Como alternativa pode-se utilizar o algoritmo de Metropolis-Hastings.

O trabalho original proposto por Metropolis et al., (1953) teve como objetivo principal calcular propriedades de algumas substâncias químicas. Passado algum tempo este trabalho, em si, teve um grande impacto na estatística Bayesiana. Hastings (1970) propôs uma generalização do algoritmo proposto por Metropolis, ficando conhecido como algoritmo Metropolis-Hastings.

Suponha que desejamos gerar valores de uma densidade a posteriori não regular $p(\theta | \mathbf{x})$. Defina um valor inicial $\theta^{(t)}$, gere $\theta^{(*)}$ de uma distribuição auxiliar (ou distribuição proposta) $q(\theta^{(t)}, \theta^{(*)})$. O novo valor $\theta^{(*)}$ é então aceito com probabilidade

$$\alpha(\theta^{(t)}, \theta^{(*)}) = \min \left(1, \frac{p(\theta^{(*)} | \mathbf{x}) q(\theta^{(t)} | \theta^{(*)})}{p(\theta^{(t)} | \mathbf{x}) q(\theta^{(*)} | \theta^{(t)})} \right).\tag{1.50}$$

Se $\alpha(\boldsymbol{\theta}^{(t)}, \boldsymbol{\theta}^{(*)}) \geq u(0, 1)$ então $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(*)}$, caso contrário $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)}$.

Este mecanismo de aceitação garante a convergência da cadeia para a distribuição de equilíbrio.

O algoritmo de Metropolis-Hastings pode ser especificado pelos seguintes passos:

1. Especifique um valor inicial $\boldsymbol{\theta}^{(0)}$ e defina o contador da cadeia $t = 0$;
2. Gere um novo valor $\boldsymbol{\theta}^{(*)}$ da densidade proposta $q(\boldsymbol{\theta}^{(t)}, \boldsymbol{\theta}^{(*)})$;
3. Calcule a probabilidade de aceitação $\alpha(\boldsymbol{\theta}^{(t)}, \boldsymbol{\theta}^{(*)})$ e amostre u de uma Uniforme(0, 1).
4. Se $\alpha(\boldsymbol{\theta}^{(t)}, \boldsymbol{\theta}^{(*)}) \geq u(0, 1)$ então $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(*)}$, caso contrário $\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)}$.
5. Altere o contador de t para $t + 1$ e volte ao passo 2 até a convergência ser alcançada.

Para t suficientemente grande, $\boldsymbol{\theta}^{(t)}$ pode ser considerada uma amostra simulada da distribuição a posteriori $p(\boldsymbol{\theta}|\mathbf{x})$.

Vale ressaltar que a distribuição proposta $q(\boldsymbol{\theta}, \cdot)$ deve ser escolhida com cautela, embora sua escolha possa ser arbitrária, alguns cuidados devem ser tomados para garantir a eficiência do algoritmo.

1.4.3 Algoritmo Híbrido

Na seção 1.4.2 foi apresentado o algoritmo de Metropolis-Hastings. Uma de suas particularidades é que a atualização da quantidade de interesse $\boldsymbol{\theta}$ é obtida em um único passo. Hastings (1970) e Tierney (1994) discutem a possibilidade de atualizar os componentes $\theta_1, \theta_2, \dots, \theta_k$ separadamente, ou seja, a cada iteração, um único componente de $\boldsymbol{\theta}$ é atualizado e a escolha da atualização é feita por uma ordem fixada dos k componentes. Por exemplo, os componentes podem ser atualizados seguindo $\theta_1 \rightarrow \theta_2 \rightarrow \dots \rightarrow \theta_k$, similarmente à equação (1.49).

A distribuição proposta agora passa a ser escrita como $q_i(\theta_i^{(t)}|\theta_i^{(*)})$ e a probabilidade de aceitação (1.50) torna-se calculável por

$$\alpha_i(\theta_i^{(t)}, \theta_i^{(*)}) = \min \left(1, \frac{p_i(\theta_i^{(*)}|\theta_{-i}^{(t)}, \mathbf{x}) q_i(\theta_i^{(t)}|\theta_i^{(*)})}{p_i(\theta_i^{(t)}|\theta_{-i}^{(t)}, \mathbf{x}) q_i(\theta_i^{(*)}|\theta_i^{(t)})} \right). \quad (1.51)$$

sendo $\theta_{-i}^{(t)} = (\theta_1^{(t+1)}, \dots, \theta_{i-1}^{(t+1)}, \theta_{i+1}^{(t)}, \dots, \theta_k^{(t)})$. Se $\alpha_i(\theta_i^{(t)}, \theta_i^{(*)}) \geq u(0, 1)$ então $\theta_i^{(t+1)} = \theta_i^{(*)}$, caso contrário $\theta_i^{(t+1)} = \theta_i^{(t)}$.

O algoritmo de Metropolis-Hastings híbrido pode ser especificado pelos seguintes passos:

1. Defina os valores iniciais $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})'$ e inicie o contador $t = 0$;
2. Inicie o contador $i = 1$;
3. Gere um novo valor $\theta_i^{(*)}$ da densidade proposta $q_i(\theta_i^{(t)}, \theta_i^{(*)})$;
4. Calcule a probabilidade de aceitação $\alpha_i(\theta_i^{(t)}, \theta_i^{(*)})$ e amostre u de uma $\text{Uniforme}(0, 1)$.
5. Se $\alpha_i(\theta_i^{(t)}, \theta_i^{(*)}) \geq u(0, 1)$ então $\theta_i^{(t+1)} = \theta_i^{(*)}$, caso contrário $\theta_i^{(t+1)} = \theta_i^{(t)}$.
6. Altere o contador de i para $i + 1$ e retorne ao passo 3 até que $i = k$. Se $i = k$, vá ao passo 7;
7. Altere o contador de t para $t + 1$ e volte ao passo 2, até a convergência ser alcançada.

De fato este foi o algoritmo apresentado originalmente no trabalho de Metropolis et al. (1957), onde a posição das moléculas eram atualizadas uma a uma. Embora Hastings (1970) discuta a possibilidade de atualizar cada componente separadamente, ao unificar seu trabalho com o de Metropolis, o mesmo introduziu o algoritmo com um único núcleo de transição.

Até aqui discutimos o caso em que a distribuição auxiliar é dada por $q(\theta^{(t)}, \theta^{(*)})$, sendo que $\theta^{(t)} = \theta^{(t-1)} + w_j$, onde w_j é uma variável aleatória com distribuição f_w independente da cadeia. Esta evolução de $\theta^{(t)}$ é também conhecida como passeio aleatório.

No entanto é possível utilizar uma distribuição auxiliar em que o valor proposto $\theta^{(*)}$ seja independente da posição anterior de θ na cadeia, desta forma $q(\theta^{(t)}, \theta^{(*)}) = f(\theta^{(*)})$. A probabilidade de aceitação pode ser escrita então na seguinte forma:

$$\alpha_i(\theta_i^{(t)}, \theta_i^{(*)}) = \min \left(1, \frac{p_i(\theta_i^{(*)} | \theta_{-i}^{(t)}, x) f_i(\theta_i^{(t)})}{p_i(\theta_i^{(t)} | \theta_{-i}^{(t)}, x) f_i(\theta_i^{(*)})} \right). \quad (1.52)$$

Neste caso os passos realizados serão os mesmos do algoritmo de Metropolis-Hastings híbrido.

Este resultado não invalida as propriedades das cadeias de Markov, pois $f(\theta^{(*)})$ é apenas a distribuição proposta que, combinada com a taxa de aceitação $\alpha(\theta^{(t)}, \theta^{(*)})$, proporciona o núcleo de transição do algoritmo. Este núcleo depende do estado anterior da cadeia e preserva a estrutura de uma cadeia de Markov (veja Gamerman e Lopes, 2006).

1.4.4 Diagnósticos de Convergência

A utilização de algoritmos que envolvem um processo estocástico, tem como foco prioritário a necessidade em verificar se a convergência para a distribuição estacionária é

obtida, ou seja, se os valores simulados podem ser considerados como uma amostra da distribuição de interesse.

Uma abordagem empírica para controlar a convergência consiste em desenhar gráficos de série temporal para os valores simulados da cadeia, procurando detectar comportamentos de não-estacionariedade.

Uma cadeia obtém uma possível convergência para a distribuição de interesse se os valores simulados mudarem rapidamente dos valores anteriores e do valor inicial. Desta forma, se a cadeia não convergir, ela ficará muito perto ou no mesmo valor, durante muitas iterações.

Ao utilizar o algoritmo de Metropolis-Hastings, Hastings (1970) sugere calcular a quantidade de valores aceitos, pois o sucesso do método depende de taxas de aceitação não muito baixas. Bennett et al. (1995), Brooks et al. (2011), entre outros autores, indicam que as taxas de aceitação devem estar aproximadamente no intervalo de 10% a 60%.

Geweke (1992) propõe um método formal para verificar a convergência das cadeias (ou sequências). O teste é baseado nas médias inicial e final de uma cadeia. Se as amostras são obtidas de uma distribuição estacionária, espera-se que as duas médias sejam similares. Para uma sequência de tamanho n , já removido a amostra de aquecimento, as médias iniciais e finais da cadeia são obtidas calculando-se:

$$v_b = \frac{1}{n_b} \sum_{j=1}^{n_b} v^{(j)} \quad e \quad v_a = \frac{1}{n_a} \sum_{j=n-n_a+1}^n v^{(j)},$$

onde $n_b + n_a < n$. Conforme n cresce n_a/n e n_b/n permanecem fixas e a estatística de Geweke é dada por

$$z_G = \frac{v_a - v_b}{\sqrt{\hat{V}ar(v_a) + \hat{V}ar(v_b)}} \xrightarrow{d} N(0, 1) \quad (1.53)$$

A estatística z_G não deve possuir grandes valores, caso a convergência seja alcançada. Portanto, grandes valores indicam falta de convergência, mas valores pequenos não implicam em convergência. Geweke (1992) sugere que sejam utilizados os valores $n_b = 0.1n$ e $n_a = 0.5n$. A principal vantagem deste método é que ele requer apenas o resultado de uma cadeia para cada parâmetro de interesse e pode ser usado em qualquer esquema MCMC.

Há outros métodos de convergência muito utilizados na literatura, dentre eles podemos citar: Gelman e Rubin (1992), Raftery e Lewis (1992) e Heidelberger (1983). Estes métodos estão implementados no R e podem ser acessados utilizando-se a biblioteca CODA (*Convergence Diagnosis and Output Analysis Software for Gibbs sampling output*) (Best et al., 1997). Em geral todos os métodos discutidos até aqui devem ser utilizados conjuntamente a fim de obtermos melhores indicações de convergência. Outros métodos podem ser vistos com mais detalhes em Gamerman e Lopes (2006).

2 Distribuição Gama Generalizada

A distribuição Gama Generalizada tem se mostrado um ótimo modelo que leva a excelentes inferências para dados de sobrevivência. Introduzida por Stacy (1962), uma v.a T terá distribuição Gama Generalizada (GG) se sua função densidade é dada por

$$f(t|\boldsymbol{\theta}) = \frac{\alpha}{\Gamma(\phi)} \mu^{\alpha\phi} t^{\alpha\phi-1} \exp(-(\mu t)^\alpha), \quad (2.1)$$

sendo $t > 0$, $\boldsymbol{\theta} = (\phi, \mu, \alpha)$. Os parâmetros de forma são $\alpha > 0$ e $\phi > 0$ e $\mu > 0$ o parâmetro de escala.

A função acumulada é dada por

$$F(t|\boldsymbol{\theta}) = \int_0^{(\mu t)^\alpha} \frac{1}{\Gamma(\phi)} w^{\phi-1} e^{-w} dw = \frac{\gamma[\phi, (\mu t)^\alpha]}{\Gamma(\phi)}, \quad (2.2)$$

sendo $\gamma[y, x] = \int_0^x w^{y-1} e^{-w} dw$ denominada de gama incompleta inferior.

Distribuições relevantes podem ser obtidas a partir da distribuição GG, como a distribuição Weibull (quando $\phi = 1$), a distribuição Gama ($\alpha = 1$), Log-Normal (caso limite quando $\phi \rightarrow \infty$) e a distribuição Normal Generalizada ($\alpha = 2$). A distribuição Normal generalizada é também uma distribuição que inclui várias distribuições conhecidas como, half-normal ($\phi = 1/2, \mu = 1/\sqrt{2}\sigma$), Rayleigh ($\phi = 1, \mu = 1/\sqrt{2}\sigma$), Maxwell-Boltzmann ($\phi = 3/2$) e chi ($\phi = k/2, k = 1, 2, \dots$).

A Figura 9 dá exemplos das formas das funções densidade de probabilidade e acumulada para diferentes valores de ϕ , μ e α .

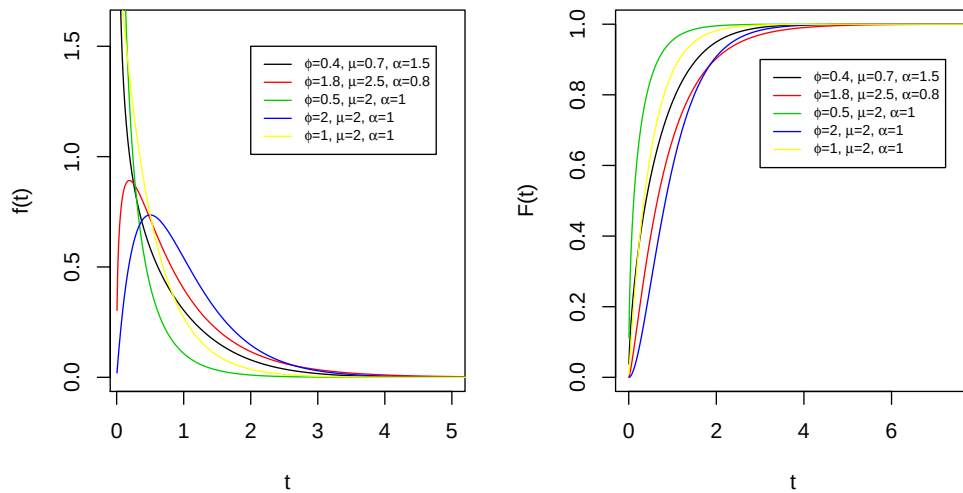


Figura 9 – Funções densidade de probabilidade e acumulada da distribuição Gama Generalizada para diferentes valores de ϕ , μ e α .

2.1 Média, Moda e Variância

A média, a variância e a moda da distribuição GG, são dadas respectivamente, por:

$$E(X) = \frac{\Gamma\left(\phi + \frac{1}{\alpha}\right)}{\mu\Gamma(\phi)}, \quad (2.3)$$

$$V(X) = \frac{1}{\mu^2} \left\{ \frac{\Gamma\left(\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)} - \left(\frac{\Gamma\left(\phi + \frac{1}{\alpha}\right)}{\Gamma(\phi)} \right)^2 \right\}, \quad (2.4)$$

$$Mo(X) = \frac{1}{\mu} \left(\phi - \frac{1}{\alpha} \right)^{\frac{1}{\alpha}}, \quad \phi > \frac{1}{\alpha}. \quad (2.5)$$

2.2 Função de Sobrevivência e de Risco

A função de sobrevivência de uma v.a. $T \sim \text{GG}(\phi, \mu, \alpha)$, com a probabilidade de uma observação não falhar até certo tempo t é

$$S(t|\boldsymbol{\theta}) = P[T > t] = \int_{(\mu t)^\alpha}^{\infty} \frac{1}{\Gamma(\phi)} w^{\phi-1} e^{-w} dw = \frac{\Gamma[\phi, (\mu t)^\alpha]}{\Gamma(\phi)}, \quad (2.6)$$

sendo $\Gamma[y, x] = \int_x^{\infty} w^{y-1} e^{-w} dw$ denominada de gama incompleta superior.

A função de risco da v.a. T é dada por

$$h(t|\boldsymbol{\theta}) = \frac{f(t|\boldsymbol{\theta})}{S(t|\boldsymbol{\theta})} = \frac{\alpha \mu^{\alpha\phi} t^{\alpha\phi-1} \exp(-(\mu t)^\alpha)}{\Gamma[\phi, (\mu t)^\alpha]}. \quad (2.7)$$

Este modelo acomoda várias formas de risco, sendo elas: constante, crescente, decrescente, forma de banheira e banheira inversa. Estas propriedades fazem com que a GG seja muito flexível para modelar dados de confiabilidade (Glaser, 1980). Um método simples para se verificar o comportamento da função de risco é obtido baseando-se nos valores de α e ϕ (Gi e Almhana, 1995). Quando $\alpha = 1$, obtemos a distribuição Gama. Assim a função de risco (2.7) passa a assumir as seguintes formas:

- Crescente: para $0 < \phi < 1$;
- Constante: para $\phi = 1$;
- Decrescente: para $\phi > 1$;

Considerando $\alpha \neq 1$, se existir uma solução estritamente positiva para $(1 - \alpha\phi)/[\alpha(\alpha - 1)]$, a função de risco possuirá:

- Forma de banheira: para $\phi > 1$;
- Forma de banheira inversa: para $0 < \phi < 1$;

A Figura 10 dá exemplos das formas das funções de sobrevivência e de risco para diferentes valores de ϕ, μ e α .

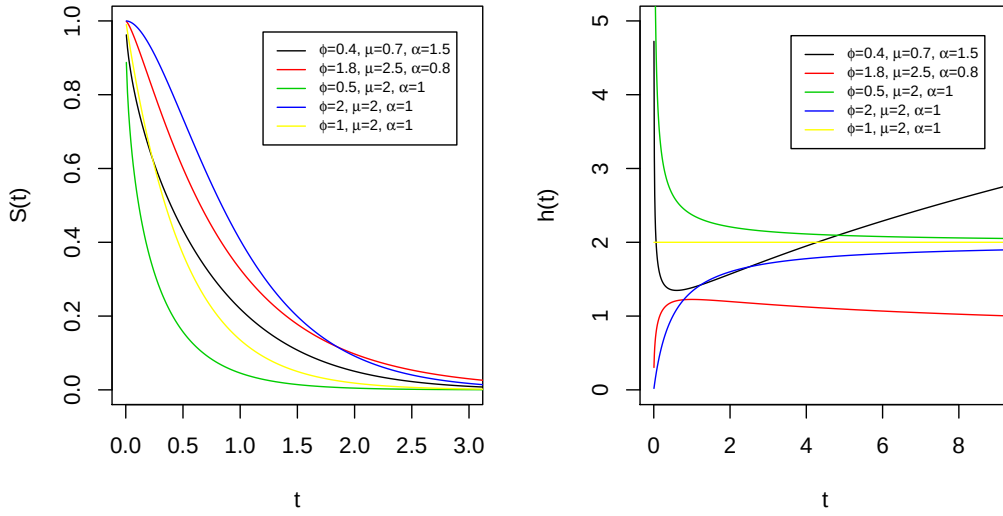


Figura 10 – Funções de sobrevivência e de risco da distribuição Gama Generalizada para diferentes valores de ϕ, μ e α .

2.3 Outros Resultados Importantes

Uma propriedade importante da distribuição GG estabelece que esta é fechada sobre transformação de potência. Desta forma, se $X \sim \text{GG}(\phi, \mu, \alpha)$ então

$$Y = X^s \sim \text{GG}\left(\phi, \mu^s, \frac{\alpha}{s}\right), \quad s > 0. \quad (2.8)$$

Um caso particular seria, $Y = X^\alpha \sim \text{Gama}(\phi, \mu^\alpha)$. Outra propriedade muito útil é que se $X \sim \text{GG}(\phi, \mu, \alpha)$ e para uma constante $c > 0$, sendo $W = cX$, então

$$W \sim \text{GG}(\phi, c\mu, \alpha). \quad (2.9)$$

Alguns momentos da distribuição GG desempenham um papel essencial na inferência estatística e na teoria de informação, dentre eles podemos citar:

$$E[T^s] = \frac{\Gamma\left[\frac{s}{a} + \phi\right]}{\mu^s \Gamma[\phi]}, \quad (2.10)$$

$$E[\log(T^s)] = s \log\left(\frac{1}{\mu}\right) + \frac{s}{\alpha} \psi(\phi). \quad (2.11)$$

A matriz de informação de Fisher para a distribuição GG é dada por

$$I(\alpha, \mu, \phi) = \begin{bmatrix} \frac{1 + 2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2}{\alpha^2} & -\frac{1 + \phi\psi(\phi)}{\mu} & -\frac{\psi(\phi)}{\alpha} \\ -\frac{1 + \phi\psi(\phi)}{\mu} & \frac{\phi\alpha^2}{\mu^2} & \frac{\alpha}{\mu} \\ -\frac{\psi(\phi)}{\alpha} & \frac{\alpha}{\mu} & \psi'(\phi) \end{bmatrix}. \quad (2.12)$$

2.4 Métodos para Gerar Valores da Distribuição Gama Generalizada

Nesta seção, serão discutidos métodos para gerar valores aleatórios da distribuição Gama Generalizada, considerando dados completos e com censura do tipo aleatória.

2.4.1 Observações Completas

Um método simples para gerar valores da distribuição GG é baseado no Teorema fundamental da transformação de probabilidades (Ross, 2006). Seja U uma v.a. com distribuição $U(0, 1)$ e $X \sim \text{GG}(\phi, \mu, \alpha)$, sendo $F(x)$ a f.d.a., então valores de x podem ser obtidos através de $X = F^{-1}(U)$, sendo $F^{-1}(x)$ a função inversa de $F(x)$. No entanto, a equação (2.2) depende de uma função gama incompleta, não conseguindo-se encontrar uma forma fechada para $X = F^{-1}(U)$.

Métodos numéricos podem ser utilizados para resolver a seguinte equação não linear:

$$U - F(x) = 0. \quad (2.13)$$

A solução desta equação será considerada uma observação de X .

No entanto, o tempo computacional gasto para resolver a equação (2.13) é alto, tornando outros métodos preferíveis para gerar-se observação da distribuição GG. Uma alternativa é considerar que $Y \sim \text{Gamma}(\phi, \mu^\alpha)$. Realizando a transformação $X = Y^{(1/\alpha)}$, pela propriedade de transformação de potencia (2.8), obtemos que $X \sim \text{GG}(\phi, \mu, \alpha)$.

Como há várias formas de gerar valores de uma distribuição gama (veja, por exemplo, Ahrens e Dieter, 1974, 1982) e diversos *softwares* possuem estes procedimentos implementados, obteremos valores da distribuição GG, realizando apenas uma transformação nos valores da distribuição gama gerados pelo *software*.

2.4.2 Observações Censuradas

Em análise de confiabilidade é comum se deparar com problemas em que uma porcentagem das observações é parcialmente conhecida. Nestes casos, dizemos que as observações são censuradas.

Alguns métodos têm sido propostos para se gerar observações com diferentes tipos de censura. Philippe (1997) descreve um procedimento para se gerar valores com censura tipo I, utilizando a distribuição gama truncada. Balakrishnan e Sandhu (1995) desenvolveram um procedimento para se gerar observações com censura progressiva tipo II. Gentle (2003) descreve algumas formas de se gerar valores com diferentes tipos de censura.

Observações com censura do tipo aleatória ocorrem com muita frequência. Podemos gerá-las através dos seguintes passos:

1. Defina o tamanho da amostra n e o número de observações censuradas n_c . A porcentagem de censura passa a ser calculada por n_c/n ;
2. Gere uma amostra da v.a. $X \sim \text{GG}(\phi, \mu, \alpha)$ de tamanho n , utilizando o método descrito na seção 2.4.1;
3. Selecione, de forma aleatória e uniforme, n_c valores distintos do conjunto $A = 1, 2, \dots, n$ e armazene em um vetor ω ;
4. Sejam x_{ω_j} , $j = 1, \dots, n_c$ os valores que serão censurados. Sabe-se que x_{ω_j} é uma observação completa, mas que, em algum momento no tempo, deveria ser censurada. Será denominado que T_{ω_j} é o tempo de censura do indivíduo ω_j . Como x_{ω_j} é uma observação completa, então $t_{\omega_j} \leq x_{\omega_j}$. Supondo que a informação x_{ω_j} foi perdida de forma uniforme ao longo do tempo, então t_{ω_j} pode ser obtido através de $T_{\omega_j} \sim \text{Uniforme}(0, X_{\omega_j})$.
5. Desta forma os valores censurados serão dados por

$$X_{\omega_j} = \min(t_{\omega_j}, x_{\omega_j}) \text{ e } \delta_{\omega_j} = 0, \quad j = 1, \dots, n_c.$$

No final teremos então uma amostra $\mathcal{L} = (x_i, \delta_i)$, em que $X \sim \text{GG}(\phi, \mu, \alpha)$ e δ é variável que indica a presença ou ausência de censura.

Exemplo 2.4.1. Uma amostra de tamanho $n = 30$ será gerada no R com parâmetros $(\phi, \mu, \alpha) = (2, 1, 0.5)$ e 20% de censura. Para isto iremos utilizar a amostra completa dada pela tabela a seguir:

6.165	0.488	1.282	3.845	7.854	6.769	5.637	4.469	0.203	34.606
26.745	9.492	3.907	32.447	6.342	14.242	0.815	1.141	54.017	9.450
11.021	1.613	17.682	9.153	1.599	6.436	51.485	4.019	3.393	0.479

Devemos selecionar $n_c = 6$ elementos que serão censurados. Ao utilizar o comando "sample" no R , obtemos de forma aleatória e uniforme n_c valores distintos, dados por $\omega = (7, 12, 20, 29, 27, 11)$. Desta forma, os valores censurados serão $X_7 = 5.637$, $X_{11} = 26.745$, $X_{12} = 9.492$, $X_{20} = 9.450$, $X_{27} = 6.436$ e $X_{29} = 3.393$.

Utilizando o passo 4 do esquema acima, os tempos de censura obtidos através de $T_{\omega_j} \sim \text{Uniforme}(0, X_{\omega_j})$ serão dados por $T_7 = 3.047$, $T_{11} = 5.523$, $T_{12} = 0.147$, $T_{20} = 0.361$, $T_{27} = 2.094$ e $T_{29} = 1.736$.

Os valores censurados serão obtidos através de $X_{\omega_j} = \min(t_{\omega_j}, x_{\omega_j})$, em que $j = 1, 2, \dots, 6$. Portanto a amostra contendo 20% de valores com censura (*) do tipo aleatória é dada por:

6.165	0.488	1.282	3.845	7.854	6.769	*3.047	4.469	0.203	34.606
*5.523	*0.147	3.907	32.447	6.342	14.242	0.815	1.141	54.017	*0.361
11.021	1.613	17.682	9.153	1.599	6.436	*2.094	4.019	*1.736	0.479

3 Análise Clássica da Distribuição Gama Generalizada

É fato que a distribuição Gama Generalizada é um ótimo modelo probabilístico, que leva a excelentes ajustes em dados reais. No entanto, em muitas aplicações os EMV's da distribuição GG, não apresentam o comportamento estável (Stacy e Mihram, 1965; Parr e Webster, 1965; Hager e Bain, 1970; Wingo, 1987; Kumar e Shukla, 2010) e os resultados obtidos, podem depender dos valores iniciais utilizados nos métodos computacionais iterativos. Desta forma, é de grande importância o desenvolvimento de técnicas exploratórias para se obter bons valores iniciais, visando-se alcançar resultados mais precisos ao se utilizar métodos computacionais iterativos.

Neste capítulo, apresentaremos os Estimadores de Máxima Verossimilhança da distribuição Gama Generalizada (2.1), considerando dados completos e com censura tipo I.

3.1 Estimador de Máxima Verossimilhança para Dados Completos

No método de máxima verossimilhança os estimadores são obtidos a partir, da maximização da função de verossimilhança.

Seja $T_1, \dots, T_n$ uma amostra aleatória tal que $T \sim \text{GG}(\alpha, \mu, \phi)$. Neste caso, a função de verossimilhança de (2.1) é

$$L(\boldsymbol{\theta}; \mathbf{t}) = \frac{\alpha^n}{\Gamma(\phi)^n} \mu^{n\alpha\phi} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\}. \quad (3.1)$$

Estabelecendo $\frac{\partial}{\partial \alpha} \log(L(\boldsymbol{\theta}; \mathbf{t}))$, $\frac{\partial}{\partial \mu} \log(L(\boldsymbol{\theta}; \mathbf{t}))$ e $\frac{\partial}{\partial \phi} \log(L(\boldsymbol{\theta}; \mathbf{t}))$ iguais a 0, obtemos as equações de verossimilhança dadas por

$$n \Gamma'(\hat{\phi}) = n \hat{\alpha} \log(\hat{\mu}) + \hat{\alpha} \sum_{i=1}^n \log(t_i) \quad (3.2)$$

$$n \hat{\phi} = \hat{\mu}^{\hat{\alpha}} \sum_{i=1}^n t_i^{\hat{\alpha}} \quad (3.3)$$

$$\frac{n}{\hat{\alpha}} + n \hat{\phi} \log(\mu) + \phi \sum_{i=1}^n \log(t_i) = \hat{\mu}^{\hat{\alpha}} \sum_{i=1}^n t_i^{\hat{\alpha}} \log(\hat{\mu} t_i), \quad (3.4)$$

cujas soluções fornecem os estimadores de máxima verossimilhança (veja, por exemplo, Stacy e Mihram, 1965, Hager e Bain, 1970). Métodos numéricos como o de Newton-Raphson são necessários para encontrar-se a solução deste sistema não linear.

Algumas manipulações algébricas podem ser feitas nas equações de verossimilhança, buscando-se simplificar a solução do sistema. A equação (3.4) pode ser rescrita através de

$$\frac{n}{\hat{\alpha}} + \phi \sum_{i=1}^n \log(t_i) = \hat{\mu}^{\hat{\alpha}} \sum_{i=1}^n t_i^{\hat{\alpha}} \log(t_i), \quad (3.5)$$

Utilizando a expressão (3.3), após algumas manipulações algébricas, temos que

$$\hat{\phi} = \hat{\mu}^{\hat{\alpha}} \frac{\sum_{i=1}^n t_i^{\hat{\alpha}}}{n} \quad (3.6)$$

Substituindo (3.6) em (3.5) e após algumas manipulações algébrica temos que

$$\hat{\mu} = \left(\frac{1}{\hat{\alpha}} \frac{n}{\sum_{i=1}^n t_i^{\hat{\alpha}} \log(t_i) - \frac{\sum_{i=1}^n t_i^{\hat{\alpha}}}{n} \sum_{i=1}^n \log(t_i)} \right)^{\frac{1}{\alpha}} \quad (3.7)$$

Substituindo (3.7) em (3.6) temos que

$$\hat{\phi} = \left(\frac{1}{\hat{\alpha}} \frac{\sum_{i=1}^n t_i^{\hat{\alpha}}}{\sum_{i=1}^n t_i^{\hat{\alpha}} \log(t_i) - \frac{\sum_{i=1}^n t_i^{\hat{\alpha}}}{n} \sum_{i=1}^n \log(t_i)} \right) \quad (3.8)$$

Desta forma utilizando os resultados (3.7) e (3.8) na equação (3.2) temos que

$$\begin{aligned} h(\alpha) = & \psi \left(\frac{1}{\hat{\alpha}} \frac{\sum_{i=1}^n t_i^{\hat{\alpha}}}{\sum_{i=1}^n t_i^{\hat{\alpha}} \log(t_i) - \frac{\sum_{i=1}^n t_i^{\hat{\alpha}}}{n} \sum_{i=1}^n \log(t_i)} \right) - \frac{\hat{\alpha} \sum_{i=1}^n \log(t_i)}{n} \\ & - \log \left(\frac{1}{\hat{\alpha}} \frac{n}{\sum_{i=1}^n t_i^{\hat{\alpha}} \log(t_i) - \frac{\sum_{i=1}^n t_i^{\hat{\alpha}}}{n} \sum_{i=1}^n \log(t_i)} \right) \end{aligned} \quad (3.9)$$

Estes novos resultados são de grande interesse prático, pois resolvendo a equação (3.9), encontramos facilmente o EMV de $\hat{\alpha}$. Com este resultado encontramos respectivamente os EMV's de ϕ e μ .

Outros métodos podem ser utilizados para se realizar inferências nos parâmetros da distribuição GG: Huang e Hwang (2006) utilizam o método dos momentos para obter inferências da distribuição GG. Khodabin e Ahmadabadi (2010) comparam este método com o método da máxima verossimilhança e concluem que, em geral, os EMV's possuem melhor desempenho, principalmente quando $n \geq 50$. Outros métodos também são propostos por Di Ciccio (1987) e Ahsanullah et al. (2013).

3.2 Estimador de Máxima Verossimilhança para Dados com Censura Tipo I

Na presença de observações com censura tipo I, a função de verossimilhança passa a ser escrita através de (1.17). Seja $T_1, \dots, T_n$ uma amostra aleatória com $T \sim GG(\alpha, \mu, \phi)$,

considerando observações com censura tipo I, a função de verossimilhança é dada por

$$L(\boldsymbol{\theta}; \mathbf{t}) = \frac{\alpha^d \mu^{d\alpha\phi}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\delta_i \alpha \phi - 1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i}. \quad (3.10)$$

O logaritmo da verossimilhança (3.10) é dado por

$$\begin{aligned} l(\boldsymbol{\theta}; \mathbf{t}) = & d \log(\alpha) + d\alpha\phi \log(\mu) - n \log(\Gamma(\phi)) + (\alpha\phi - 1) \sum_{i=1}^n \delta_i \log(t_i) \\ & - \mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha + \sum_{i=1}^n \{(1 - \delta_i) \log(\Gamma[\phi, (\mu t_i)^\alpha])\}. \end{aligned} \quad (3.11)$$

Objetivando-se encontrar as equações de verossimilhança, serão utilizados os seguintes resultados:

$$\frac{\partial}{\partial k} \Gamma[k, x] = \int_x^\infty w^{k-1} \log(w) e^{-w} dw = \Psi(k, x) \text{ e } \frac{\partial}{\partial x} \Gamma[k, x] = -x^{k-1} e^{-x}. \quad (3.12)$$

Estabelecendo $\frac{\partial}{\partial \alpha} \log(L(\boldsymbol{\theta}; \mathbf{t}))$, $\frac{\partial}{\partial \mu} \log(L(\boldsymbol{\theta}; \mathbf{t}))$ e $\frac{\partial}{\partial \phi} \log(L(\boldsymbol{\theta}; \mathbf{t}))$ iguais a zero, obtemos as equações de verossimilhança dadas por

$$\begin{aligned} \sum_{i=1}^n \left\{ (1 - \delta_i) \left[\frac{(\mu t_i)^{\alpha\phi} e^{-(\mu t_i)^\alpha} \log(\mu t_i)}{\Gamma[\phi, \mu t_i^\alpha]} \right] \right\} = & \frac{d}{\alpha} + d\phi \log(\mu) \\ & + \phi \sum_{i=1}^n \delta_i \log(t_i) - \mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \log(\mu t_i) \end{aligned} \quad (3.13)$$

$$\frac{d\alpha\phi}{\mu} - \alpha\mu^{\alpha-1} \sum_{i=1}^n \delta_i t_i^\alpha = \sum_{i=1}^n \left\{ (1 - \delta_i) \left[\frac{\alpha t_i (\mu t_i)^{\alpha\phi-1} e^{-(\mu t_i)^\alpha}}{\Gamma[\phi, \mu t_i^\alpha]} \right] \right\} \quad (3.14)$$

$$d\alpha \log(\mu) - n\psi(\phi) + \alpha \sum_{i=1}^n \delta_i \log(t_i) = - \sum_{i=1}^n \left\{ (1 - \delta_i) \left[\frac{\Psi[\phi, \mu t_i^\alpha]}{\Gamma[\phi, \mu t_i^\alpha]} \right] \right\} \quad (3.15)$$

cujas soluções fornecem os estimadores de máxima verossimilhança de ϕ , μ e α . Métodos numéricos como o de Newton-Raphson são necessários para encontrar a solução deste sistema não linear.

As equações de verossimilhança, considerando observações com censura do tipo aleatória, são equivalentes às obtidas considerando censura tipo I, sendo que a censura tipo I, é um caso particular da censura aleatória (veja, Lawless, 1982).

Outros tipos de censura podem ser considerados ao se realizar inferências nos parâmetros da distribuição GG. Chen, e Lio (2009) obtêm as equações de verossimilhança para os casos em que as observações possuem censura progressiva tipo II (veja, por exemplo, Balakrishnan e Aggarwala, 2000).

4 Análise Bayesiana da Distribuição Gama Generalizada

A descoberta de regiões para os valores reais dos parâmetros do modelo, também podem ser de grande interesse sob o paradigma Bayesiano (Noortwijk, 2001; Dadpay et al., 2007; Mukherjee et al., 2010; Chang e Kim, 2011; Santos, 2010), onde diferentes distribuições a priori, ou diferentes valores dos hiperparâmetros, podem levar a inferências bem diferentes (médias a posteriori e intervalos de credibilidade).

Nesta seção, serão obtidas as distribuições a priori e a posteriori para a distribuição Gama Generalizada, discutidas na seção 1.3, considerando-se dados completos e com censura tipo aleatória.

4.1 Densidades a Posteriori Considerando Observações Completas

4.1.1 Densidade a Posteriori com a Priori Uniforme

A distribuição a priori conjunta de $\theta = (\phi, \mu, \alpha)$ pode ser obtida, a partir de distribuições a priori uniformes marginais e independentes. Desta forma, a distribuição a priori conjunta é dada por

$$\pi_U(\phi, \mu, \alpha) \propto \text{cte}, \quad (4.1)$$

sendo $a_1 < \phi < a_2$, $b_1 < \mu < b_2$ e $c_1 < \alpha < c_2$.

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a priori não-informativa uniforme é proporcional ao produto da função de verossimilhança (3.1) e da distribuição a priori (4.1), resultando em

$$p_U(\phi, \mu, \alpha | \mathbf{t}) \propto \frac{\alpha^n}{\Gamma(\phi)^n} \mu^{n\alpha\phi} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\}. \quad (4.2)$$

As distribuições a posteriori condicionais de ϕ, μ e α são dadas a seguir:

$$p_U(\alpha | \phi, \mu, \mathbf{t}) \propto \alpha^n \mu^{n\alpha\phi} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\},$$

$$p_U(\phi | \mu, \alpha, \mathbf{t}) \propto \frac{\mu^{n\alpha\phi}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\}, \quad (4.3)$$

$$p_U(\mu|\phi, \alpha, \mathbf{t}) \sim \text{GG} \left(n\phi + \frac{1}{\alpha}, \left(\sum_{i=1}^n t_i^\alpha \right)^{1/\alpha}, \alpha \right).$$

Os hiperparâmetros das distribuições a priori podem ser obtidos através dos dados, utilizando-se o método dos momentos (priori empírica). Sejam, λ_i e σ_i^2 , a média e a variância de θ_i , como $\theta_i \sim \text{Uniforme}(a_i, b_i)$, então

$$\lambda_i = \frac{1}{2}(a_i + b_i) \quad \text{e} \quad \sigma_i^2 = \frac{1}{12}(b_i - a_i)^2, \quad i = 1, \dots, 3.$$

Após algumas manipulações algébricas, os hiperparâmetros podem ser encontrados através de

$$a_i = \max \left(0, \lambda_i - \sqrt{3\sigma_i^2} \right) \quad \text{e} \quad b_i = \lambda_i + \sqrt{3\sigma_i^2}, \quad i = 1, \dots, 3. \quad (4.4)$$

em que $\max(x, y)$ é o valor máximo entre x e y . Esta restrição é necessária pois em alguns casos a_i poderia obter valores negativos.

4.1.2 Densidade a Posteriori com a Priori Gama

A distribuição Gama Generalizada possui parâmetros reais não negativos. Uma distribuição a priori conjunta de ϕ, μ e α pode ser obtida através do produto de distribuições gama (1.5). Utilizando a equação (1.25), sendo $\boldsymbol{\theta} = (\phi, \mu, \alpha)$, a distribuição a priori é dada por

$$\pi_G(\phi, \mu, \alpha) \propto \phi^{a_1-1} \mu^{a_2-1} \alpha^{a_3-1} \exp(-b_1\phi - b_2\mu - b_3\alpha). \quad (4.5)$$

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a distribuição a priori Gama, é proporcional ao produto da função de verossimilhança (3.1) e da distribuição a priori (4.5), resultando em

$$\begin{aligned} p_G(\phi, \mu, \alpha|\mathbf{t}) &\propto \frac{\alpha^{n+a_3-1} \mu^{n\alpha\phi+a_2-1} \phi^{a_1-1}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\{-b_1\phi - b_2\mu - b_3\alpha\} \times \\ &\times \exp\left\{-\mu^\alpha \sum_{i=1}^n t_i^\alpha\right\}. \end{aligned} \quad (4.6)$$

As distribuições a posteriori condicionais de ϕ, μ e α são dadas a seguir:

$$\begin{aligned} p_G(\alpha|\phi, \mu, \mathbf{t}) &\propto \alpha^{n+a_3-1} \mu^{n\alpha\phi+a_2-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\left\{-b_3\alpha - \mu^\alpha \sum_{i=1}^n t_i^\alpha\right\}, \\ p_G(\phi|\mu, \alpha, \mathbf{t}) &\propto \frac{\mu^{n\alpha\phi+a_2-1} \phi^{a_1-1}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp\{-b_1\phi\}, \\ p_G(\mu|\phi, \alpha, \mathbf{t}) &\propto \mu^{n\alpha\phi+a_2-1} \exp\left\{-b_2\mu - \mu^\alpha \sum_{i=1}^n t_i^\alpha\right\}. \end{aligned} \quad (4.7)$$

Os hiperparâmetros das distribuições a priori podem ser obtidos através dos dados, utilizando-se o método dos momentos (priori empírica). Sejam, λ_i e σ_i^2 , a média e a variância de θ_i , como $\theta_i \sim \text{Gamma}(a_i, b_i)$, então

$$\lambda_i = \frac{a_i}{b_i} \quad \text{e} \quad \sigma_i^2 = \frac{a_i}{b_i^2}, \quad i = 1, \dots, 3.$$

Após algumas manipulações algébricas, os hiperparâmetros podem ser encontrados através de

$$a_i = \frac{\lambda_i^2}{\sigma_i^2} \quad \text{e} \quad b_i = \frac{\lambda_i}{\sigma_i^2}, \quad i = 1, \dots, 3. \quad (4.8)$$

4.1.3 Densidade a Posteriori com a Regra de Jeffreys

Outra distribuição a priori não-informativa, que representa uma situação com pouca informação a priori sobre os parâmetros de interesse, pode ser obtida utilizando-se a regra de Jeffreys (1.25). Como os parâmetros da distribuição Gama Generalizada estão contidos em intervalos $(0, \infty)$, a distribuição a priori utilizando a regra de Jeffreys é

$$\pi_R(\phi, \mu, \alpha) \propto (\phi\mu\alpha)^{-1}. \quad (4.9)$$

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a distribuição a priori obtida pela regra de Jeffreys, é proporcional ao produto da função de verossimilhança (3.1) e da distribuição a priori (4.9), resultando em

$$p_R(\phi, \mu, \alpha | \mathbf{t}) = \frac{1}{d_1} \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\}. \quad (4.10)$$

No entanto a densidade a posteriori (4.10) é imprópria, pois

$$d_1 = \int_{\mathcal{A}} \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} = \infty \quad (4.11)$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$. A demonstração de (4.11) pode ser verificada no apêndice A.

4.1.4 Densidade a Posteriori com a Priori de Jeffreys

A priori de Jeffreys é definida a partir da raiz quadrada do determinante da matriz de informação de Fisher $I(\boldsymbol{\theta})$. Calculando-se a raiz quadrada do determinante de $I(\phi, \mu, \alpha)$ dado em (2.12), temos

$$\pi_J(\phi, \mu, \alpha) \propto \frac{\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\mu}. \quad (4.12)$$

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a priori Jeffreys, é proporcional ao produto da função de verossimilhança (3.1) e da distribuição a priori (4.12) resultando em

$$p_J(\phi, \mu, \alpha | \mathbf{t}) = \frac{1}{d_2} \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\}. \quad (4.13)$$

No entanto a densidade a posteriori (4.13) é imprópria, pois

$$d_2 = \int_{\mathcal{A}} \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} = \infty, \quad (4.14)$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$. A demonstração de (4.14) pode ser verificada no apêndice A.

4.1.5 Densidade a Posteriori com a Priori de Zellner

A distribuição a priori de Zellner (1977, 1984), mais conhecida como *Maximal Data Information Prior* (MDIP), pode ser obtida através de

$$\pi_Z(\phi, \mu, \alpha) \propto \exp(H(\phi, \mu, \alpha)). \quad (4.15)$$

Primeiramente deve-se encontrar a solução da medida de informação, dada por

$$H(\phi, \mu, \alpha) = \int_0^\infty \log \left(\frac{\alpha}{\Gamma(\phi)} \mu^{\alpha\phi} t^{\alpha\phi-1} \exp(-(\mu t)^\alpha) \right) f(t | (\phi, \mu, \alpha)) dt.$$

Após algumas manipulações algébricas, $H(\phi, \mu, \alpha)$ pode ser escrita como:

$$H(\phi, \mu, \alpha) = \log(\alpha) - \log(\Gamma(\phi)) + \phi\psi(\phi) + \log(\mu) - \frac{\psi(\phi)}{\alpha} - \phi.$$

Portanto a priori de Zellner (4.15) da distribuição Gama Generalizada (2.2) é dada por

$$\pi_Z(\phi, \mu, \alpha) \propto \frac{\alpha\mu}{\Gamma(\phi)} \exp \left(\psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right). \quad (4.16)$$

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a priori MDIP, é proporcional ao produto da função de verossimilhança (3.1) e da distribuição a priori (4.16), resultando em

$$p_Z(\phi, \mu, \alpha | \mathbf{t}) = \frac{1}{d_3} \frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \times \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\}. \quad (4.17)$$

No entanto a densidade a posteriori (4.17) é imprópria, pois

$$d_3 = \int_{\mathcal{A}} \frac{\alpha^{n+1} \mu^{n\alpha\phi+1}}{\Gamma(\phi)^{n+1}} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} d\boldsymbol{\theta} = \infty. \quad (4.18)$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$. A demonstração de (4.18) pode ser verificada no apêndice A.

4.1.6 Densidade a Posteriori com a Regra de Jeffreys Modificada

Uma modificação na priori obtida pela regra de Jeffreys (4.9) foi realizada. Desta forma, foi possível construir uma distribuição a priori conjunta que, combinada com a função de verossimilhança, resultasse em uma distribuição a posteriori própria. Esta priori é dada

$$\pi_M(\phi, \mu, \alpha) \propto \frac{1}{\phi \mu \alpha^{\frac{1}{2} + \alpha/(1+\alpha)}}. \quad (4.19)$$

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a distribuição a priori obtida pela regra de Jeffreys modificada, é proporcional ao produto da função de verossimilhança (3.1) e da distribuição a priori (4.19), resultando em

$$p_M(\phi, \mu, \alpha | \mathbf{t}) = \frac{1}{d_4} \frac{\alpha^{n - \frac{1}{2} - \alpha/(1+\alpha)}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\}. \quad (4.20)$$

A distribuição a posteriori é própria, pois

$$d_4 = \int_{\mathcal{A}} \frac{\alpha^{n - \frac{1}{2} - \alpha/(1+\alpha)}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} < \infty, \quad (4.21)$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$. A demonstração de (4.21) pode ser verificada no apêndice B.

As distribuições a posteriori condicionais de ϕ, μ e α são dadas a seguir:

$$\begin{aligned} p_M(\alpha | \phi, \mu, \mathbf{t}) &\propto \alpha^{n - \frac{1}{2} - \alpha/(1+\alpha)} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\}, \\ p_M(\phi | \mu, \alpha, \mathbf{t}) &\propto \frac{\mu^{n\alpha\phi-1}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\}, \\ p_M(\mu | \phi, \alpha, \mathbf{t}) &\sim \text{GG} \left(n\phi, \left(\sum_{i=1}^n t_i^\alpha \right)^{1/\alpha}, \alpha \right). \end{aligned} \quad (4.22)$$

4.2 Densidades a Posteriori Considerando Observações Censuradas

4.2.1 Densidade a Posteriori com a Priori Uniforme

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a priori uniforme é proporcional ao produto da função de verossimilhança (3.10) e da distribuição a priori (4.1), resultando em

$$p_U(\phi, \mu, \alpha | \mathbf{t}, \boldsymbol{\delta}) \propto \frac{\alpha^d \mu^{d\alpha\phi}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\delta_i(\alpha\phi-1)} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i}. \quad (4.23)$$

sendo $a_1 < \phi < b_1, a_2 < \mu < b_2$ e $a_3 < \alpha < b_3$.

As distribuições a posteriori condicionais de ϕ, μ e α são dadas a seguir:

$$\begin{aligned} p_U(\alpha | \phi, \mu, \mathbf{t}, \boldsymbol{\delta}) &\propto \alpha^d \mu^{d\alpha\phi} \left\{ \prod_{i=1}^n t_i^{\delta_i(\alpha\phi-1)} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i}, \\ p_U(\phi | \mu, \alpha, \mathbf{t}, \boldsymbol{\delta}) &\propto \frac{\mu^{d\alpha\phi}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\delta_i(\alpha\phi-1)} \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i}, \\ p_U(\mu | \phi, \alpha, \mathbf{t}, \boldsymbol{\delta}) &\propto \mu^{d\alpha\phi} \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i}. \end{aligned} \quad (4.24)$$

4.2.2 Densidade a Posteriori com a Priori Gama

A distribuição a posteriori conjunta para ϕ, μ e α , utilizando-se a distribuição a priori Gama, é proporcional ao produto da função de verossimilhança (3.10) e da distribuição a priori (4.5), resultando em

$$\begin{aligned} p_G(\phi, \mu, \alpha | \mathbf{t}, \boldsymbol{\delta}) &\propto \frac{\alpha^{d+a_3-1} \mu^{d\alpha\phi+a_2-1} \phi^{a_1-1}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\delta_i(\alpha\phi-1)} \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i} \times \\ &\times \exp \left\{ -b_1\phi - b_2\mu - b_3\alpha - \mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\}. \end{aligned} \quad (4.25)$$

As distribuições a posteriori condicionais de ϕ, μ e α são dadas a seguir:

$$\begin{aligned} p_G(\alpha | \phi, \mu, \mathbf{t}, \boldsymbol{\delta}) &\propto \alpha^{d+a_3-1} \mu^{d\alpha\phi+a_2-1} \phi^{a_1-1} \left\{ \prod_{i=1}^n t_i^{\delta_i(\alpha\phi-1)} \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i} \times \\ &\times \exp \left\{ -b_3\alpha - \mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\}, \\ p_G(\phi | \mu, \alpha, \mathbf{t}, \boldsymbol{\delta}) &\propto \frac{\mu^{d\alpha\phi+a_2-1} \phi^{a_1-1}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\delta_i(\alpha\phi-1)} \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i} \times \\ &\times \exp \{-b_1\phi\}, \end{aligned} \quad (4.26)$$

$$p_G(\mu|\phi, \alpha, \mathbf{t}, \boldsymbol{\delta}) \propto \mu^{d\alpha\phi+a_2-1} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i} \exp \left\{ -b_2\mu - \mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\}.$$

4.3 Algoritmo Híbrido (MCMC)

Devido à impossibilidade de se encontrar formas fechadas para as distribuições a posteriori condicionais de ϕ, μ e α , citadas neste capítulo, sob o contexto bayesiano, será utilizado o algoritmo híbrido, buscando se obter as características de interesse das distribuições marginais a posteriori de ϕ, μ e α .

A distribuição proposta $q_i(\theta_i^{(t)}|\theta_i^{(*)})$ utilizada, será uma distribuição Normal truncada entre $(0, \infty)$ com média a_i e desvio-padrão b_i , e denominada $\text{Nt}(a_i, b_i)$.

O algoritmo de Metropolis-Hastings híbrido, utilizando-se diferentes distribuições a posteriori, é especificado pelos seguintes passos:

1. Defina os valores iniciais $\boldsymbol{\theta}^{(0)} = (\alpha^{(0)}, \phi^{(0)}, \mu^{(0)})'$ e inicie o contador $t = 0$;
2. Gere um novo valor $\alpha^{(*)}$ da densidade proposta $\text{Nt}(a_1, b_1)$;
3. Calcule a probabilidade de aceitação

$$\alpha_1(\alpha^{(t)}, \alpha^{(*)}) = \min \left(1, \frac{p_1(\alpha^{(*)}|\phi^{(t)}, \mu^{(t)}, \mathbf{x}) \, \text{dNt}(\alpha^{(t)}, a_1, b_1)}{p_1(\alpha^{(t)}|\phi^{(t)}, \mu^{(t)}, \mathbf{x}) \, \text{dNt}(\alpha^{(*)}, a_1, b_1)} \right),$$

e amostrasse u de uma $\text{Uniforme}(0, 1)$;

4. Se $\alpha_1(\alpha^{(t)}, \alpha^{(*)}) \geq u(0, 1)$ então $\alpha^{(t+1)} = \alpha^{(*)}$, caso contrário $\alpha^{(t+1)} = \alpha^{(t)}$;
5. Gere um novo valor $\phi^{(*)}$ da densidade proposta $\text{Nt}(a_2, b_2)$;
6. Calcule a probabilidade de aceitação

$$\alpha_2(\phi^{(t)}, \phi^{(*)}) = \min \left(1, \frac{p_2(\phi^{(*)}|\alpha^{(t+1)}, \mu^{(t)}, \mathbf{x}) \, \text{dNt}(\phi^{(t)}, a_2, b_2)}{p_2(\phi^{(t)}|\alpha^{(t+1)}, \mu^{(t)}, \mathbf{x}) \, \text{dNt}(\phi^{(*)}, a_2, b_2)} \right),$$

e amostrasse u de uma $\text{Uniforme}(0, 1)$;

7. Se $\alpha_2(\phi^{(t)}, \phi^{(*)}) \geq u(0, 1)$ então $\phi^{(t+1)} = \phi^{(*)}$, caso contrário $\phi^{(t+1)} = \phi^{(t)}$;
8. Gere um novo valor $\mu^{(*)}$ da densidade proposta $\text{Nt}(a_3, b_3)$;
9. Calcule a probabilidade de aceitação

$$\alpha_3(\mu^{(t)}, \mu^{(*)}) = \min \left(1, \frac{p_3(\mu^{(*)}|\alpha^{(t+1)}, \phi^{(t+1)}, \mathbf{x}) \, \text{dNt}(\mu^{(t)}, a_3, b_3)}{p_3(\mu^{(t)}|\alpha^{(t+1)}, \phi^{(t+1)}, \mathbf{x}) \, \text{dNt}(\mu^{(*)}, a_3, b_3)} \right),$$

e amostrasse u de uma $\text{Uniforme}(0, 1)$;

10. Se $\alpha_3(\mu^{(t)}, \mu^{(*)}) \geq u(0, 1)$ então $\mu^{(t+1)} = \mu^{(*)}$, caso contrário $\mu^{(t+1)} = \mu^{(t)}$;
11. Altere o contador de t para $t + 1$ e volte ao passo 2, até a convergência ser alcançada.

Ao se utilizar as distribuições a posteriori condicionais, obtidas da distribuição Gama Generalizada, há grande dificuldade em se efetuar o calculo da razão

$$\frac{p_i(\theta_i^{(*)}|\theta_{-i}^{(t)}, \mathbf{x})}{p_i(\theta_i^{(t)}|\theta_{-i}^{(t)}, \mathbf{x})},$$

esta dificuldade se deve ao fato que $p_i(\theta_i^{(*)}|\theta_{-i}^{(t)}, \mathbf{x})$ possui, em muitos casos, os termos $\Gamma(n\phi)$ e $(\Gamma(\phi)^n)^{-1}$. Dependendo dos valores de ϕ e n os *softwares* utilizados não conseguem computar tais termos.

Por exemplo, considerando $\phi = 3$ e $n = 50$, os valores obtidos serão, $\Gamma(150) = 3.81e^{260}$ e $(\Gamma(3)^{50})^{-1} = 8.88e^{-16}$. Para valores de $n \geq 100$, o R não conseguirá calcular tais termos, retornando o símbolo ∞ no primeiro caso e 0 para o segundo caso.

Está dificuldade pode ser contornada, reescrevendo-se a forma em que calculamos a probabilidade de aceitação, com

$$\alpha_i(\theta_i^{(t)}, \theta_i^{(*)}) = \min(1, e^{\Omega}),$$

$$\Omega = \log(p_i(\theta_i^{(*)}|\theta_{-i}^{(t)}, x)) + \log(q_i(\theta_i^{(t)}|\theta_i^{(*)})) - \log(p_i(\theta_i^{(t)}|\theta_{-i}^{(t)}, x)) - \log(q_i(\theta_i^{(*)}|\theta_i^{(t)})).$$

O método se mostra satisfatório, pois é fácil calcular os termos $n \log(\Gamma(\phi))$ e $\log(\Gamma(n\phi))$. Para se calcular $\log(\Gamma(n\phi))$ podemos utilizar a aproximação

$$\log(\Gamma(x)) = x \log(x) - x - \frac{1}{2} \log\left(\frac{x}{2\pi}\right) + \frac{1}{12x} - \frac{1}{360x^3} + \frac{1}{1260x^5} + \text{Erro}. \quad (4.27)$$

5 Equações Úteis para Encontrar Bons Valores Iniciais

É comum se deparar com dificuldades computacionais ao se realizar inferências nos parâmetros da distribuição Gama Generalizada. As estimativas dos parâmetros obtidas utilizando-se a abordagem Clássica ou Bayesiana, considerando dados completos e censurados, podem depender dos valores iniciais utilizados nos métodos computacionais iterativos ou também da distribuição a priori utilizada.

Desta forma, é de grande importância o desenvolvimento de técnicas exploratórias para se obter bons valores iniciais. Serão propostos a seguir diferentes técnicas para se obter bons valores iniciais que serão, utilizados tanto na abordagem clássica quanto na bayesiana.

5.1 Análise Clássica

Para se utilizar o método da máxima verossimilhança, considerando-se dados completos, foram obtidos na seção 3.1 novos resultados de grande importância computacional em que, resolvendo a equação (3.9), encontra-se facilmente o EMV de α e com este resultado obtêm-se através de (3.8 e 3.7), respectivamente, os EMV's de ϕ e μ .

No entanto, quando utilizamos observações censuradas, as equações de verossimilhança (3.13, 3.14 e 3.15) não podem ser simplificadas similarmente, como foi feito nas equações de verossimilhança (3.2, 3.3 e 3.4). Desta forma, as equações obtidas considerando-se observações censuradas são muito instáveis e as estimativas obtidas de ϕ , μ e α podem depender dos valores iniciais, utilizados no método iterativo.

Valores preditos para as observações censuradas podem ser aproximados utilizando-se uma distribuição exponencial com densidade $f(t) = \lambda e^{-\lambda t}$, $t > 0$. A esperança condicional $E(T_i | T_i > t^*)$ é dada por

$$E(T_i | T_i > t^*) = \frac{\lambda \int_{t^*}^{\infty} x e^{-\lambda x} dx}{e^{-\lambda t^*}} = t^* + \frac{1}{\lambda}. \quad (5.1)$$

Para se obter o valor de λ , utiliza-se o estimador de máxima verossimilhança $\hat{\lambda} = n / \sum_{i=1}^n t_i$ baseado apenas nas observações completas. Por meio da equação (5.1) obtemos os valores preditos aproximados para as observações censuradas t^* .

Utilizando-se as observações completas e as preditas, através das equações (3.7, 3.8 e 3.9), encontramos bons valores iniciais $\tilde{\phi}$, $\tilde{\mu}$ e $\tilde{\alpha}$. Estes valores obtidos podem ser

utilizados para inicialização dos métodos iterativos, fazendo com que, ao considerarmos dados censurados, obtenhamos boas estimativas para ϕ , μ e α .

Bons valores iniciais para se utilizar nas equações (3.13, 3.14 e 3.15) também podem ser obtidos, removendo-se as observações censuradas e calculando-se $\tilde{\phi}$, $\tilde{\mu}$ e $\tilde{\alpha}$ através das equações (3.7, 3.8 e 3.9). Embora haja perda de informação ao se remover as observações censuradas, como se trata de uma análise preliminar, os valores $\tilde{\phi}$, $\tilde{\mu}$ e $\tilde{\alpha}$ podem ser utilizados apenas para se inicializar o método iterativo.

Vale salientar que, em muitos casos, há grande dificuldade em se calcular os estimadores dos parâmetros, quando as observações t_i são muito grandes. Uma alternativa seria realizar uma transformação linear nos dados, através de $t_i/100$ ou t_i/c , sendo c uma constante maior que zero. Desta forma, utilizando-se a propriedade (2.9), os estimadores de ϕ , μ e α serão dados por $\hat{\phi}$, $\hat{\mu}/c$ e $\hat{\alpha}$.

5.2 Análise Bayesiana

Utilizando-se a abordagem Bayesiana, diferentes distribuições a priori podem levar a inferências bem diferentes das esperadas, principalmente quando utilizamos distribuições a priori não-informativas. Será descrita a seguir uma técnica exploratória para se obter bons valores iniciais. Com estes valores será possível construir distribuições a priori empíricas para os parâmetros ϕ , μ e α .

A distribuição a posteriori conjunta de ϕ , μ e α , utilizando-se a distribuição a priori obtida pela regra de Jeffreys modificada, é dada em (4.20). A distribuição a posteriori conjunta de α e ϕ é obtida integrando-se (4.20) em relação a μ .

$$p_M(\phi, \alpha | \mathbf{t}) = \frac{1}{d_4} \frac{\alpha^{n-\frac{1}{2}-\alpha/(1+\alpha)}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \int_0^\infty \mu^{n\alpha\phi-1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu. \quad (5.2)$$

Utilizando-se o resultado

$$\int_0^\infty x^{\alpha-1} e^{-\beta x} dx = \frac{\Gamma(\alpha)}{\beta^\alpha} \quad (5.3)$$

e considerando-se a transformação $v = \mu^\alpha$, obtemos

$$\begin{aligned} \int_0^\infty \mu^{n\alpha\phi-1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu &= \frac{1}{\alpha} \int_0^\infty v^{n\phi-1} \exp \left\{ -v \sum_{i=1}^n t_i^\alpha \right\} dv \\ &= \frac{\Gamma(n\phi)}{\alpha (\sum_{i=1}^n t_i^\alpha)^{n\phi}}. \end{aligned} \quad (5.4)$$

Substituindo-se a solução da integral (5.4) em (5.2), a distribuição a posteriori conjunta de α e ϕ é dada por

$$p_M(\phi, \alpha | \mathbf{t}) = \frac{1}{d_4} \frac{\alpha^{n-\frac{3}{2}-\alpha/(1+\alpha)}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}}. \quad (5.5)$$

O logaritmo da distribuição a posteriori conjunta de α e ϕ é dado por

$$\begin{aligned} \log(p_M(\alpha, \phi|\mathbf{t})) = & \left(n - \frac{3}{2} - \frac{\alpha}{(1+\alpha)}\right) \log(\alpha) + \log(\Gamma(n\phi)) - \log(\phi) - \log(d_4) \\ & + (\alpha\phi - 1) \sum_{i=1}^n \log(t_i) - n \log(\Gamma(\phi)) - n\phi \log\left(\sum_{i=1}^n t_i^\alpha\right). \end{aligned} \quad (5.6)$$

As derivadas parciais de primeira ordem do $\log(p_M(\alpha, \phi|\mathbf{t}))$, com respeito a α e ϕ , são dadas respectivamente por

$$\begin{aligned} \frac{\partial \log(p_M(\alpha, \phi|\mathbf{t}))}{\partial \alpha} &= \kappa(\alpha) + \phi \sum_{i=1}^n \log(t_i) - n\phi \left(\frac{\sum_{i=1}^n t_i^\alpha \log(t_i)}{\sum_{i=1}^n t_i^\alpha} \right), \\ \frac{\partial \log(p_M(\alpha, \phi|\mathbf{t}))}{\partial \phi} &= n\psi(n\phi) + \alpha \sum_{i=1}^n \log(t_i) - \frac{1}{\phi} - n\psi(\phi) - n \log\left(\sum_{i=1}^n t_i^\alpha\right), \end{aligned} \quad (5.7)$$

sendo $\kappa(\alpha) = \left(n - \frac{3}{2} - \frac{\alpha}{(1+\alpha)}\right) \frac{1}{\alpha} - \frac{\alpha}{(1+\alpha)^2} \log(\alpha)$.

As modas $\tilde{\phi}$ e $\tilde{\alpha}$ da distribuição a posteriori conjunta $p_M(\alpha, \phi|\mathbf{t})$ são obtidas resolvendo-se as seguintes equações:

$$\frac{\partial \log(p_M(\alpha, \phi|\mathbf{t}))}{\partial \alpha} = 0 \quad \text{e} \quad \frac{\partial \log(p_M(\alpha, \phi|\mathbf{t}))}{\partial \phi} = 0.$$

Resolvendo $\partial \log(p_M(\alpha, \phi|\mathbf{t}))/\partial \alpha$, obtemos a equação

$$\tilde{\phi} = \frac{\kappa(\tilde{\alpha})}{\tilde{\alpha} \left(\frac{n \sum_{i=1}^n t_i^{\tilde{\alpha}} \log(t_i)}{\sum_{i=1}^n t_i^{\tilde{\alpha}}} - \sum_{i=1}^n \log(t_i) \right)}. \quad (5.8)$$

Resolvendo $\partial \log(p_M(\alpha, \phi|\mathbf{t}))/\partial \phi$, obtemos $\tilde{\alpha}$ substituindo-se ϕ por $\tilde{\phi}$ e resolvendo-se a seguinte equação em α , dada por

$$h(\alpha) = n\psi(n\tilde{\phi}) + \alpha \sum_{i=1}^n \log(t_i) - \frac{1}{\tilde{\phi}} - \psi(\tilde{\phi}) - \log\left(\sum_{i=1}^n t_i^\alpha\right) = 0. \quad (5.9)$$

Utilizando-se métodos iterativos como Newton-Raphson obtemos os valores de $\tilde{\alpha}$. Substituindo-se $\tilde{\alpha}$ na equação (5.8), encontra-se $\tilde{\phi}$. O valor inicial de μ pode ser obtido resolvendo-se a equação

$$\frac{\partial \log(p_M(\phi, \mu, \alpha|\mathbf{t}))}{\partial \mu} = (n\tilde{\phi}\tilde{\alpha} - 1) - \tilde{\alpha}\tilde{\mu}^\alpha \sum_{i=1}^n t_i^{\tilde{\alpha}} = 0$$

e após algumas manipulação algébricas,

$$\tilde{\mu} = \left(\frac{n\tilde{\alpha}\tilde{\phi} - 1}{\tilde{\alpha} \sum_{i=1}^n t_i^{\tilde{\alpha}}} \right)^{\frac{1}{\tilde{\alpha}}}, \quad \forall \tilde{\mu} > 0 \text{ e } n\tilde{\alpha}\tilde{\phi} > 1. \quad (5.10)$$

Quando $n\tilde{\alpha}\tilde{\phi} < 1$, podemos obter $\tilde{\mu}$ por meio da equação (3.3), tal que, após algumas manipulações algébricas, obtemos

$$\tilde{\mu} = \left(\frac{n\tilde{\phi}}{\sum_{i=1}^n t_i^{\tilde{\alpha}}} \right)^{\frac{1}{\tilde{\alpha}}}. \quad (5.11)$$

Com esta abordagem encontramos facilmente bons valores iniciais $\tilde{\phi}$, $\tilde{\mu}$ e $\tilde{\alpha}$. Estes resultados são de grande interesse prático pois, podem ser utilizados em uma análise Bayesiana para a construção de distribuições a priori empíricas dos parâmetros ϕ , μ e α e também como valores para iniciar o algoritmo híbrido a ser descrito na próxima seção.

6 Análise Numérica

Neste capítulo, serão realizadas simulações via método de Monte Carlo. Os objetivos destas simulações serão:

1. Comparar o desempenho dos estimadores de máxima verossimilhança para dados completos, obtidos em (3.2, 3.3 e 3.4) e (3.7, 3.8 e 3.9).
2. Verificar o efeito dos valores iniciais obtidos das equações (3.7, 3.8 e 3.9), considerando-se as observações preditas (5.1) nas estimativas de máxima verossimilhança, levando-se em conta dados censurados (3.13, 3.14 e 3.15).
3. Comparar o efeito dos valores iniciais (5.8, 5.9 e 5.10) nas estimativas a posteriori para dados completos e censurados, considerando-se diferentes distribuições a priori.

Para se obter as estatísticas das distribuições a posteriori, serão utilizados os métodos MCMC, com foco no algoritmo híbrido. Também será investigada a influência do tamanho da amostra na acurácia dos estimadores.

O seguinte procedimento foi adotado no estudo:

1. Determine os valores de $\boldsymbol{\theta} = (\phi, \mu, \alpha)$.
2. Especifique o tamanho da amostra n .
3. Gere valores de uma distribuição GG(ϕ, μ, α) com tamanho n .
4. Utilizando os dados obtidos no passo 3, calcule as estimativas a posteriori $\hat{\phi}$, $\hat{\mu}$ e $\hat{\alpha}$, via MCMC ou os EMV's para os parâmetros ϕ , μ e α .
5. Repita o procedimento 3 e 4 N vezes.

Os valores escolhidos para se realizar tal procedimento serão, $\boldsymbol{\theta} = ((0.5, 0.5, 3), (2, 1, 0.5), (4, 2, 2), (0.4, 1.5, 5))$, $N = 1000$ e $n = (50, 100, 200)$. A semente utilizada para gerar os valores aleatórios no R é 2013.

A comparação entre os métodos será feita através do cálculo das médias e dos desvios-padrão das N estimativas, obtidas através dos EMV's e das médias a posteriori. Espera-se que o melhor método de estimação, possua as médias das N estimativas, mais perto dos verdadeiros valores de $\boldsymbol{\theta}$ com menores desvios-padrão.

Serão calculadas também as probabilidades de cobertura dos parâmetros, através dos intervalos de confiança assintóticos (inferência Clássica) e dos intervalos de credibilidade

(abordagem Bayesiana). Para um grande número de experimentos, utilizando-se um nível de confiança de 95%, as frequências dos intervalos que cobriram os verdadeiros valores de θ devem ser de aproximadamente 95%.

6.1 Análise Clássica

6.1.1 Observações Completas

Nesta seção, iremos realizar simulações numéricas, utilizando-se a inferência Clássica com foco no método da máxima verossimilhança, considerando-se observações completas.

Ao utilizar-se o método de Newton-Raphson (1.19), para encontrar a solução do sistema (3.2, 3.3 e 3.4), é necessário especificar os valores iniciais para inicialização do procedimento iterativo. Espera-se que o método, após algumas iterações, leve aos verdadeiros valores de $\hat{\theta}$.

Na ausência de informação sobre quais valores iniciais devem ser utilizados no método iterativo, serão gerados para ϕ , μ e α valores aleatórios de distribuições uniformes com intervalos entre $[0, 5)$. Portanto os valores utilizados para a inicialização do método de Newton-Raphson serão $\theta^{(0)} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$. Este procedimento será denominado de Método 1 e o pacote do R utilizado para resolver o método iterativo será o maxNR.

Como foi demonstrado, as equações de verossimilhança (3.2, 3.3 e 3.4) podem ser simplificadas através de (3.7, 3.8 e 3.9). Será utilizado o pacote multiroot, disponível no R, para resolver a equação (3.9). A solução desta equação acarreta na obtenção do EMV de α e com este resultado obtêm-se facilmente, através de (3.7 e 3.8), os EMV's de μ e ϕ . Este procedimento será denominado de Método 2.

Na Tabela 1 temos as médias e os desvios-padrão das estimativas de 1000 amostras obtidas utilizando os EMV's, calculados através dos Métodos 1 e 2 para diferentes valores de θ e n . Na Tabela 2, podemos ver as probabilidades de cobertura com o nível de confiança de 95%.

Tabela 1 – Estimativas das médias e desvios-padrão dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, através dos Métodos 1 e 2 para diferentes valores de θ e n .

θ	Método 1			Método 2		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	1,946(1,531)	1,869(1,430)	1,779(1,250)	0,725(0,745)	0,562(0,320)	0,521(0,176)
$\mu = 0,5$	2,004(3,965)	1,912(3,986)	1,654(2,139)	0,694(1,125)	0,538(0,170)	0,513(0,066)
$\alpha = 3$	1,887(1,310)	1,775(1,024)	1,727(0,885)	3,344(1,776)	3,292(1,177)	3,160(0,776)
$\phi = 2$	2,664(1,215)	2,583(1,088)	2,665(1,103)	1,682(1,046)	1,925(0,953)	2,083(0,839)
$\mu = 1$	5,581(7,043)	5,371(7,478)	5,508(7,559)	2,048(4,326)	2,435(4,590)	2,490(4,259)
$\alpha = 0,5$	0,542(0,395)	0,540(0,371)	0,495(0,260)	0,779(1,018)	0,592(0,208)	0,531(0,126)
$\phi = 4$	4,611(2,754)	4,687(2,648)	4,705(2,741)	4,880(4,121)	5,031(4,102)	4,881(3,472)
$\mu = 2$	2,658(3,658)	2,662(2,710)	2,758(2,922)	3,336(4,409)	3,438(4,442)	3,089(3,527)
$\alpha = 2$	2,265(0,970)	2,106(0,647)	2,076(0,585)	2,342(0,880)	2,257(0,840)	2,153(0,698)
$\phi = 0,4$	1,094(0,968)	1,050(0,957)	0,983(0,970)	0,707(0,712)	0,501(0,244)	0,444(0,131)
$\mu = 1,5$	2,318(1,724)	2,386(3,729)	2,352(4,507)	1,842(1,230)	1,589(0,213)	1,540(0,107)
$\alpha = 5$	3,548(2,003)	3,539(4,282)	3,526(1,492)	4,446(2,231)	4,783(1,436)	4,901(0,933)

Tabela 2 – Probabilidades de cobertura com o nível de confiança de 95% dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ utilizando-se os Métodos 1 e 2.

θ	Método 1			Método 2		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	97,30%	98,10%	98,50%	91,80%	90,30%	92,00%
$\mu = 0,5$	98,00%	98,40%	98,70%	93,20%	91,30%	93,40%
$\alpha = 3$	46,50%	37,60%	28,90%	92,50%	94,90%	95,30%
$\phi = 2$	90,00%	89,00%	90,90%	76,80%	84,20%	89,20%
$\mu = 1$	87,00%	86,50%	88,50%	61,70%	72,60%	80,50%
$\alpha = 0,5$	100,00%	96,40%	94,70%	100,00%	99,00%	98,90%
$\phi = 4$	91,80%	93,60%	91,00%	88,70%	84,40%	86,00%
$\mu = 2$	89,80%	90,90%	88,60%	83,60%	80,50%	82,00%
$\alpha = 2$	100,00%	99,90%	99,60%	100,00%	100,00%	98,80%
$\phi = 0,4$	98,30%	98,80%	98,80%	98,70%	98,30%	99,10%
$\mu = 1,5$	99,20%	99,00%	98,90%	98,90%	98,40%	99,10%
$\alpha = 5$	63,90%	52,30%	47,80%	89,50%	92,30%	94,60%

Pode-se observar pelos resultados obtidos na Tabela 1, que as estimativas de máxima verossimilhança são muito influenciadas pelos valores iniciais. Por exemplo, utilizando-se o método 1, para $\mu = 1$ e $n = 50$, a média das 1000 estimativas é de $\mu = 5,581$, com desvio-padrão $\sigma_\mu = 7,043$. Estes resultados devem-se ao fato das equações de verossimilhança (3.2, 3.3 e 3.4) possuírem diferentes pontos de máximos locais, fazendo com que em muitos casos, o método de Newton-Raphson não encontre os pontos de máximos globais.

Utilizando-se o método 2, observa-se que o método proposto retorna em todos os casos boas estimativas de ϕ, μ e α . As probabilidades de cobertura dos parâmetros mostram que, embora alguns casos estejam abaixo do ideal, em sua maioria são superiores a 90%.

Com estes resultados pode-se constatar que, ao utilizarmos o método 2, conseguimos obter facilmente boas inferências para os parâmetros da distribuição Gama Generalizada.

6.1.2 Observações Censuradas

Nesta seção iremos realizar simulações numéricas utilizando-se a inferência Clássica, com foco no método da máxima verossimilhança, considerando observações censuradas. O método de Newton-Raphson (1.19), será utilizado para encontrar a solução do sistema (3.13, 3.14 e 3.15). No entanto, para encontrar tal solução, é necessário especificar valores iniciais para inicialização do procedimento iterativo.

Similarmente ao procedimento utilizado para observações completas, na ausência de informação sobre quais valores iniciais devem ser utilizados no método iterativo, serão gerados valores aleatórios tais que $\tilde{\phi} \sim U(0, 5)$, $\tilde{\mu} \sim U(0, 5)$ e $\tilde{\alpha} \sim U(0, 5)$. Portanto os valores utilizados para a inicialização do método de Newton-Raphson serão $\boldsymbol{\theta}^{(0)} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$. Este procedimento será denominado de Método 1.

Na Tabela 3, temos as médias, os desvios-padrão e as probabilidades de cobertura com o nível de confiança de 95% das estimativas de 1000 amostras obtidas, utilizando-se os EMV's, calculados através do Método 1 para diferentes valores de $\boldsymbol{\theta}$, n e 20% de censura.

Tabela 3 – Estimativas das médias e desvios-padrão dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de $\boldsymbol{\theta}$, 20% de censura e valores iniciais obtidos através do método 1.

	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\boldsymbol{\theta}$	Média(Desvio P.) EMV	Média(Desvio P.) EMV	Média(Desvio P.) EMV	Prob. de Cobertura	Prob. de Cobertura	Prob. de Cobertura
$\phi = 0,5$	1,156(1,420)	1,112(1,364)	0,973(1,151)	70,20%	63,80%	61,60%
$\mu = 0,5$	1,242(2,770)	1,256(2,710)	1,032(2,330)	79,70%	72,80%	63,40%
$\alpha = 3$	3,827(4,221)	3,253(2,504)	3,327(4,594)	69,30%	63,80%	63,10%
$\phi = 2$	1,737(1,191)	1,757(1,055)	1,866(0,977)	59,90%	60,30%	59,00%
$\mu = 1$	2,208(4,727)	1,872(4,011)	1,883(3,809)	51,60%	49,10%	51,50%
$\alpha = 0,5$	1,017(2,103)	0,785(0,821)	0,669(0,691)	79,70%	73,00%	67,20%
$\phi = 4$	3,144(2,994)	3,072(2,620)	3,153(2,647)	56,00%	53,60%	46,60%
$\mu = 2$	1,961(2,326)	1,853(2,106)	1,921(2,291)	52,10%	48,70%	42,80%
$\alpha = 2$	3,805(6,086)	3,437(5,512)	3,321(6,103)	78,90%	72,70%	60,20%
$\phi = 0,4$	0,667(0,882)	0,659(0,936)	0,535(0,544)	77,40%	74,00%	75,00%
$\mu = 1,5$	1,842(1,617)	1,872(2,110)	1,608(0,894)	83,10%	77,50%	77,20%
$\alpha = 5$	6,266(4,290)	5,686(3,180)	5,379(2,613)	79,40%	75,80%	76,60%

Embora as estimativas estejam mais concentradas, pode-se observar que os EMV's são muito influenciados pelos valores iniciais. Mesmo quando o tamanho amostral é grande ($n = 200$), as equações (3.13, 3.14 e 3.15) não retornam boas estimativas de ϕ , μ e α . É importante ressaltar que utilizando-se o método 1, a probabilidade de cobertura tende a diminuir quando o tamanho amostral cresce, mostrando desta forma, que este método não retorna bons resultados.

Na seção 5.1, discutiu-se a possibilidade de se obter bons valores iniciais, por meio da remoção das observações censuradas, a fim de se calcular $\tilde{\phi}$, $\tilde{\mu}$ e $\tilde{\alpha}$ através das equações (3.7, 3.8 e 3.9). Estes resultados podem ser utilizados nas equações (3.13, 3.14 e 3.15),

possibilitando a obtenção dos EMV's de ϕ , μ e α . Tal procedimento será denominado de Método 3.

Discutiu-se também a possibilidade de se utilizar as observações completas e as preditas (5.1) nas equações (3.7, 3.8 e 3.9), a fim de se obter bons valores iniciais $\tilde{\phi}$, $\tilde{\mu}$ e $\tilde{\alpha}$. Este procedimento será denominado de Método 4.

Nas Tabelas 4 e 5 temos as médias e os desvios-padrão das estimativas de 1000 amostras obtidas, calculando-se os EMV's através dos Método 3 e 4 para diferentes valores de θ , n e 20% de censura. Na Tabela 6, podemos ver as probabilidades de cobertura com o nível de confiança de 95%.

Tabela 4 – Estimativas das médias e desvios-padrão dos valores iniciais $\theta^{(0)} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, obtidos através do método 3 e dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, considerando diferentes valores de θ e 20% de censura.

Mét. 3	$n = 50$		$n = 100$		$n = 200$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) EMV	Média(Desvio P.) Val .In	Média(Desvio P.) EMV	Média(Desvio P.) Val .In	Média(Desvio P.) EMV
$\phi = 0,5$	0,888(1,041)	0,865(1,061)	0,653(0,449)	0,634(0,417)	0,546(0,203)	0,539(0,198)
$\mu = 0,5$	0,843(1,883)	0,814(2,042)	0,588(0,356)	0,553(0,257)	0,523(0,080)	0,502(0,071)
$\alpha = 3$	2,900(1,503)	3,249(2,615)	3,028(1,129)	3,215(1,194)	3,092(0,774)	3,260(0,821)
$\phi = 2$	1,413(0,941)	1,353(0,947)	1,704(0,878)	1,629(0,878)	1,894(0,791)	1,810(0,812)
$\mu = 1$	1,239(2,927)	1,102(3,453)	1,636(3,574)	1,443(4,641)	1,850(3,522)	1,588(4,214)
$\alpha = 0,5$	0,911(1,074)	1,108(2,118)	0,647(0,251)	0,683(0,289)	0,566(0,142)	0,597(0,158)
$\phi = 4$	3,359(3,362)	3,389(3,430)	3,800(3,527)	3,803(3,548)	3,976(3,011)	3,975(3,016)
$\mu = 2$	2,258(3,159)	2,266(3,319)	2,527(3,597)	2,509(3,715)	2,455(2,748)	2,412(2,774)
$\alpha = 2$	2,996(1,184)	3,030(1,217)	2,695(1,043)	2,715(1,044)	2,419(0,798)	2,439(0,799)
$\phi = 0,4$	0,773(0,790)	0,750(0,775)	0,533(0,264)	0,530(0,262)	0,449(0,143)	0,449(0,140)
$\mu = 1,5$	1,922(1,360)	1,850(1,279)	1,616(0,235)	1,583(0,216)	1,543(0,116)	1,517(0,107)
$\alpha = 5$	4,191(1,598)	4,593(2,619)	4,615(1,244)	4,838(1,391)	4,892(0,944)	5,098(1,004)

Tabela 5 – Estimativas das médias e desvios-padrão dos valores iniciais $\theta^{(0)} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, obtidos através do método 4 e dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, considerando diferentes valores de θ e 20% de censura.

Mét. 4	$n = 50$		$n = 100$		$n = 200$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) EMV	Média(Desvio P.) Val .In	Média(Desvio P.) EMV	Média(Desvio P.) Val .In	Média(Desvio P.) EMV
$\phi = 0,5$	0,831(0,698)	0,975(0,962)	0,639(0,263)	0,697(0,427)	0,561(0,148)	0,571(0,195)
$\mu = 0,5$	0,640(0,909)	0,803(1,314)	0,516(0,106)	0,576(0,235)	0,487(0,051)	0,512(0,070)
$\alpha = 3$	2,747(0,888)	3,123(2,385)	2,947(0,646)	3,038(1,304)	3,082(0,497)	3,099(0,792)
$\phi = 2$	1,081(0,776)	1,745(1,077)	1,098(0,606)	1,809(0,877)	1,118(0,422)	1,884(0,685)
$\mu = 1$	0,511(1,934)	1,807(4,773)	0,371(1,147)	1,409(3,001)	0,261(0,385)	1,119(1,546)
$\alpha = 0,5$	1,055(0,499)	0,841(1,208)	0,921(0,309)	0,641(0,288)	0,842(0,194)	0,573(0,150)
$\phi = 4$	4,802(3,388)	4,410(4,306)	6,147(3,372)	5,724(4,517)	7,902(3,516)	7,778(5,013)
$\mu = 2$	3,501(4,292)	3,297(4,785)	4,683(4,746)	4,294(5,547)	6,967(6,258)	6,553(7,247)
$\alpha = 2$	1,940(0,838)	2,910(2,026)	1,560(0,521)	2,186(1,125)	1,291(0,325)	1,692(0,727)
$\phi = 0,4$	0,918(0,771)	0,622(0,778)	0,768(0,457)	0,462(0,314)	0,678(0,225)	0,413(0,152)
$\mu = 1,5$	1,843(1,263)	1,761(1,314)	1,649(0,457)	1,542(0,268)	1,559(0,170)	1,494(0,111)
$\alpha = 5$	3,541(1,343)	6,715(5,345)	3,602(1,048)	6,009(2,806)	3,604(0,748)	5,667(1,735)

Tabela 6 – Probabilidades de cobertura com o nível de confiança de 95% dos EMV's encontrados para 1000 amostras, utilizando-se os métodos 3 e 4, considerando-se 20% de censura e diferentes valores de θ e n .

θ	Método 3			Método 4		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	96,00%	98,30%	94,50%	94,70%	96,30%	97,60%
$\mu = 0,5$	95,10%	94,50%	90,50%	93,70%	94,40%	94,90%
$\alpha = 3$	94,00%	95,70%	97,30%	89,60%	94,40%	96,40%
$\phi = 2$	64,60%	74,90%	77,10%	73,20%	78,30%	82,40%
$\mu = 1$	43,60%	54,40%	59,80%	59,50%	63,80%	70,30%
$\alpha = 0,5$	94,50%	93,90%	91,60%	92,90%	93,50%	92,60%
$\phi = 4$	68,80%	64,80%	69,60%	56,40%	51,10%	41,20%
$\mu = 2$	63,30%	61,60%	65,70%	55,60%	51,90%	42,60%
$\alpha = 2$	90,60%	88,00%	84,40%	78,10%	63,00%	41,00%
$\phi = 0,4$	97,40%	99,40%	99,30%	86,10%	88,50%	92,50%
$\mu = 1,5$	97,80%	98,80%	98,10%	88,70%	89,90%	90,30%
$\alpha = 5$	92,40%	95,10%	96,80%	88,60%	93,20%	95,80%

Por meio dos resultados obtidos nas Tabelas 4 e 5, conclui-se que ao se utilizar o método 3 para obtenção dos valores iniciais, obtemos melhores estimativas dos EMV's de ϕ, μ e α , considerando censura do tipo aleatória.

É importante salientar que, utilizando-se o método 3, as probabilidades de cobertura dos parâmetros tende a aumentar quando o tamanho amostral cresce, aproximando-se dos valores esperados (95% de confiança).

Através dos resultados pode-se concluir que o método 3 possibilita a obtenção de melhores valores iniciais, e que ao serem utilizados no método iterativo (Newton-Raphson), retornam boas estimativas para os parâmetros ϕ, μ e α .

6.2 Análise Bayesiana

Por meio da abordagem Bayesiana, as escolhas de diferentes distribuições a priori podem levar a inferências bem diferentes. Portanto, serão apresentados resultados simulados, obtidos de densidades a posteriori, utilizando-se diferentes distribuições a priori empíricas e não-informativas, considerando-se dados completos e com censura aleatória.

Aplicando-se os procedimentos descritos no início do capítulo 6 e verificadas as condições de convergência das cadeias, obtemos 1000 estimativas de ϕ, μ e α através de diferentes amostras simuladas.

Para cada amostra simulada foram realizadas 8500 iterações por meio dos métodos MCMC. Como amostra de aquecimento, descartamos 1000 valores iniciais. Buscando-se diminuir as autocorrelações das cadeias, foram salvos valores com espaços de tamanho 15, obtendo-se, no final, três cadeias de tamanho 500 cada. Com estes resultados calculamos

as estimativas das médias a posteriori de ϕ , μ e α .

Para verificar se as convergências das cadeias foram obtidas, combinamos o critério de Geweke, implementado no pacote coda do R, com a taxa de aceitação. Por meio da estatística dada em (1.53), se o valor z_G estiver contido no intervalo $[-1, 96, 1, 96]$, temos indicação de convergência. As taxas de aceitação devem estar também no intervalo de 10% a 60%.

Nos casos em que estes dois critérios foram respeitados, consideramos que as cadeias alcançaram a convergência.

6.2.1 Observações Completas

6.2.1.1 Prioris Empíricas

As densidades a posteriori utilizadas nesta seção, são obtidas considerando-se como distribuição a priori as distribuições Gama e Uniforme, apresentadas na seções 4.1.1 e 4.1.2. A escolha destas distribuições a priori deve-se ao fato que, combinadas com o método dos momentos (4.4 e 4.8), possibilitam facilmente a obtenção de distribuições a priori empíricas.

Utilizando-se as equações (5.8, 5.9 e 5.10), encontramos facilmente bons valores iniciais $\tilde{\theta} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$. Estes valores podem ser utilizados no método dos momentos (4.4 e 4.8) como média $\lambda = \tilde{\theta}$ e como variância $\sigma_{\theta}^2 = \tilde{\theta}$ para obtenção dos hiperparâmetros da distribuição Gama ou Uniforme. Iremos também considerar como variância $\sigma_{\theta}^2 = (1, 1, 1)$ e $\sigma_{\theta}^2 = (10, 10, 10)$, verificando quais destes valores produzem melhores resultados.

As amostras serão geradas no R, utilizando-se a semente 2013. Através do pacote R2OpenBugs o software R irá utilizar o OpenBUGS para gerar as cadeias das distribuições marginais a posteriori de ϕ , μ e α via MCMC. Estas cadeias são utilizadas no R para calcularmos as estimativas das médias a posteriori de ϕ , μ e α e para verificarmos se as cadeias geradas obtiveram a convergência. É importante ressaltar que os valores iniciais $\tilde{\theta}$ também serão aplicados, para a inicialização do algoritmo híbrido, minimizando assim problemas de convergência das cadeias geradas.

Desta forma será possível realizar a interação entre os softwares R e OpenBUGS para se executar o procedimento iterativo, descrito no início da seção 6. O programa desenvolvido pode ser visualizado no Apêndice D.

Utilizando-se a distribuição a priori Gama (empírica) e considerando-se diferentes valores de θ e n nas Tabelas 7-9, temos as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori, encontradas para 1000 amostras simuladas. Nas Tabelas 10-11, podemos ver as probabilidades de cobertura dos parâmetros com o nível de

confiança de 95%.

Tabela 7 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$.

$n = 50$	$\sigma_\theta^2 = (1, 1, 1)$		$\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_\theta^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,674(0,184)	0,741(0,300)	0,676(0,193)	0,729(0,327)	0,667(0,190)	0,893(0,603)
$\mu = 0, 5$	0,572(0,097)	0,624(0,166)	0,572(0,101)	0,623(0,177)	0,568(0,100)	0,879(0,611)
$\alpha = 3$	2,572(0,439)	2,767(0,634)	2,567(0,442)	2,983(0,858)	2,592(0,432)	3,423(1,466)
$\phi = 2$	1,306(0,743)	1,498(0,670)	1,284(0,718)	1,387(0,668)	1,308(0,745)	1,824(0,676)
$\mu = 1$	0,847(2,309)	1,146(2,248)	0,773(2,034)	0,937(2,020)	0,835(2,271)	3,126(2,280)
$\alpha = 0, 5$	0,738(0,216)	0,750(0,272)	0,744(0,218)	0,779(0,283)	0,740(0,217)	0,775(0,540)
$\phi = 4$	1,911(1,213)	3,462(1,026)	1,912(1,280)	2,398(1,316)	1,893(1,304)	3,378(1,497)
$\mu = 2$	1,260(0,546)	1,355(0,561)	1,260(0,563)	1,461(0,635)	1,255(0,582)	2,021(0,840)
$\alpha = 2$	3,462(1,026)	3,321(0,966)	3,486(1,036)	3,245(0,930)	3,526(1,057)	2,910(0,865)
$\phi = 0, 4$	0,629(0,135)	0,641(0,212)	0,629(0,137)	0,636(0,264)	0,625(0,128)	0,706(0,416)
$\mu = 1, 5$	1,683(0,142)	1,701(0,205)	1,684(0,145)	1,712(0,256)	1,678(0,138)	1,837(0,481)
$\alpha = 5$	3,772(0,551)	4,004(0,697)	3,767(0,554)	4,454(1,082)	3,779(0,545)	4,981(1,630)

Tabela 8 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$.

$n = 100$	$\sigma_\theta^2 = (1, 1, 1)$		$\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_\theta^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,684(0,126)	0,678(0,240)	0,685(0,129)	0,661(0,258)	0,679(0,125)	0,690(0,372)
$\mu = 0, 5$	0,569(0,060)	0,583(0,117)	0,569(0,061)	0,578(0,124)	0,568(0,062)	0,635(0,290)
$\alpha = 3$	2,489(0,304)	2,805(0,568)	2,495(0,311)	2,981(0,747)	2,502(0,307)	3,238(1,086)
$\phi = 2$	1,590(0,768)	1,733(0,690)	1,573(0,726)	1,656(0,661)	1,600(0,770)	1,979(0,652)
$\mu = 1$	1,267(2,976)	1,528(2,909)	1,095(2,223)	1,279(2,198)	1,270(2,915)	2,879(2,735)
$\alpha = 0, 5$	0,634(0,164)	0,628(0,173)	0,632(0,158)	0,636(0,162)	0,631(0,165)	0,598(0,168)
$\phi = 4$	2,623(1,744)	2,857(1,700)	2,560(1,730)	3,037(1,725)	2,594(1,810)	3,828(1,880)
$\mu = 2$	1,558(1,009)	1,658(1,009)	1,533(1,026)	1,759(1,087)	1,546(0,971)	2,226(1,160)
$\alpha = 2$	2,958(0,939)	2,859(0,876)	2,984(0,913)	2,809(0,821)	2,998(0,940)	2,600(0,799)
$\phi = 0, 4$	0,636(0,099)	0,590(0,166)	0,635(0,095)	0,555(0,204)	0,635(0,098)	0,552(0,277)
$\mu = 1, 5$	1,680(0,102)	1,652(0,155)	1,680(0,098)	1,634(0,186)	1,679(0,100)	1,646(0,278)
$\alpha = 5$	3,698(0,422)	4,143(0,650)	3,701(0,424)	4,650(1,052)	3,707(0,417)	5,112(1,475)

Tabela 9 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

$n = 200$	$\sigma_{\theta}^2 = (1, 1, 1)$		$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_{\theta}^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,687(0,084)	0,605(0,169)	0,688(0,086)	0,593(0,185)	0,685(0,084)	0,582(0,216)
$\mu = 0, 5$	0,569(0,038)	0,545(0,072)	0,570(0,039)	0,543(0,078)	0,569(0,038)	0,544(0,126)
$\alpha = 3$	2,456(0,216)	2,889(0,498)	2,454(0,224)	2,999(0,629)	2,454(0,221)	3,116(0,750)
$\phi = 2$	1,822(0,708)	1,917(0,639)	1,783(0,673)	1,846(0,613)	1,801(0,722)	2,071(0,626)
$\mu = 1$	1,500(2,887)	1,723(2,824)	1,286(2,113)	1,482(2,092)	1,547(3,191)	2,686(3,062)
$\alpha = 0, 5$	0,566(0,123)	0,559(0,115)	0,572(0,122)	0,572(0,116)	0,571(0,125)	0,546(0,112)
$\phi = 4$	3,276(1,968)	3,491(1,908)	3,186(1,886)	3,638(1,869)	3,201(2,027)	4,176(2,044)
$\mu = 2$	1,854(1,283)	1,956(1,275)	1,802(1,193)	2,046(1,249)	1,836(1,442)	2,409(1,549)
$\alpha = 2$	2,520(0,699)	2,454(0,650)	2,558(0,709)	2,441(0,643)	2,569(0,720)	2,323(0,628)
$\phi = 0, 4$	0,639(0,062)	0,527(0,121)	0,639(0,060)	0,488(0,141)	0,639(0,061)	0,467(0,156)
$\mu = 1, 5$	1,684(0,066)	1,602(0,106)	1,683(0,064)	1,574(0,117)	1,683(0,066)	1,561(0,134)
$\alpha = 5$	3,642(0,292)	4,369(0,622)	3,647(0,288)	4,821(0,958)	3,645(0,287)	5,125(1,223)

Tabela 10 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$.

θ	$\sigma_{\theta}^2 = (1, 1, 1)$			$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0, 5$	94,00%	94,60%	95,50%	95,30%	95,80%	95,40%
$\mu = 0, 5$	93,00%	94,30%	96,20%	94,90%	95,40%	95,20%
$\alpha = 3$	94,90%	94,00%	96,00%	95,70%	95,00%	94,80%
$\phi = 2$	86,60%	84,10%	81,80%	72,20%	78,90%	82,40%
$\mu = 1$	86,10%	84,10%	81,60%	69,60%	79,50%	82,00%
$\alpha = 0, 5$	85,80%	83,90%	82,50%	73,70%	80,20%	83,20%
$\phi = 4$	33,70%	47,70%	57,20%	55,70%	67,10%	74,30%
$\mu = 2$	35,70%	48,20%	57,60%	56,10%	68,00%	74,50%
$\alpha = 2$	40,50%	49,90%	58,80%	58,10%	67,90%	74,90%
$\phi = 0, 4$	89,30%	88,20%	89,90%	93,70%	94,40%	93,70%
$\mu = 1, 5$	88,90%	88,10%	89,90%	92,40%	92,60%	93,20%
$\alpha = 5$	86,30%	85,80%	88,80%	93,70%	93,70%	93,20%

Tabela 11 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Gama com $\sigma_\theta^2 = (10, 10, 10)$.

	$\sigma_\theta^2 = (10, 10, 10)$		
θ	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	96,60%	95,50%	95,30%
$\mu = 0,5$	95,20%	95,60%	95,60%
$\alpha = 3$	96,00%	95,30%	95,40%
$\phi = 2$	93,90%	91,50%	90,90%
$\mu = 1$	93,80%	91,90%	90,40%
$\alpha = 0,5$	93,80%	92,40%	91,40%
$\phi = 4$	93,30%	87,70%	85,70%
$\mu = 2$	91,80%	87,70%	85,70%
$\alpha = 2$	92,60%	87,70%	86,10%
$\phi = 0,4$	95,80%	94,90%	94,90%
$\mu = 1,5$	94,90%	94,40%	94,80%
$\alpha = 5$	96,60%	95,30%	94,60%

Observa-se através das Tabelas 7-9 que, utilizando-se os valores calculados através das equações (5.8, 5.9 e 5.10) no método dos momentos (4.8) como média $\lambda = \tilde{\theta}$ e como variância $\sigma_\theta^2 = \tilde{\theta}$ para obtenção dos hiperparâmetros da distribuição Gama, obtemos boas estimativas das médias a posteriori de ϕ, μ e α , por meio da abordagem Bayesiana. Consideramos também diferentes valores para σ_θ^2 e, embora todos tenham retornado boas estimativas, o que se mostrou mais eficiente foi quando utilizou-se como variância $\sigma_\theta^2 = \tilde{\theta}$.

É importante salientar que utilizando-se $\sigma_\theta^2 = \tilde{\theta}$, as probabilidades de cobertura dos parâmetros tendem a crescer para 95%, conforme aumenta-se o tamanho de n .

Empregando-se a distribuição a priori uniforme conjunta, os valores dos hiperparâmetros podem ser obtidos através das equações (5.8, 5.9 e 5.10) e do método dos momentos (4.4), utilizando-se como média $\lambda = \tilde{\theta}$, como variância $\sigma_\theta^2 = \tilde{\theta}$, $\sigma_\theta^2 = (1, 1, 1)$ e $\sigma_\theta^2 = (10, 10, 10)$.

Nas Tabelas 12-14, temos as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas, utilizando-se a distribuição a priori Uniforme (empírica) e considerando-se diferentes valores de θ e n . Nas Tabelas 15-16, podemos ver as probabilidades de cobertura dos parâmetros com o nível de confiança de 95%.

Tabela 12 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$.

$n = 50$	$\sigma_\theta^2 = (1, 1, 1)$		$\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_\theta^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,675(0,197)	1,067(0,414)	0,673(0,187)	0,957(0,423)	0,670(0,192)	1,635(1,009)
$\mu = 0, 5$	0,572(0,104)	0,809(0,252)	0,569(0,099)	0,744(0,250)	0,568(0,101)	1,508(0,886)
$\alpha = 3$	2,581(0,447)	2,328(0,645)	2,571(0,437)	2,595(0,812)	2,587(0,442)	2,563(1,310)
$\phi = 2$	1,293(0,745)	1,933(0,545)	1,286(0,730)	1,690(0,614)	1,304(0,725)	2,544(0,459)
$\mu = 1$	0,845(2,353)	1,442(2,236)	0,826(2,261)	1,158(2,249)	0,808(2,159)	3,050(1,864)
$\alpha = 0, 5$	0,741(0,216)	0,583(0,142)	0,745(0,222)	0,649(0,192)	0,734(0,215)	0,500(0,179)
$\phi = 4$	1,916(1,242)	2,440(1,180)	1,894(1,256)	2,767(1,429)	1,907(1,214)	4,557(1,311)
$\mu = 2$	1,259(0,548)	1,451(0,562)	1,254(0,544)	1,608(0,710)	1,256(0,507)	2,518(0,841)
$\alpha = 2$	3,472(1,029)	3,060(0,832)	3,503(1,041)	2,936(0,795)	3,470(1,029)	2,207(0,510)
$\phi = 0, 4$	0,629(0,136)	0,825(0,315)	0,632(0,138)	0,823(0,361)	0,624(0,133)	1,106(0,700)
$\mu = 1, 5$	1,684(0,146)	1,880(0,315)	1,686(0,146)	1,899(0,367)	1,679(0,146)	2,324(0,823)
$\alpha = 5$	3,752(0,556)	3,591(0,776)	3,762(0,556)	3,941(1,041)	3,770(0,543)	4,097(1,513)

Tabela 13 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$.

$n = 100$	$\sigma_\theta^2 = (1, 1, 1)$		$\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_\theta^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,684(0,131)	0,862(0,363)	0,685(0,138)	0,807(0,380)	0,681(0,137)	1,002(0,704)
$\mu = 0, 5$	0,569(0,063)	0,685(0,200)	0,570(0,066)	0,658(0,203)	0,569(0,067)	0,894(0,588)
$\alpha = 3$	2,495(0,306)	2,559(0,612)	2,493(0,311)	2,776(0,789)	2,506(0,316)	2,895(1,113)
$\phi = 2$	1,574(0,747)	2,065(0,527)	1,571(0,750)	1,901(0,620)	1,591(0,756)	2,566(0,460)
$\mu = 1$	1,182(2,671)	1,740(2,527)	1,177(2,782)	1,557(2,756)	1,219(2,793)	3,283(2,437)
$\alpha = 0, 5$	0,634(0,161)	0,532(0,096)	0,635(0,160)	0,567(0,120)	0,632(0,161)	0,471(0,089)
$\phi = 4$	2,606(1,755)	3,078(1,656)	2,578(1,713)	3,545(1,892)	2,663(1,839)	5,062(1,899)
$\mu = 2$	1,543(0,961)	1,737(0,956)	1,534(0,982)	1,987(1,172)	1,601(1,322)	2,864(1,534)
$\alpha = 2$	2,955(0,917)	2,677(0,721)	2,982(0,927)	2,566(0,751)	2,926(0,908)	2,084(0,567)
$\phi = 0, 4$	0,634(0,095)	0,685(0,244)	0,636(0,097)	0,640(0,281)	0,634(0,096)	0,682(0,432)
$\mu = 1, 5$	1,679(0,097)	1,743(0,233)	1,680(0,100)	1,715(0,270)	1,680(0,098)	1,800(0,481)
$\alpha = 5$	3,705(0,414)	3,912(0,686)	3,698(0,417)	4,398(1,007)	3,706(0,422)	4,744(1,368)

Tabela 14 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$.

$n = 200$	$\sigma_\theta^2 = (1, 1, 1)$		$\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_\theta^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,685(0,084)	0,656(0,231)	0,683(0,081)	0,633(0,239)	0,685(0,083)	0,656(0,303)
$\mu = 0, 5$	0,568(0,039)	0,573(0,110)	0,568(0,038)	0,565(0,111)	0,569(0,038)	0,588(0,183)
$\alpha = 3$	2,459(0,224)	2,820(0,533)	2,461(0,220)	2,965(0,668)	2,456(0,220)	3,005(0,787)
$\phi = 2$	1,796(0,684)	2,155(0,493)	1,782(0,684)	2,056(0,561)	1,815(0,707)	2,568(0,458)
$\mu = 1$	1,418(2,775)	1,925(2,639)	1,389(2,657)	1,811(2,617)	1,510(2,997)	3,401(2,684)
$\alpha = 0, 5$	0,567(0,118)	0,505(0,070)	0,570(0,120)	0,523(0,085)	0,566(0,120)	0,458(0,063)
$\phi = 4$	3,244(1,953)	3,671(1,840)	3,111(1,849)	4,102(1,949)	3,262(1,935)	5,436(2,100)
$\mu = 2$	1,834(1,258)	2,025(1,243)	1,764(1,189)	2,263(1,318)	1,845(1,377)	3,098(1,630)
$\alpha = 2$	2,544(0,716)	2,367(0,577)	2,588(0,699)	2,273(0,584)	2,536(0,718)	1,968(0,541)
$\phi = 0, 4$	0,639(0,061)	0,554(0,145)	0,638(0,058)	0,501(0,164)	0,639(0,062)	0,491(0,193)
$\mu = 1, 5$	1,683(0,065)	1,626(0,133)	1,681(0,063)	1,588(0,143)	1,683(0,066)	1,585(0,179)
$\alpha = 5$	3,643(0,293)	4,253(0,572)	3,646(0,287)	4,767(0,873)	3,643(0,290)	5,023(1,134)

Tabela 15 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$.

θ	$\sigma_\theta^2 = (1, 1, 1)$			$\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0, 5$	88,40%	89,20%	94,10%	91,40%	90,60%	94,70%
$\mu = 0, 5$	85,80%	89,80%	94,20%	89,30%	90,70%	93,60%
$\alpha = 3$	89,30%	90,60%	94,30%	92,60%	91,20%	94,40%
$\phi = 2$	93,90%	91,30%	89,20%	67,70%	76,10%	83,80%
$\mu = 1$	93,30%	90,50%	88,70%	45,50%	66,80%	82,20%
$\alpha = 0, 5$	92,40%	90,80%	88,80%	74,50%	81,40%	85,60%
$\phi = 4$	23,60%	38,10%	50,60%	37,70%	57,30%	71,90%
$\mu = 2$	32,30%	43,90%	54,00%	43,80%	59,50%	72,90%
$\alpha = 2$	46,00%	51,90%	59,60%	55,70%	65,70%	75,70%
$\phi = 0, 4$	78,90%	82,20%	85,50%	89,40%	92,50%	95,50%
$\mu = 1, 5$	80,80%	83,50%	87,20%	87,70%	91,10%	94,80%
$\alpha = 5$	68,50%	75,60%	81,40%	90,80%	93,80%	94,70%

Tabela 16 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori Uniforme com $\sigma_\theta^2 = (10, 10, 10)$.

	$\sigma_\theta^2 = (10, 10, 10)$		
θ	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	82,70%	88,30%	92,20%
$\mu = 0,5$	82,10%	89,10%	92,20%
$\alpha = 3$	82,10%	89,10%	92,20%
$\phi = 2$	96,90%	93,30%	91,80%
$\mu = 1$	97,30%	94,50%	91,20%
$\alpha = 0,5$	97,50%	94,90%	92,00%
$\phi = 4$	98,80%	93,30%	84,80%
$\mu = 2$	98,50%	93,10%	85,10%
$\alpha = 2$	98,30%	92,90%	85,60%
$\phi = 0,4$	88,20%	93,00%	93,60%
$\mu = 1,5$	87,60%	92,60%	93,60%
$\alpha = 5$	89,20%	94,00%	94,10%

Similarmente aos resultados obtidos utilizando-se a distribuição a priori Gama, observa-se através das Tabelas 12-14 que, utilizando-se os valores calculados através das equações (5.8, 5.9 e 5.10) no método dos momentos (4.4) como $\lambda = \tilde{\theta}$ e como variância $\sigma_\theta^2 = \tilde{\theta}$ para obtenção dos hiperparâmetros da distribuição Uniforme, obtemos boas estimativas das médias a posteriori de ϕ, μ e α , por meio da abordagem Bayesiana.

É importante salientar que, utilizando-se $\lambda = \tilde{\theta}$, as probabilidades de cobertura dos parâmetros tendem a crescer para 95%, conforme aumenta-se o tamanho de n .

No entanto, se compararmos os resultados obtidos empregando-se as duas distribuições a priori conjuntas (Gama e Uniforme), verificamos que a distribuição a priori Gama proporciona melhores estimativas das médias a posteriori de ϕ, μ e α quando $\lambda = \tilde{\theta}$ e $\sigma_\theta^2 = \tilde{\theta}$.

6.2.1.2 Prioris Não-Informativas

É de grande interesse prático o uso de distribuições a priori não-informativas ao se realizar uma análise Bayesiana, principalmente nos casos em que não dispomos de informação a priori sobre os parâmetros de interesse.

No entanto, há grande dificuldade em se estimar os parâmetros da distribuição Gama Generalizada, utilizando-se tais distribuições a priori, sendo que, em muitos casos, estas distribuições podem levar a inferências bem diferentes das esperadas.

Por exemplo, utilizando-se o OpenBUGS para gerar as cadeias das distribuições marginais a posteriori de ϕ, μ e α via MCMC, na Tabela 17, obtivemos as médias e os desvios-padrão das estimativas das médias a posteriori encontradas para 1000 amostras simuladas, considerando-se diferentes valores de θ e n .

Como distribuição a priori consideramos no primeiro caso que $\theta_i \sim \text{Uniforme}(0, 40)$, $i = 1, \dots, 3$ e no segundo $\theta_i \sim \text{Gamma}(0.01, 0.01)$. As frequências das probabilidades de cobertura com o nível de confiança de 95% também estão disponíveis na Tabela 18.

Tabela 17 – Médias e desvios-padrão das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e considerando as distribuição a priori $\theta_i \sim \text{Uniforme}(0, 40)$ e $\theta_i \sim \text{Gamma}(0.01, 0.01)$.

θ	$\theta_i \sim \text{Uniforme}(0, 40)$			$\theta_i \sim \text{Gamma}(0.01, 0.01)$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0, 5$	2,945(2,255)	1,264(1,325)	0,669(0,409)	1,052(0,934)	0,719(0,499)	0,580(0,222)
$\mu = 0, 5$	7,454(6,442)	2,114(3,490)	0,646(0,708)	1,785(2,534)	0,772(1,079)	0,549(0,193)
$\alpha = 3$	3,465(3,673)	3,141(1,740)	3,033(0,867)	8,762(44,485)	3,955(8,192)	3,177(0,916)
$\phi = 2$	3,768(0,492)	3,647(0,461)	3,489(0,541)	2,071(0,801)	2,218(0,771)	2,301(0,765)
$\mu = 1$	17,498(2,174)	16,497(2,976)	14,824(4,213)	8,349(6,313)	8,487(7,385)	7,836(8,336)
$\alpha = 0, 5$	0,389(0,119)	0,383(0,063)	0,391(0,064)	1,296(9,974)	0,584(0,189)	0,531(0,119)
$\phi = 4$	12,174(4,141)	9,902(4,541)	7,406(3,935)	5,230(3,067)	5,147(3,008)	5,035(2,662)
$\mu = 2$	13,923(5,260)	10,279(6,170)	6,276(5,051)	5,359(4,587)	4,536(4,236)	3,664(3,247)
$\alpha = 2$	1,737(1,652)	1,783 (1,132)	1,859(0,676)	3,521(4,630)	2,540(0,966)	2,247(0,668)
$\phi = 0, 4$	1,807(1,833)	0,773 0,813)	0,488(0,200)	0,864(0,845)	0,548(0,363)	0,452(0,168)
$\mu = 1, 5$	5,246(5,029)	2,210 2,073)	1,587(0,214)	2,634(2,582)	1,730(0,853)	1,551(0,143)
$\alpha = 5$	6,177(4,707)	5,500 2,853)	5,211(1,536)	9,773(11,967)	6,348(4,353)	5,426(1,680)

Tabela 18 – Probabilidades de cobertura com o nível de confiança de 95% das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$ e considerando as distribuição a priori $\theta_i \sim \text{Uniforme}(0, 40)$ e $\theta_i \sim \text{Gamma}(0.01, 0.01)$.

θ	$\theta_i \sim \text{Uniforme}(0, 40)$			$\theta_i \sim \text{Gamma}(0.01, 0.01)$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0, 5$	72,80%	88,70%	92,80%	93,10%	94,70%	95,40%
$\mu = 0, 5$	72,80%	88,10%	92,40%	93,40%	94,50%	96,20%
$\alpha = 3$	71,70%	88,70%	92,70%	93,60%	94,60%	95,10%
$\phi = 2$	74,50%	72,40%	68,10%	97,40%	97,60%	96,10%
$\mu = 1$	73,30%	71,50%	67,00%	97,80%	97,80%	96,50%
$\alpha = 0, 5$	78,80%	74,20%	69,80%	97,60%	97,80%	96,50%
$\phi = 4$	68,10%	68,00%	76,10%	95,70%	95,60%	92,40%
$\mu = 2$	68,30%	67,70%	75,80%	95,40%	95,90%	91,50%
$\alpha = 2$	68,80%	68,50%	75,80%	95,00%	95,50%	91,40%
$\phi = 0, 4$	84,60%	91,10%	93,80%	93,50%	94,30%	94,20%
$\mu = 1, 5$	84,70%	90,60%	94,00%	94,20%	93,70%	94,30%
$\alpha = 5$	85,40%	91,30%	93,60%	94,50%	94,60%	94,10%

Pode-se observar pelos resultados demonstrados na Tabela 17, que as estimativas das médias a posteriori são muito influenciadas pelas distribuições a priori não-informativas, levando-se a inferências bem diferentes das esperadas. Por exemplo, para $\mu = 2$, $n = 50$ e utilizando-se como distribuição a priori $\theta_2 \sim \text{Uniforme}(0, 40)$, a média das 1000 estimativas é de $\mu = 13,923$, com desvio-padrão $\sigma_\mu = 5,260$.

Outras distribuições a priori não-informativas podem ser utilizadas para se realizar inferências nos parâmetros da distribuição Gama Generalizada. Dentre elas podemos citar

as prioris: de Jeffreys (4.12), a MDIP (4.16) e a obtida pela Regra de Jeffreys (4.9). No entanto demonstramos que, tais distribuições a priori levam a distribuições a posteriori impróprias e não devem ser utilizadas em uma análise Bayesiana.

Uma modificação na priori obtida através da Regra de Jeffreys, foi realizada e demonstramos que utilizando-se tal distribuição a priori, obtemos uma densidade a posteriori própria.

Contudo, para se alcançar a convergência das cadeias através dos métodos MCMC, utilizando-se as densidades a posteriori condicionais (4.22), é necessário que durante a realização do MCMC, os valores propostos $\theta^{(*)}$ das distribuições auxiliares sejam obtidos independentemente das posições anteriores de θ . Desta forma, a probabilidade de aceitação passa a ser calculada por meio de (1.52).

Os valores das cadeias serão gerados através do algoritmo descrito na seção 4.3. As distribuições auxiliares, utilizadas para gerar os possíveis valores das cadeias de α , ϕ e μ , serão respectivamente $Nt(\tilde{\alpha}, \sigma_\alpha)$, $Nt(\tilde{\phi}, \sigma_\phi)$ e $Nt(\tilde{\mu}, \sigma_\mu)$. Os hiperparâmetros $\tilde{\theta} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ serão obtidos através das equações (5.8, 5.9 e 5.10) e os valores de $\sigma = (\sigma_\alpha, \sigma_\phi, \sigma_\mu)$ foram selecionados de forma que a taxa de aceitação fique entre 10% e 60%.

Por meio da Tabela 19, temos as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas, utilizando-se a densidade a posteriori, obtida através Regra de Jeffreys Modificada, considerando-se diferentes valores de θ e n . Através da Tabela 20, podemos ver as probabilidades de cobertura dos parâmetros com o nível de confiança de 95%.

Tabela 19 – Médias e desvios-padrão das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$, utilizando-se a distribuição a priori obtida pela Regra de Jeffreys Modificada e considerando-se diferentes valores de θ e n .

	$n = 50$		$n = 100$		$n = 200$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,692(0,237)	0,733(0,312)	0,719(0,137)	0,716(0,214)	0,690(0,091)	0,606(0,149)
$\mu = 0, 5$	0,583(0,141)	0,605(0,161)	0,589(0,064)	0,591(0,092)	0,570(0,041)	0,541(0,060)
$\alpha = 3$	2,482(0,371)	2,462(0,543)	2,373(0,249)	2,429(0,399)	2,440(0,209)	2,709(0,392)
$\phi = 2$	1,262(0,617)	1,416(0,512)	1,485(0,548)	1,601(0,435)	1,523(0,356)	1,573(0,287)
$\mu = 1$	0,526(0,774)	0,565(0,678)	0,650(0,690)	0,682(0,583)	0,548(0,309)	0,564(0,249)
$\alpha = 0, 5$	0,739(0,206)	0,660(0,149)	0,638(0,143)	0,591(0,097)	0,603(0,091)	0,585(0,068)
$\phi = 4$	1,859(1,117)	2,066(1,089)	2,676(1,499)	2,793(1,451)	3,581(1,142)	3,628(1,116)
$\mu = 2$	1,232(0,435)	1,300(0,429)	1,538(0,669)	1,578(0,652)	1,865(0,508)	1,885(0,498)
$\alpha = 2$	3,432(0,970)	3,164(0,868)	2,780(0,720)	2,678(0,666)	2,201(0,359)	2,177(0,339)
$\phi = 0, 4$	0,627(0,133)	0,584(0,234)	0,639(0,101)	0,532(0,186)	0,647(0,062)	0,512(0,117)
$\mu = 1, 5$	1,683(0,145)	1,658(0,208)	1,685(0,104)	1,609(0,159)	1,689(0,066)	1,589(0,099)
$\alpha = 5$	3,767(0,554)	4,149(1,029)	3,695(0,410)	4,376(0,938)	3,614(0,277)	4,313(0,603)

Tabela 20 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a priori obtida pela Regra de Jeffreys Modificada.

θ	Priori Regra de Jeffreys Modificada		
	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	93,90%	91,00%	90,20%
$\mu = 0,5$	92,60%	89,80%	91,30%
$\alpha = 3$	93,20%	90,70%	92,00%
$\phi = 2$	75,00%	80,60%	73,90%
$\mu = 1$	73,80%	81,20%	73,20%
$\alpha = 0,5$	76,90%	82,40%	75,90%
$\phi = 4$	21,10%	27,70%	39,60%
$\mu = 2$	20,00%	24,50%	39,30%
$\alpha = 2$	30,00%	31,60%	51,30%
$\phi = 0,4$	94,10%	91,60%	88,30%
$\mu = 1,5$	92,70%	90,90%	88,80%
$\alpha = 5$	93,50%	91,90%	88,70%

Observa-se através da Tabela 19, que utilizando-se os valores iniciais obtidos através das equações (5.8, 5.9 e 5.10) nas distribuições auxiliares para gerar os possíveis valores das cadeias dos parâmetros de interesse, obtemos boas estimativas da média a posteriori de ϕ, μ e α .

As probabilidades de cobertura dos parâmetros mostraram que, embora alguns casos estejam abaixo do ideal, em muitos estas são superiores a 90%. No entanto pode-se observar que para alguns valores de θ as probabilidades de cobertura, ou estão muito baixas, ou diminuem quando o tamanho amostral aumenta.

Desta forma, ao compararmos estes resultados com os obtidos utilizando-se as distribuições a priori empíricas (Gama e Uniforme), podemos concluir que a distribuição a priori Gama proporciona melhores estimativas das médias a posteriori ϕ, μ e α quando $\lambda = \tilde{\theta}$ e $\sigma_{\theta}^2 = \tilde{\theta}$. Outra vantagem do uso da distribuição a priori Gama, é a facilidade em se implementar computacionalmente o modelo no OPENBUGS.

6.2.2 Observações Censuradas

As densidades a posteriori utilizadas nesta seção são obtidas considerando-se como distribuição a priori as distribuições Gama e Uniforme apresentadas na seções 4.2.1 e 4.2.2. Similarmente ao caso em que consideramos observações completas, a escolha destas distribuições a priori devem-se ao fato de que, combinadas com o método dos momentos (4.4 e 4.8), possibilitam facilmente a obtenção de distribuições a priori empíricas.

Removendo-se as observações censuradas e utilizando-se as equações (5.8, 5.9 e 5.10), encontramos facilmente bons valores iniciais $\tilde{\theta} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$. Embora haja perda de informação ao se remover tais observações censuradas, esta abordagem mostrou-se muito

eficiente para se obter as estimativas a posteriori de ϕ, μ e α .

Desta forma, podemos utilizar os valores $\tilde{\phi}, \tilde{\mu}$ e $\tilde{\alpha}$ no método dos momentos (4.4 e 4.8) como média $\lambda = \tilde{\theta}$ e como variância $\sigma_{\theta}^2 = \tilde{\theta}$ para obtenção dos hiperparâmetros da distribuição Gama ou Uniforme. Iremos também considerar como variância $\sigma_{\theta}^2 = (1, 1, 1)$ e $\sigma_{\theta}^2 = (10, 10, 10)$, verificando quais destes valores produzem melhores resultados.

Assim como na seção 6.2.1.1, as amostras serão geradas no R, utilizando-se a semente 2013. Será realizada a interação entre os softwares R e OpenBUGS para gerar as cadeias das distribuições marginais a posteriori de ϕ, μ e α via MCMC. É importante ressaltar que os valores iniciais $\tilde{\theta}$ também serão aplicados para a inicialização do método MCMC. O programa desenvolvido pode ser visualizado no Apêndice D.

Utilizando-se a distribuição a priori Gama (empírica) e considerando-se diferentes valores de θ , n e 20% de censura nas Tabelas 21-23 temos as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas. Nas Tabelas 24-25, podemos ver as probabilidades de cobertura dos parâmetros com o nível de confiança de 95%.

Tabela 21 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

$n = 50$	$\sigma_{\theta}^2 = (1, 1, 1)$		$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_{\theta}^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0,5$	0,656(0,186)	0,757(0,302)	0,662(0,196)	0,744(0,330)	0,653(0,187)	0,941(0,638)
$\mu = 0,5$	0,564(0,096)	0,602(0,152)	0,567(0,101)	0,603(0,166)	0,563(0,100)	0,897(0,625)
$\alpha = 3$	2,635(0,465)	2,840(0,636)	2,623(0,471)	3,098(0,869)	2,648(0,462)	3,657(1,573)
$\phi = 2$	1,220(0,737)	1,456(0,707)	1,202(0,710)	1,348(0,699)	1,238(0,755)	1,792(0,740)
$\mu = 1$	0,813(2,742)	1,057(2,698)	0,768(2,564)	0,880(2,559)	0,882(2,887)	3,072(2,867)
$\alpha = 0,5$	0,774(0,233)	0,819(0,321)	0,776(0,233)	0,845(0,332)	0,769(0,233)	0,931(0,846)
$\phi = 4$	1,701(1,093)	1,969(1,088)	1,679(1,104)	2,166(1,142)	1,680(1,074)	3,254(1,366)
$\mu = 2$	1,177(0,436)	1,253(0,445)	1,175(0,479)	1,342(0,529)	1,174(0,439)	1,944(0,750)
$\alpha = 2$	3,641(1,014)	3,505(0,957)	3,668(1,024)	3,431(0,920)	3,678(1,035)	3,026(0,849)
$\phi = 0,4$	0,622(0,159)	0,669(0,229)	0,624(0,165)	0,665(0,284)	0,617(0,148)	0,759(0,445)
$\mu = 1,5$	1,682(0,173)	1,688(0,223)	1,685(0,181)	1,704(0,279)	1,677(0,162)	1,860(0,503)
$\alpha = 5$	3,831(0,610)	4,041(0,718)	3,824(0,614)	4,503(1,055)	3,844(0,606)	5,060(1,595)

Tabela 22 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

$n = 100$	$\sigma_{\theta}^2 = (1, 1, 1)$		$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_{\theta}^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,694(0,180)	0,727(0,287)	0,694(0,178)	0,703(0,310)	0,688(0,161)	0,762(0,471)
$\mu = 0, 5$	0,575(0,094)	0,583(0,141)	0,576(0,094)	0,578(0,153)	0,572(0,080)	0,672(0,394)
$\alpha = 3$	2,505(0,360)	2,843(0,606)	2,501(0,357)	3,070(0,830)	2,516(0,359)	3,447(1,338)
$\phi = 2$	1,518(0,808)	1,710(0,775)	1,511(0,796)	1,642(0,778)	1,540(0,838)	1,984(0,763)
$\mu = 1$	1,297(3,354)	1,487(3,303)	1,289(3,472)	1,396(3,444)	1,432(3,662)	2,967(3,462)
$\alpha = 0, 5$	0,662(0,182)	0,675(0,222)	0,662(0,179)	0,687(0,220)	0,660(0,182)	0,651(0,235)
$\phi = 4$	2,411(1,623)	2,667(1,586)	2,363(1,605)	2,850(1,618)	2,455(1,770)	3,781(1,904)
$\mu = 2$	1,454(0,809)	1,534(0,807)	1,438(0,796)	1,633(0,859)	1,494(0,969)	2,186(1,165)
$\alpha = 2$	3,092(0,964)	3,004(0,905)	3,131(0,980)	2,965(0,887)	3,106(1,007)	2,690(0,861)
$\phi = 0, 4$	0,632(0,102)	0,625(0,176)	0,629(0,099)	0,581(0,211)	0,629(0,098)	0,591(0,310)
$\mu = 1, 5$	1,680(0,104)	1,645(0,155)	1,676(0,101)	1,621(0,179)	1,677(0,101)	1,657(0,309)
$\alpha = 5$	3,726(0,451)	4,146(0,642)	3,738(0,448)	4,736(1,056)	3,734(0,444)	5,274(1,573)

Tabela 23 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

$n = 200$	$\sigma_{\theta}^2 = (1, 1, 1)$		$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_{\theta}^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,684(0,095)	0,626(0,188)	0,683(0,092)	0,602(0,196)	0,684(0,094)	0,597(0,246)
$\mu = 0, 5$	0,569(0,044)	0,536(0,076)	0,569(0,044)	0,528(0,078)	0,568(0,044)	0,534(0,119)
$\alpha = 3$	2,466(0,249)	2,982(0,583)	2,471(0,250)	3,155(0,738)	2,467(0,249)	3,357(0,998)
$\phi = 2$	1,766(0,745)	1,903(0,732)	1,741(0,734)	1,847(0,722)	1,752(0,773)	2,047(0,726)
$\mu = 1$	1,528(3,200)	1,648(3,156)	1,461(3,003)	1,551(2,979)	1,636(3,525)	2,577(3,390)
$\alpha = 0, 5$	0,584(0,136)	0,592(0,149)	0,586(0,135)	0,599(0,147)	0,588(0,141)	0,581(0,153)
$\phi = 4$	3,086(2,008)	3,336(1,966)	3,040(2,005)	3,521(2,008)	3,063(2,058)	4,170(2,152)
$\mu = 2$	1,792(1,424)	1,877(1,422)	1,778(1,476)	1,997(1,563)	1,799(1,565)	2,409(1,711)
$\alpha = 2$	2,665(0,802)	2,604(0,752)	2,687(0,800)	2,569(0,731)	2,682(0,802)	2,413(0,715)
$\phi = 0, 4$	0,637(0,066)	0,559(0,127)	0,637(0,066)	0,510(0,153)	0,634(0,063)	0,478(0,167)
$\mu = 1, 5$	1,681(0,070)	1,593(0,106)	1,681(0,070)	1,563(0,123)	1,678(0,069)	1,542(0,135)
$\alpha = 5$	3,659(0,314)	4,370(0,605)	3,658(0,307)	4,913(0,976)	3,665(0,308)	5,351(1,297)

Tabela 24 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$.

θ	$\sigma_\theta^2 = (1, 1, 1)$			$\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	95,10%	92,20%	94,70%	95,70%	93,50%	94,80%
$\mu = 0,5$	96,40%	95,70%	96,00%	97,10%	95,70%	94,70%
$\alpha = 3$	96,00%	94,50%	96,20%	97,20%	95,70%	95,90%
$\phi = 2$	86,80%	82,00%	80,00%	71,50%	78,00%	78,40%
$\mu = 1$	84,60%	80,20%	77,90%	62,90%	69,70%	75,10%
$\alpha = 0,5$	85,50%	82,40%	79,40%	72,80%	76,30%	78,90%
$\phi = 4$	28,60%	45,20%	53,10%	50,20%	64,40%	70,70%
$\mu = 2$	26,80%	43,30%	50,40%	50,00%	63,80%	70,10%
$\alpha = 2$	32,80%	45,50%	51,50%	54,00%	65,10%	70,60%
$\phi = 0,4$	87,80%	85,20%	86,20%	94,10%	93,70%	94,30%
$\mu = 1,5$	91,10%	91,20%	93,30%	94,90%	96,00%	95,90%
$\alpha = 5$	87,30%	88,60%	91,40%	94,80%	95,50%	96,00%

Tabela 25 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Gama com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$.

θ	$\sigma_\theta^2 = (10, 10, 10)$		
	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	96,30%	94,00%	93,80%
$\mu = 0,5$	96,70%	94,80%	94,40%
$\alpha = 3$	97,70%	95,20%	93,90%
$\phi = 2$	93,10%	90,70%	86,80%
$\mu = 1$	93,20%	90,60%	86,40%
$\alpha = 0,5$	93,70%	91,30%	87,40%
$\phi = 4$	94,20%	88,50%	84,40%
$\mu = 2$	94,00%	87,80%	83,50%
$\alpha = 2$	94,50%	88,40%	83,50%
$\phi = 0,4$	96,70%	95,60%	95,50%
$\mu = 1,5$	96,60%	96,60%	95,40%
$\alpha = 5$	97,10%	97,20%	95,60%

Através dos resultados obtidos nas Tabelas 21-23 observa-se que, utilizando-se a abordagem descrita no início da seção para obtenção dos hiperparâmetros da distribuição Gama, obtemos boas estimativas das médias a posteriori ϕ, μ e α por meio da abordagem Bayesiana, considerando-se observações censuradas. Consideramos também diferentes valores para σ_θ^2 e, embora todos tenham retornado boas estimativas, o que se mostrou mais eficiente foi o que teve como variância $\sigma_\theta^2 = \tilde{\theta}$.

Similarmente ao caso em que consideramos observações completas, utilizando-se $\sigma_\theta^2 = \tilde{\theta}$, as probabilidades de cobertura dos parâmetros tendem a crescer para 95%, conforme aumenta-se o tamanho de n .

Considerando-se a distribuição a priori uniforme conjunta, os valores dos hiperpa-

râmetros podem ser obtidos semelhantemente ao método anterior, só que considerando-se o método dos momentos (4.4) e utilizando-se como média $\sigma_{\theta}^2 = \tilde{\theta}$ e como variância $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

Nas Tabelas 26-28, temos as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas, utilizando-se a distribuição a priori Uniforme (empírica) e considerando-se diferentes valores de θ , n e 20% de censura. Nas Tabelas 29-30, podemos ver as probabilidades de cobertura dos parâmetros com o nível de confiança de 95%.

Tabela 26 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 50$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

$n = 50$	$\sigma_{\theta}^2 = (1, 1, 1)$		$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_{\theta}^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,659(0,200)	1,088(0,400)	0,657(0,201)	0,968(0,431)	0,665(0,209)	1,823(1,065)
$\mu = 0, 5$	0,565(0,105)	0,776(0,227)	0,565(0,107)	0,713(0,238)	0,566(0,108)	1,606(0,878)
$\alpha = 3$	2,640(0,470)	2,372(0,643)	2,632(0,464)	2,679(0,814)	2,619(0,473)	2,553(1,358)
$\phi = 2$	1,205(0,715)	1,959(0,550)	1,221(0,745)	1,730(0,671)	1,220(0,722)	2,680(0,497)
$\mu = 1$	0,778(2,660)	1,324(2,571)	0,834(2,796)	1,133(2,779)	0,771(2,607)	3,022(2,323)
$\alpha = 0, 5$	0,779(0,234)	0,599(0,158)	0,774(0,232)	0,669(0,203)	0,770(0,225)	0,496(0,151)
$\phi = 4$	1,716(1,148)	2,263(1,100)	1,687(1,068)	2,539(1,253)	1,754(1,215)	4,499(1,260)
$\mu = 2$	1,189(0,496)	1,360(0,505)	1,174(0,441)	1,480(0,585)	1,213(0,781)	2,450(0,952)
$\alpha = 2$	3,644(1,032)	3,205(0,841)	3,661(1,020)	3,075(0,783)	3,595(1,033)	2,225(0,481)
$\phi = 0, 4$	0,623(0,160)	0,865(0,327)	0,624(0,157)	0,862(0,364)	0,623(0,168)	1,251(0,752)
$\mu = 1, 5$	1,684(0,179)	1,871(0,325)	1,682(0,171)	1,888(0,367)	1,682(0,184)	2,434(0,861)
$\alpha = 5$	3,815(0,619)	3,615(0,793)	3,825(0,617)	3,950(1,008)	3,821(0,623)	4,004(1,480)

Tabela 27 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 100$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

$n = 100$	$\sigma_{\theta}^2 = (1, 1, 1)$		$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_{\theta}^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,689(0,176)	0,925(0,396)	0,690(0,177)	0,863(0,418)	0,684(0,178)	1,156(0,833)
$\mu = 0, 5$	0,573(0,092)	0,686(0,211)	0,574(0,095)	0,660(0,224)	0,573(0,096)	1,000(0,687)
$\alpha = 3$	2,513(0,359)	2,569(0,637)	2,513(0,360)	2,811(0,828)	2,513(0,365)	2,969(1,275)
$\phi = 2$	1,483(0,771)	2,116(0,585)	1,504(0,794)	1,946(0,693)	1,537(0,799)	2,696(0,501)
$\mu = 1$	1,151(3,076)	1,676(2,956)	1,246(3,368)	1,574(3,341)	1,283(3,230)	3,316(2,854)
$\alpha = 0, 5$	0,668(0,179)	0,548(0,127)	0,664(0,178)	0,587(0,151)	0,657(0,179)	0,472(0,117)
$\phi = 4$	2,410(1,634)	2,913(1,539)	2,376(1,576)	3,318(1,745)	2,415(1,599)	4,883(1,721)
$\mu = 2$	1,457(0,781)	1,628(0,773)	1,443(0,776)	1,833(0,944)	1,457(0,822)	2,654(1,097)
$\alpha = 2$	3,105(0,990)	2,804(0,786)	3,124(0,988)	2,687(0,782)	3,090(0,978)	2,161(0,609)
$\phi = 0, 4$	0,631(0,100)	0,750(0,269)	0,628(0,097)	0,685(0,292)	0,627(0,098)	0,772(0,516)
$\mu = 1, 5$	1,679(0,104)	1,761(0,249)	1,677(0,102)	1,718(0,267)	1,674(0,100)	1,860(0,554)
$\alpha = 5$	3,728(0,459)	3,850(0,716)	3,735(0,448)	4,393(0,999)	3,753(0,463)	4,790(1,443)

Tabela 28 – Médias e desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = 200$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

$n = 200$	$\sigma_{\theta}^2 = (1, 1, 1)$		$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		$\sigma_{\theta}^2 = (10, 10, 10)$	
θ	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post	Média(Desvio P.) Val .In	Média(Desvio P.) Dis. Post
$\phi = 0, 5$	0,680(0,093)	0,694(0,258)	0,682(0,095)	0,665(0,273)	0,683(0,094)	0,708(0,405)
$\mu = 0, 5$	0,567(0,044)	0,568(0,118)	0,568(0,045)	0,560(0,120)	0,569(0,044)	0,607(0,261)
$\alpha = 3$	2,479(0,249)	2,873(0,569)	2,473(0,247)	3,063(0,746)	2,469(0,247)	3,178(0,987)
$\phi = 2$	1,759(0,741)	2,231(0,575)	1,729(0,725)	2,093(0,633)	1,772(0,741)	2,666(0,510)
$\mu = 1$	1,523(3,216)	1,965(3,087)	1,440(3,092)	1,773(3,038)	1,535(3,166)	3,322(2,830)
$\alpha = 0, 5$	0,583(0,132)	0,513(0,090)	0,589(0,134)	0,539(0,108)	0,581(0,132)	0,463(0,083)
$\phi = 4$	3,069(1,975)	3,535(1,880)	3,035(1,917)	4,056(2,077)	3,041(1,959)	5,290(2,100)
$\mu = 2$	1,765(1,257)	1,936(1,252)	1,751(1,312)	2,232(1,478)	1,770(1,432)	2,971(1,664)
$\alpha = 2$	2,671(0,802)	2,481(0,648)	2,678(0,794)	2,354(0,666)	2,660(0,772)	2,032(0,576)
$\phi = 0, 4$	0,635(0,065)	0,599(0,167)	0,637(0,067)	0,541(0,188)	0,634(0,065)	0,519(0,235)
$\mu = 1, 5$	1,680(0,069)	1,629(0,146)	1,682(0,070)	1,591(0,158)	1,678(0,070)	1,582(0,220)
$\alpha = 5$	3,663(0,309)	4,222(0,566)	3,658(0,316)	4,773(0,885)	3,664(0,306)	5,171(1,174)

Tabela 29 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$.

θ	$\sigma_{\theta}^2 = (1, 1, 1)$			$\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0, 5$	89,40%	88,70%	93,80%	92,60%	90,70%	92,80%
$\mu = 0, 5$	90,50%	91,70%	95,30%	93,10%	92,90%	94,60%
$\alpha = 3$	92,50%	91,50%	95,50%	94,00%	92,70%	94,10%
$\phi = 2$	95,00%	90,60%	87,10%	73,90%	84,00%	86,80%
$\mu = 1$	93,60%	90,90%	87,70%	39,80%	59,10%	78,00%
$\alpha = 0, 5$	93,30%	90,80%	87,20%	79,70%	82,80%	86,40%
$\phi = 4$	17,20%	34,60%	44,80%	32,40%	52,80%	68,40%
$\mu = 2$	23,80%	37,60%	45,90%	38,40%	55,00%	68,10%
$\alpha = 2$	40,50%	49,80%	54,20%	52,80%	61,80%	70,90%
$\phi = 0, 4$	73,80%	74,80%	78,10%	89,20%	92,00%	94,50%
$\mu = 1, 5$	83,30%	87,60%	91,70%	90,00%	94,20%	95,70%
$\alpha = 5$	69,90%	72,50%	83,20%	92,30%	94,60%	95,50%

Tabela 30 – Probabilidades de cobertura com o nível de confiança de 95% das estimativas das médias a posteriori encontradas para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e 20% de censura, utilizando-se a distribuição a priori Uniforme com $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$.

	$\sigma_\theta^2 = (10, 10, 10)$		
θ	$n = 50$	$n = 100$	$n = 200$
$\phi = 0,5$	81,80%	87,30%	91,40%
$\mu = 0,5$	84,30%	88,80%	91,60%
$\alpha = 3$	85,00%	88,30%	91,30%
$\phi = 2$	96,20%	92,50%	90,30%
$\mu = 1$	97,50%	94,60%	92,80%
$\alpha = 0,5$	97,20%	94,40%	91,40%
$\phi = 4$	99,50%	96,40%	86,90%
$\mu = 2$	99,30%	96,60%	87,30%
$\alpha = 2$	98,80%	96,30%	87,20%
$\phi = 0,4$	89,30%	92,80%	94,90%
$\mu = 1,5$	89,10%	94,20%	94,30%
$\alpha = 5$	90,40%	94,60%	95,30%

De modo semelhante aos resultados obtidos utilizando-se a distribuição a priori Gama, observa-se através das Tabelas 26-28 que, ao se obter os hiperparâmetros da distribuição Uniforme, utilizando-se uma abordagem semelhante a descrita no início da seção, só que considerando o método dos momentos (4.4), obtemos boas estimativas das médias a posteriori de ϕ, μ e α , por meio da abordagem Bayesiana, considerando-se observações censuradas.

Consideramos também diferentes valores para σ_θ^2 e, embora todos tenham retornado boas estimativas, o que se mostrou mais eficiente foi o que teve como variância $\sigma_\theta^2 = \tilde{\theta}$. Constata-se também que ao se utilizar $\sigma_\theta^2 = \tilde{\theta}$ as probabilidades de cobertura dos parâmetros tendem a crescer para 95%, conforme aumenta-se o tamanho de n .

No entanto, se compararmos os resultados obtidos empregando-se as duas distribuições a priori conjuntas (Gama e Uniforme), verificamos que, quando consideramos observações censuradas, ao se utilizar a distribuição a priori Gama e $\lambda = \tilde{\theta}$ e $\sigma_\theta^2 = \tilde{\theta}$, obtemos melhores estimativas das médias a posteriori de ϕ, μ e α .

7 Aplicações

Neste capítulo, serão apresentados alguns exemplos de aplicações utilizando-se dados da literatura e dados reais. Iremos considerar a distribuição Gama Generalizada para analisar tais conjuntos de dados e iremos comparar os resultados obtidos com outros modelos, se possível, como as distribuições Weibull, Gama e Log-normal, através dos testes AIC, BIC e DIC.

Será utilizada a abordagem Bayesiana para se obter as estimativas dos parâmetros das distribuições citadas acima.

Objetivando-se realizar inferências nos parâmetros da distribuição Gama Generalizada através da abordagem Bayesiana, iremos utilizar a distribuição a priori Gama (empírica), sendo os hiperparâmetros obtidos através dos métodos dos momentos (4.8), considerando-se $\lambda = \tilde{\theta}$ e $\sigma_{\theta}^2 = \tilde{\theta}$.

Para verificar o desempenho das outras distribuições, consideramos a distribuição a priori Gama em que $\theta_1 \sim \text{Gama}(0.1, 0.1)$, $\theta_1 \sim \text{Gama}(0.1, 0.1)$, para as distribuições Weibull, Gama e Log-normal.

Para cada conjunto de dados, foram realizadas 30000 iterações por meio dos métodos MCMC, considerando-se as quatro distribuições a posteriori. Buscando-se diminuir as autocorrelações das cadeias, foram salvos valores com espaços de tamanho 15, obtendo-se no final para cada parâmetro, uma cadeia de tamanho 2000. Com estes resultados calculamos as estimativas a posteriori dos parâmetros de interesse.

Para verificar se as convergências das cadeias foram alcançadas, utilizamos o critério de Geweke, por meio da estatística dada em (1.53). Obtivemos indicação de convergência nos casos em que o valor z_G esteve contido no intervalo $[-1, 96, 1, 96]$. Outros critérios também foram utilizados para se verificar a indicação de convergência, como os gráficos de séries temporais e os gráficos de autocorrelação amostral.

7.1 Dados de um Experimento Industrial

Nesta seção, será considerado um conjunto de dados de tempo de vida de 30 unidades em um experimento industrial. O conjunto de dados não possui valores censurados sendo descrito por meio da Tabela 31:

Tabela 31 – Conjunto de dados de tempo de vida de 30 unidades em um experimento industrial.

275	13	147	23	181	30	65	10	300	173
106	300	300	212	300	300	300	2	261	293
88	274	28	143	300	23	300	80	245	266

Para se obter as estimativas das médias a posteriori de ϕ, μ e α , é necessário encontrar os bons valores iniciais $\tilde{\theta} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, sendo estes obtidos através das soluções das equações (5.8, 5.9 e 5.10). Desta forma, para o conjunto de dados analisado, encontramos que $\tilde{\phi} = 0.2147$, $\tilde{\mu} = 0.002864$ e $\tilde{\alpha} = 5.116$.

Utilizando-se o método dos momentos (4.8) e considerando-se com média $\lambda = \tilde{\theta}$ e variância $\sigma_{\theta}^2 = \tilde{\theta}$, obtemos que $\phi \sim \text{Gama}(0.2147, 1)$, $\mu \sim \text{Gama}(0.002864, 1)$ e $\alpha \sim \text{Gama}(5.116, 1)$.

Os gráficos contendo amostras geradas das densidades marginais a posteriori de ϕ, μ e α , por meio dos métodos MCMC, os gráficos de séries temporais e os de autocorrelações amostrais podem ser observados através da Figura 11.

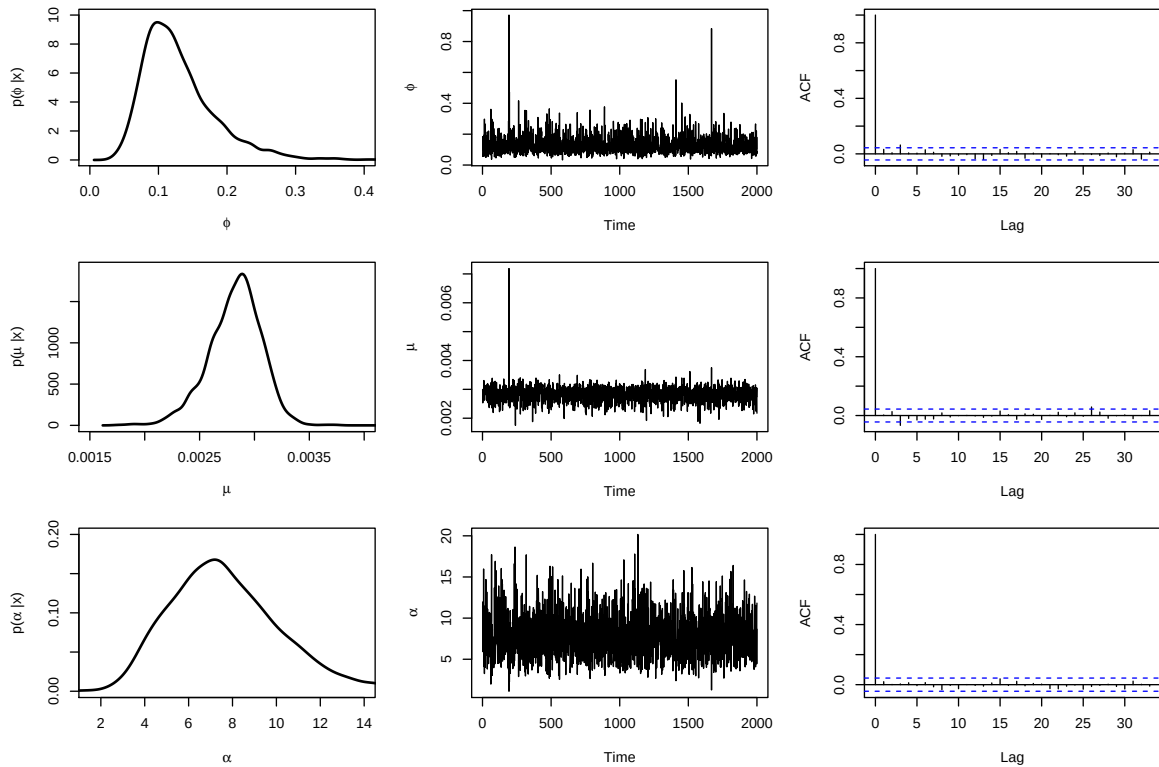


Figura 11 – Densidades marginais a posteriori, gráficos de séries de temporais e gráficos de autocorrelação para os valores gerados de ϕ, μ e α , utilizando-se os dados de tempo de vida de 30 unidades em um experimento industrial.

Observa-se pela Figura 11 e também por meio do diagnóstico de Geweke (1992), que houve indicação de convergência das cadeias para as distribuições de equilíbrio. Os

resultados das cadeias simuladas podem ser supostos como amostras das distribuições marginais a posteriori de ϕ , μ e α . As estimativas a posteriori de θ estão disponíveis na Tabela 32.

Tabela 32 – Médias a posteriori, desvios-padrão a posteriori, erros de Monte Carlo e intervalos de credibilidade de 95% dos parâmetros ϕ , μ e α

θ	Média	Desvio P.	Erro MC	$IC_{95\%}(\theta)$
ϕ	0,130500	0,059450	0,001225	(0,059550; 0,263000)
μ	0,002824	2,591E-4	5,268E-6	(0,002292; 0,003241)
α	7,734000	2,587000	0,048990	(3,66000; 13,620000)

A Tabela 33 apresenta os resultados dos Critérios DIC, BIC e AIC de diferentes distribuições de probabilidade, considerando-se os dados de um experimento industrial.

Tabela 33 – Resultados dos Critérios DIC, BIC e AIC de diferentes distribuições de probabilidade, considerando-se os dados de um experimento industrial.

Critérios	G. Generalizada	Weibull	Gama	Log-normal
DIC	358.50	363,60	374,50	386,10
BIC	364.79	385,70	377,18	389,08
AIC	360.58	382,90	374,37	386,28

Baseado em qualquer um dos critérios utilizados, podemos concluir através da Tabela 33, que a distribuição Gama Generalizada foi a que melhor se ajustou aos dados industriais.

Pode-se observar através da Figura 12 o gráfico TTT-plot, a função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada.

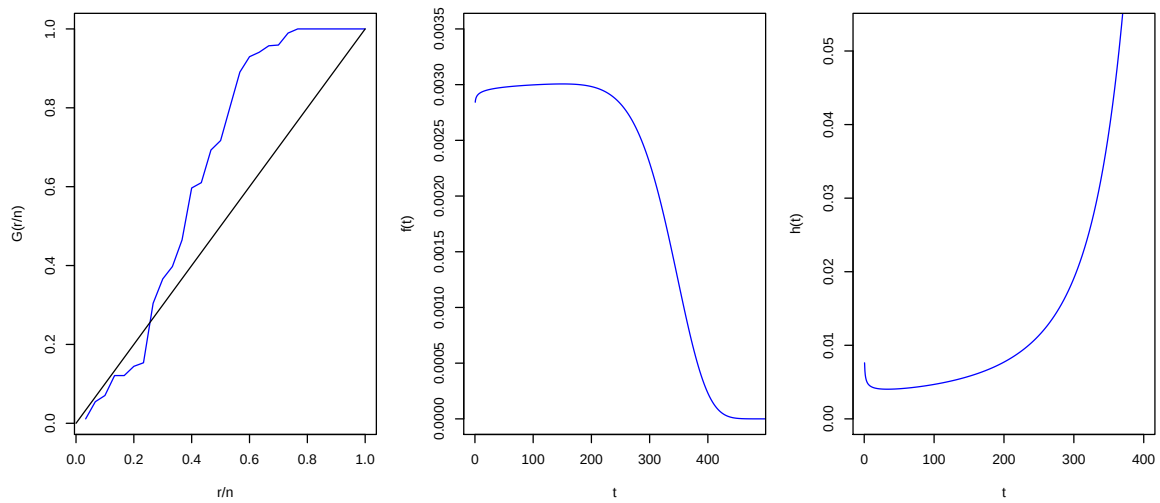


Figura 12 – TTT-plot, função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada, utilizando-se os dados de tempo de vida de 30 unidades em um experimento industrial.

Através dos resultados obtidos, concluímos que ao se utilizar as equações iniciais (5.8,5.9,5.10), encontramos bons valores iniciais que auxiliam na obtenção das estimativas a posteriori dos parâmetros da distribuição Gama Generalizada, sendo este modelo o que melhor se ajusta aos dados industriais.

7.2 Dados de Precipitação Pluvial Total Mensal

Os dados de precipitação pluviométrica utilizados nesta seção, foram obtidos da Estação Meteorológica da Faculdade de Ciências e Tecnologia da Universidade Estadual Paulista, campus de Presidente Prudente (SP), disponíveis na Tabela 33 e na página virtual do SIGRH-SP - Sistema de Informações para o Gerenciamento de Recursos Hídricos do Estado de São Paulo, da estação medidora D8-003, no município de Presidente Prudente - SP. Os valores obtidos são referentes ao mês de maio e compreendem um período de 56 anos (1947 a 2003). Desta forma, o conjunto de dados de precipitação pluvial total mensal (em mm), representando uma amostra, é composto por meio do valor total (em mm), registrado no mês de maio para cada ano.

Tabela 34 – Conjunto de dados de precipitação pluvial total mensal (em mm), referente ao mês de maio, considerando um período de 56 anos (1947 a 2003).

15,5	3,0	23,0	21,8	150,6	31,0	46,7	22,8	39,4	0,1	48,0
238,6	51,4	187,0	19,9	85,5	86,8	51,6	24,3	86,0	22,4	46,8
157,8	101,8	0,1	35,6	54,7	64,2	104,1	83,6	72,0	77,0	13,7
89,6	43,2	110,6	97,2	65,4	0,1	83,3	158,0	29,7	59,3	129,7
160,7	126,6	47,2	75,5	13,0	171,3	70,8	75,5	61,2	41,5	81,4
55,4	37,8	114,7	144,4	34,8	172,1					

Para se obter as estimativas das médias a posteriori de ϕ , μ e α , é necessário se encontrar os bons valores iniciais $\tilde{\theta} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, sendo estes obtidos através das soluções das equações (5.8, 5.9 e 5.10). Desta forma, para o conjunto de dados de precipitação pluvial total mensal, encontramos que $\tilde{\phi} = 0,5348$, $\tilde{\mu} = 0,0081$ e $\tilde{\alpha} = 1,7866$.

Utilizando-se o método dos momentos (4.8) e considerando-se como média $\lambda = \tilde{\theta}$ e variância $\sigma_{\theta}^2 = \tilde{\theta}$ obtemos que $\phi \sim \text{Gama}(0,5348, 1)$, $\mu \sim \text{Gama}(0,0081, 1)$ e $\alpha \sim \text{Gama}(1,7866, 1)$.

Os gráficos contendo amostras geradas das densidades marginais a posteriori de ϕ , μ e α , por meio dos métodos MCMC, os gráficos de séries temporais e os de autocorrelações amostrais, podem ser observados através da Figura 13.

Observa-se pela Figura 13 e também por meio do diagnostico de Geweke (1992), que houve indicação de convergência das cadeias para as distribuições de equilíbrio. Os resultados das cadeias simuladas podem ser supostos como amostras das distribuições marginais a posteriori de ϕ , μ e α . As estimativas a posteriori de θ estão disponíveis na Tabela 35.

Tabela 35 – Médias a posteriori, Desvios-padrão a posteriori, Erros de Monte Carlo e Intervalos de credibilidade de 95% dos parâmetros ϕ , μ e α

θ	Média	Desvio P.	Erro MC	$IC_{95\%}(\theta)$
ϕ	0,36080	0,15140	0,003020	(0,15170; 0,74520)
μ	0,00685	0,00144	2,846E-5	(0,00502; 0,01033)
α	2,58100	0,80520	0,015600	(1,36000; 4,58400)

A Tabela 36 apresenta os resultados dos Critérios DIC, BIC e AIC de diferentes distribuições de probabilidade, considerando-se os dados de precipitação pluvial total mensal.

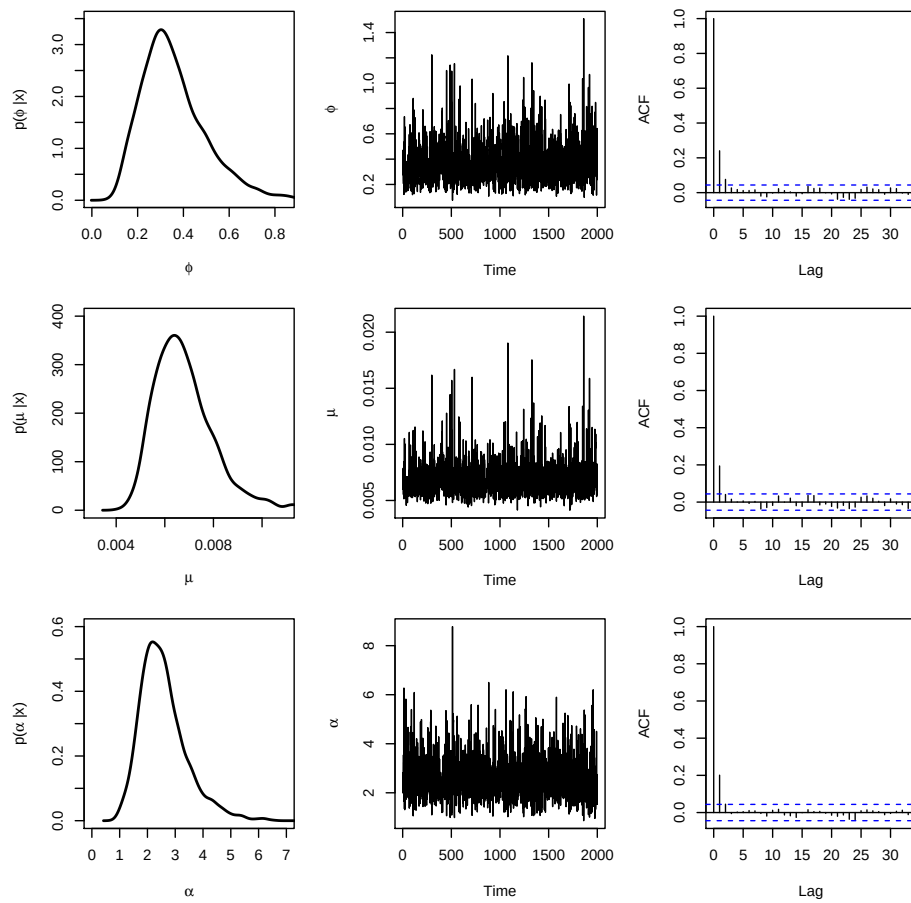


Figura 13 – Densidades marginais a posteriori, gráficos de séries de temporais e gráficos de autocorrelação para os valores gerados de ϕ , μ e α utilizando-se os dados de precipitação pluvial total mensal.

Tabela 36 – Resultados dos Critérios DIC, BIC e AIC de diferentes distribuições de probabilidade, considerando-se os dados de precipitação pluvial total mensal.

Critérios	G. Generalizada	Weibull	Gama	Log-normal
DIC	639,80	644,70	648,10	691,70
BIC	648,04	652,29	652,40	695,91
AIC	641,71	648,07	648,18	691,69

Baseado em qualquer um dos critérios utilizados, podemos concluir, através da Tabela 35, que a distribuição Gama Generalizada foi a que melhor se ajustou aos dados de precipitação pluvial total mensal.

Pode-se observar através da Figura 14, o gráfico TTT-plot, a função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada.

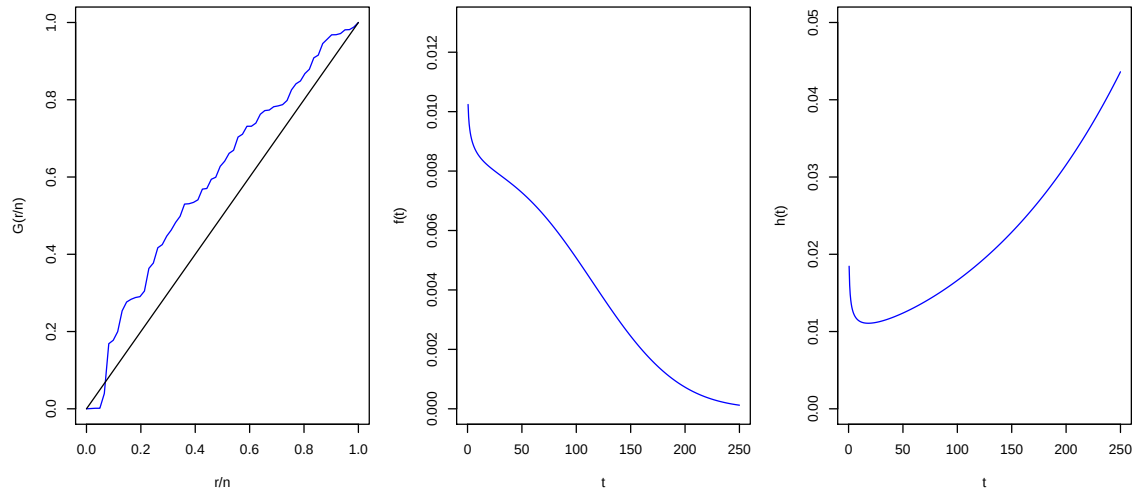


Figura 14 – TTT-plot, função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada, utilizando-se os dados de precipitação pluvial total mensal.

Por meio dos resultados obtidos concluímos que, ao se utilizar as equações iniciais (5.8,5.9,5.10) encontramos bons valores iniciais que auxiliam na obtenção das estimativas a posteriori dos parâmetros da distribuição Gama Generalizada, sendo este um modelo consistente para se ajustar a dados climatológicos.

7.3 Dados de Sobrevida de Pacientes com Câncer de Pulmão

Em estudos médicos, é comum se deparar com dados que possuem censura aleatória. Nesta seção iremos analisar os dados disponíveis em Gupta, et al. (1999). Estes dados foram obtidos da Veterans Administration e consistem em 97 pacientes com câncer de pulmão, sendo que os mesmos não tiveram tratamento prévio até descobrir a doença. Após efetuar o tratamento, os dados representam o tempo de sobrevida (em dias) destes pacientes com câncer. Na Tabela 36, temos os tempos de sobrevida (em dias) de 97 pacientes com câncer de pulmão (+ indica a presença de censura).

Tabela 37 – Os tempos de sobrevida (em dias) de 97 pacientes com câncer de pulmão (+ indica a presença de censura).

72	228	10	110	314	+100	42	144	30	384
4	13	+123	+97	59	117	151	22	18	139
20	31	52	18	51	122	27	54	7	63
392	92	35	117	132	162	3	95	162	216
553	278	260	156	+182	143	105	103	112	+87
242	111	587	389	33	25	357	467	1	30
283	25	21	13	87	7	24	99	8	99
61	25	95	80	29	24	+83	31	51	52
73	8	36	48	7	140	186	19	45	80
52	53	15	133	111	378	49			

Para se obter as estimativas das médias a posteriori de ϕ, μ e α , é necessário encontrar-se os bons valores iniciais $\tilde{\theta} = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, sendo estes obtidos através das soluções das equações (5.8, 5.9 e 5.10). Como os dados possuem observações censuradas, iremos utilizar o método 3, discutido na seção 6.1.2 para se obter $\tilde{\theta}$. Desta forma, para os tempos de sobrevida de pacientes com câncer de pulmão, encontramos que $\tilde{\phi} = 1,62556$, $\tilde{\mu} = 0,02066$ e $\tilde{\alpha} = 0,70045$.

Utilizando-se o método dos momentos (6.1.2) e considerando-se com média $\lambda = \tilde{\theta}$ e variância $\sigma_{\theta}^2 = \tilde{\theta}$, obtemos que $\phi \sim \text{Gama}(1.62556, 1)$, $\mu \sim \text{Gama}(0.02066, 1)$ e $\alpha \sim \text{Gama}(0.70045, 1)$.

Os gráficos contendo amostras geradas das densidades marginais a posteriori de ϕ, μ e α , por meio dos métodos MCMC, os gráficos de séries temporais e os de autocorrelações amostrais, podem ser observados através da Figura 15.

Observa-se pela Figura 15 e também por meio do diagnóstico de Geweke (1992), que houve indicação de convergência das cadeias para as distribuições de equilíbrio. Os resultados das cadeias simuladas podem ser supostos como amostras das distribuições marginais a posteriori de ϕ, μ e α .

Devido a assimetria obtida na distribuição marginal a posteriori de μ , serão calculadas as estimativas das medianas a posteriori de ϕ, μ e α . As estimativas a posteriori de θ estão disponíveis na Tabela 38.

Tabela 38 – Medianas a posteriori, desvios-padrão a posteriori, erros de Monte Carlo e intervalos de credibilidade de 95% dos parâmetros ϕ, μ e α

θ	Mediana	Desvio P.	Erro MC	$IC_{95\%}(\theta)$
ϕ	2.0080	0.9923	0.046720	(0.71270; 4.56900)
μ	0.0307	0.1741	0.007203	(0.00579; 0.50480)
α	0.6198	0.1994	0.008612	(0.39380; 1.16000)

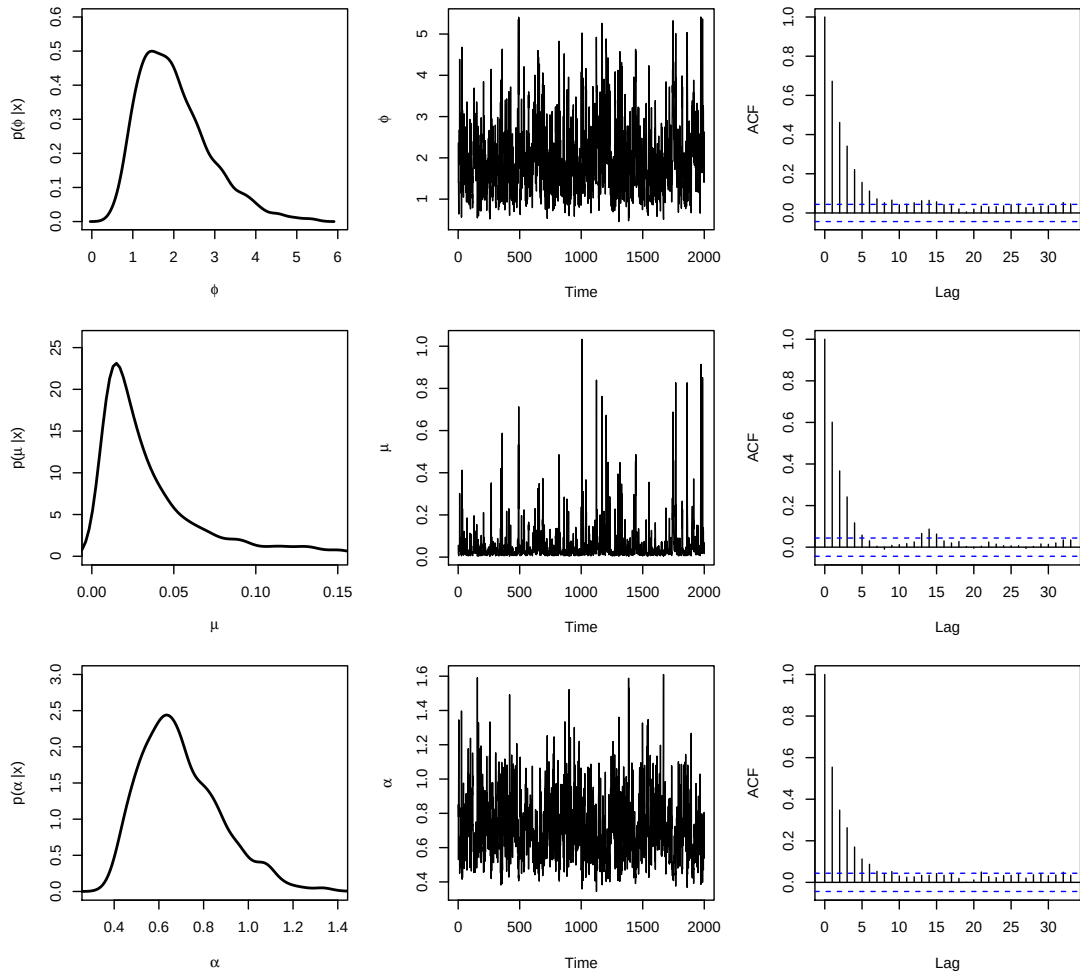


Figura 15 – Densidades marginais a posteriori, gráficos de séries de temporais e gráficos de autocorrelação para os valores gerados de ϕ , μ e α , utilizando-se os dados de tempo de vida de 30 unidades em um experimento industrial.

Pode-se observar, através da Figura 16 a seguir, o gráfico TTT-plot, a função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada.

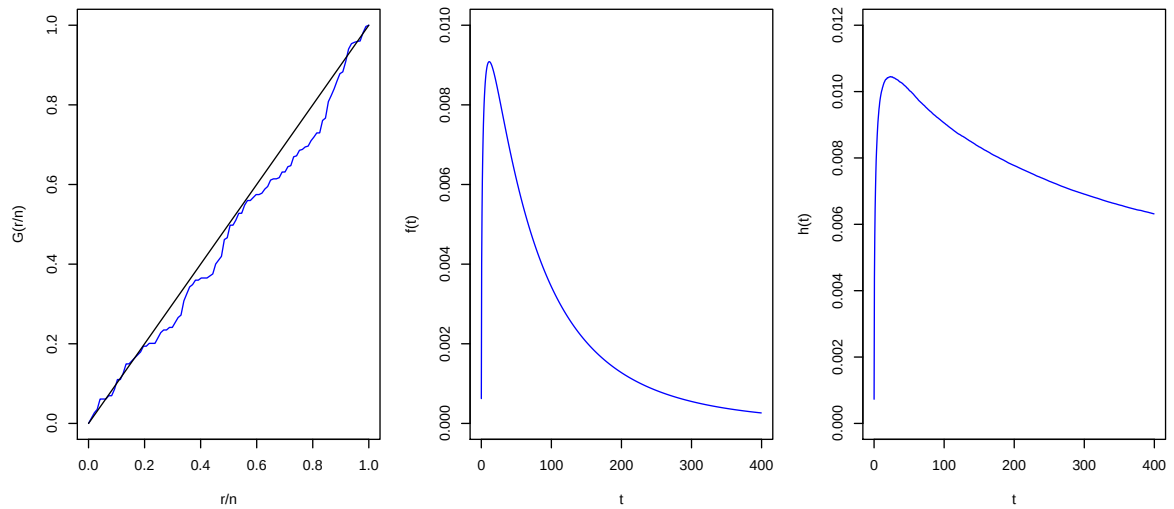


Figura 16 – TTT-plot, função densidade de probabilidade e a função de risco, ambas ajustadas através da distribuição Gama Generalizada, utilizando-se os tempos de sobrevida de 97 pacientes com câncer de pulmão.

Através dos resultados obtidos pelo TTT-plot pode-se observar que o conjunto de dados possui forma de banheira inversa. Como os modelos Weibull, Gama e Log-normal não possuem este tipo de função de risco, não devemos considerá-los ao se realizar os testes de seleção de modelos. Ajustada a distribuição Gama Generalizada aos tempos de sobrevida de pacientes com câncer de pulmão, pode-se observar facilmente a forma da função de risco, mostrando-se que este é um ótimo modelo para descrever dados médicos.

Considerações Finais e Propostas Futuras

Considerações Finais

Por meio da pesquisa realizada podemos constatar que a distribuição Gama Generalizada é um modelo muito flexível para modelar dados de confiabilidade, de climatologia e dados médicos, pois acomoda várias formas da função de risco, sendo estas: constante, crescente, decrescente, forma de banheira e banheira inversa.

Demonstramos que ao se utilizar a inferência Clássica e considerando-se observações completas é possível reduzir-se as equações de verossimilhança, eliminando-se assim a instabilidade dos estimadores no método iterativo, possibilitando assim facilmente a obtenção de boas estimativas para os parâmetros do modelo.

Considerando-se a abordagem Bayesiana, propomos diferentes distribuições a priori, empíricas ou não-informativas, sendo que seus efeitos foram estudados nas estimativas a posteriori dos parâmetros de interesse.

Dentre as prioris não-informativas obtidas, demonstramos que as prioris: de Jeffreys, MDIP e a obtida pela Regra de Jeffreys, levam a distribuições a posteriori impróprias e não devem ser utilizadas em uma análise Bayesiana. Uma modificação na priori obtida através da Regra de Jeffreys foi realizada e constatamos que, utilizando tal distribuição a priori, obtemos uma densidade a posteriori própria.

Propomos também, técnicas exploratórias, visando-se obter bons valores iniciais que, combinados com o método dos momentos, possibilitaram a construção de diferentes distribuições a priori empíricas.

Neste trabalho, também foram propostos diferentes métodos exploratórios, a fim de se obter bons valores iniciais para serem utilizados nos procedimentos de estimação dos parâmetros, considerando-se dados censurados (censura aleatória), sob o enfoque Clássico e Bayesiano.

Através de análises numéricas, verificamos que, ao se utilizar a distribuição a priori Gama (empírica), obtemos excelentes estimativas das médias a posteriori de ϕ , μ e α , principalmente quando utilizamos como média $\lambda = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e como variância $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$, tanto para observações completas, quanto para observações censuradas. Estas estimativas obtidas, se mostraram superiores às alcançadas por meio da inferência Clássica.

Estes resultados mostram-se de grande interesse prático, pois possibilitarão, a fácil utilização da distribuição Gama Generalizada em diversas áreas de aplicação.

Propostas Futuras

Uma característica muito importante da análise de sobrevivência é a presença de uma ou mais covariáveis, onde vários fatores de risco podem prolongar ou diminuir a sobrevida dos pacientes ou componentes industriais.

Outro fator importante que poderá ser explorado é a presença de fração de cura com a distribuição Gama Generalizada, onde parte dos pacientes são curados ou imunes e nunca terão a ocorrência do evento de interesse (ver por exemplo, Maller e Zhou, 1996; ou Lambert et al., 2010). Neste caso, assumimos que a população é dividida em dois grupos de indivíduos: um grupo de indivíduos curados com probabilidade π e um grupo de indivíduos suscetíveis com uma função de sobrevivência própria dada por $S_0(\mathbf{t}) = P(T > t)$ onde T denota o tempo de vida do indivíduo com probabilidade $1 - \pi$. Nesse modelo de mistura, temos:

$$S(\mathbf{t}) = \pi + (1 - \pi)S_0(\mathbf{t}) \quad (7.1)$$

onde $\pi \in (0, 1)$ e $S_0(\mathbf{t})$ é a função de sobrevivência para os indivíduos suscetíveis (observações não censuradas).

Referências

- AALO, V. A.; PIBOONGUNGON, T.; ISKANDER, C.D. Bit-Error Rate of Binary Digital Modulation Schemes in Generalized Gamma Fading Channels. *IEEE Comm. Letters*, v. 9, p. 139-141, 2005.
- AHRENS, J. H.; DIETER, U. Computer methods for sampling from gamma, beta, Poisson and binomial distributions. *Computing*, v. 12, p. 223-246, 1974.
- AHRENS, J. H.; DIETER, U. Generating gamma variates by a modified rejection technique. *Communications of the ACM*, vol. 25, p. 47-54, 1982.
- AHSANULLAH, M.; MASWADAH, M.; SEHAM A. M. Kernel Inference on the Generalized Gamma Distribution Based on Generalized Order Statistics. *Journal of Statistical Theory and Applications*, v. 12, n. 2, p. 152-172, 2013.
- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v. 19, n. 6, p. 716-723, 1974.
- BALAKRISHNAN, N.; SANDHU, R. A. A simple simulation algorithm for generating progressive Type-II censored samples. *The American Statistician*, v. 49, p. 229-230, 1995
- BALAKRISHNAN, N.; AGGARWALA, R. *Progressive Censoring: Theory, Methods and Applications*, Birkhauser, Boston, 2000.
- BALAKRISHNAN, N.; PENG, Y. Generalized gamma frailty model, *Statist. Med.* v. 25, p. 2797-2816, 2006.
- BARLOW, R. E.; CAMPO, R. A. Total Time on Test processes and applications to failure data analysis. In *Reliability and fault tree analysis*. SIAM, Pennsylvania, 1975.
- BAYES, T.R. (1763) *Essay towards solving a problem in the doctrine of changes*. Reprinted in *Biometrika*, v. 45, p. 243-315, 1958.
- BENNETT, J. E.; RACINE-POON, A.; WAKEFIELD, J. C. MCMC for nonlinear hierarchical models. In *Markov Chain Monte Carlo in Practice* (eds W. R. Gilks, S. Richardson and D. J. Spiegelhalter), Chapman and Hall, London, p. 339-357, 1996.
- BERGER, J. O.; BERNARDO, J. M. On the Development of the Reference Prior Method. *Fourth Valencia International Meeting on Bayesian Statistics*, Spain, 1992.

- BERNARDO, J. M. Reference Posterior Distributions for Bayesian Inference. *Journal Royal Statistical Society*, v.41, n. 2, p. 113-147, 1979.
- BEST, N.; COWLES, M.; VINES, K. CODA: Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output. Cambridge, UK, 1997.
- BOX, G. E. P.; TIAO, G. C. Bayesian Inference in Statistical Analysis. Reading: Addison-Wesley, p. 588, 1973.
- BROOKS, S.; GELMAN, A.; JONES, G. L; MENG, X. L. Handbook on Markov Chain Monte Carlo: Methods and Applications. Chapman & Hall/CRC Press. 2011.
- CASELLA, G.; BERGER, R. Statistical Inference (2nd ed.), Belmont, CA: Duxbury, 2002.
- CHANG, H.; KIM, B. H. Non-informative priors in the generalized gamma stress-strength systems. *IIE Transactions*, v.43, p. 797-804, 2011.
- CHEN, D. G.; LIO, Y. A Note on the Maximum Likelihood Estimation for the Generalized Gamma Distribution Parameters under Progressive Type-II Censoring. *International Journal of Intelligent Technology and Applied Statistics*, v. 2, n. 2, p. 57-64, 2009.
- CHIB, S.; GREENBERG, E. Understanding The Metropolis-Hasting Algorithm. *The American Statistician*, v. 49, n. 4, p. 327-335, 1995.
- COLOSIMO, E. A.; GIOLO, S. R. Análise de sobrevivência aplicada. São Paulo: Edgar Blucher, 2006, 392p.
- DADPAY, A.; SOOFI, E. S; SOYER, R. Information measures for generalized gamma family. *Journal of Econometrics*, v. 138 , p. 568-585, 2007.
- DEGROOT, M. H.; SCHERVISH, M. J. Probability and Statistics. Addison-Wesley, Reading, Massachusetts, 2002.
- DI CICCIO, T. Approximate inference for the generalized gamma distribution. *Technometrics*, v. 29, n. 1, p. 33-40, 1987.
- FOLLAND, G. B. Real Analysis - Modern Techniques and Their Applications, John Wiley & Sons, 1999.
- GAMERMAN, D.; LOPES, H. F. Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference. 2nd edition. Chapman&Hall/CRC, 2006.
- GELFAND, A. E.; SMITH, F. M. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, n. 85, p. 398-409, 1990.

- GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, v.7, p. 457-472, 1992.
- GEMAN, S.; GEMAN, D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 6, p. 721-741, 1984.
- GENTLE, J.E. *Random Number Generation and Monte Carlo Methods*, 2nd ed., Springer, 2003.
- GEWEKE, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments (with discussion). (eds. Bernardo, J.M., et al.) *Bayesian Statistics 4*, Clarendon Press, Oxford, UK, p.169-193, 1992.
- GIA, T. P.; ALMHANA, J. The Generalized Gamma Distribution: Its Hazard Rate and Stress-Strength Model. *IEEE Transactions on reliability*, v. 44, no. 3, Sept, 1995.
- GLASER, R.E. Bathtub and related failure rate characterization. *J. Amer. Statistical Assoc.*, v. 75, p. 667-672, 1980.
- GUPTA, R. C.; AKMAN, H. O.; LVIN, S. A study of log-logistic model in survival analysis, *Biometrical Journal*, v. 41(4), p. 431-443, 1999.
- GUPTA, R. D.; KUNDU, D. Generalized exponential distributions. *Aust. N.Z. J. Stat.*, v. 41, p. 173-188, 1999.
- GUPTA, R. D.; KUNDU, D. Generalized exponential distributions: different methods of estimation. *J. Stat. Comp. Simul.*, v. 69, p. 315-338, 2001.
- HAGER, H. W.; BAIN L. J. Inferential procedures for the generalized gamma distribution. *Journal of the American Statistical Association*, v. 65, p. 1601-1609, 1970.
- HAMADA, M.; WILSON, A. G.; REESE, C. S.; and MARTZ, H. F. *Bayesian Reliability*, Springer, New York, 2008.
- HARTER, H. L. Maximum-likelihood estimation of the parameters of a four-parameter generalized gamma population from complete and censored samples. *Technometrics*, v. 9, p. 159-165, 1967.
- HASTINGS, W. K. Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, v. 57, p. 97-109, 1970.
- HEIDELBERGER, W. Simulation run length control in the presence of an initial transient. *Operations Research*, v.31, p. 1109-1144, 1983.

- HUANG, P; HWANG, T. On new moment estimation of parameters of the Generalized Gamma distribution using it's characterization. *Taiwanese journal of Mathematics*, v. 10, n. 4, p. 1083-1093, 2006.
- JAYNES, E. T. Discussion, in Bernardo, J.M.; DeGroot, M.H.; Lindley, D.V.; Smith, A.F.M.; *Bayesian Statistics*, Valencia, Spain: University Press, p. 618-229, 1980.
- JEFFEYS, H. *Theory of probability*. 3rd rev. ed. London: Oxford U. Press, 1967.
- KASS, R. E.; WASSERMAN, L. The selection of prior distributions by formal rules. *Journal of the American Statistical Association*. v. 91, n. 435, p. 1343-1370, 1996.
- KHODABIN M; AHMADABADI A. Some properties of generalized gamma distribution. *J. Mathematical Sciences*. v. 4, p. 9-28, 2010.
- KUMAR, V.; SHUKLA, G. Maximum Likelihood Estimation in Generalized Gamma Type Model. *Journal of Reliability and Statistical Studies*, v. 3, n. 1, p. 43-51, 2010.
- LAMBERT, P. C.; DICKMAN, P. W.; WESTON, C. L.; THOMPSON, J. R. Estimating the cure fraction in population-based cancer studies by using finite mixture models. *Journal of the Royal Statistical Society. Series C. Applied Statistics*, v. 59, p. 35-55, 2010.
- LAPLACE, P. S. (1774) *Mémoire sur la Probabilité des Causes par les Évènements*, translated by S. M. Stigler as "Memoir on the Causes of Events", *Statistical Science*, v. 20, p. 32-49, 2005.
- LAWLESS, J. F. *Statistical models and methods for lifetime data*. New York: John Wiley and Sons, 580 p, 1982.
- LI, H. C.; HONG, W.; WU, Y. R.; FAN, P. Z. On the empirical statistical modeling of SAR images with generalized Gamma distribution. *IEEE J. Sel. Topics Signal Process.*, v. 5, n. 3, p. 386-397, Jun, 2011.
- LIMA, E. L. - *Análise Real*, vol. 1, Rio de Janeiro, IMPA. *Coleção Matemática Universitária*, 11.ed., 189p, 2011.
- MALLER, R. A.; ZHOU, X. *Survival analysis with long-term survivors*. John Wiley & Sons Ltd., Chichester, United Kingdom, 1996.
- MARANI, A.; LAVAGNINI, I; BUTTAZZONI, C. Statistical study of air pollutant concentrations via generalized gamma distributions. *JAPCA*, v. 36, p. 1250-1254, 1986.
- MEEKER, W.Q.; ESCOBAR, L. A. *Statistical Methods for Reliability Data*. John Wiley & Sons, 1998.

- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N., Teller, A. H. & Teller, E. Equations of state calculations by fast computing machine. *Journal of Chemical Physics*, v. 21, p. 1087-1091, 1953.
- MUKHERJEE, B.; GUPTA, A.; UPADHYAY, S. K. A Bayesian study for the comparison of generalized gamma model with its components. *Sankhya*, v.72, p. 154-174, 2010.
- NAIR, N.U., SANKARAN, P.G., KUMAR, B.V. Total time on test transforms of orderand its implications in reliability analysis. *Journal of Applied Probability*, v. 45, p. 1126-1139, 2008.
- NELSON, W. *Applied Life Data Analysis*. John Wiley & Sons, New York, 662 p,1982.
- NOORTWIJK, J. M. Bayes Estimates of Flood Quantiles using the Generalised Gamma Distribution. *System and Bayesian Reliability; Essays in Honor of Professor Richard E. Barlow on His 70th Birthday*, Singapore: World Scientific Publishing, p.351-374, 2001.
- PARR, V. B.; WEBSTER, J. T. A method for discriminating between failure density functions used in reliability predictions. *Technometrics*, v. 7, p. 1-10, 1965.
- PHILIPPE, A. Simulation of right and left truncated gamma distributions by mixtures. *Statistics and Computing* , v. 7, p. 173-181, 1997.
- RAFTERY, A. L.; LEWIS, S. How many iterations in the Gibbs sampler? *Bayesian Statistics*, v.4, p. 763-774, 1992.
- RAJU, B. I.; SRINIVASAN, M. A. Statistics of envelope of high-frequency ultrasonic backscatter from human skin in vivo. *IEEE Trans. Ultrason., Ferroelect. Freq. Contr.*, v. 49, n. 7, p. 871-882, 2002.
- ROSS, S.M. *Simulation*, 4th ed, United States of America Elsevier Academic Press, 2006.
- SANTOS, C. A. Dados de sobrevivência multivariados na presença de covariáveis e observações censuradas: uma abordagem bayesiana. 2010. 156 f. Tese (Doutorado em Estatística) - Centro de Ciências Exatas e de Tecnologia - Universidade Federal de São Carlos, São Carlos, 2010.
- SCHWARZ, G. Estimating the dimension of a model. *The Annals of Statistics*, v. 6, p. 461-464, 1978.
- SHANNON, C. E. A Mathematical theory of communication. *Bell System Technical Journal*, v.27, p 623-659, 1948.
- SHAWKY, A.I.; BAKOBAN, R.A. Conditional expectation of certain distributions of record values.*Int. J. Math. Anal.*, v. 3, no. 17, p. 829-838, 2009.

- SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society*, v. 64 , p. 583-639, 2002.
- STACY, E. W. A generalization of the gamma distribution. *Annals of Mathematical Statistics*, v. 28, p. 1187-1192, 1962.
- STACY, E. W.; MIHRAM, G. A. Parameter estimation for a generalized gamma distribution. *Technometrics*, v.7, p. 349-358, 1965.
- TAHAI, A.; MEYER, M. A revealed preference study of management journals direct influences. *Strategic Management Journal*, v. 20, p. 279-296, 1999.
- TIERNEY, L. Markov Chains for Exploring Posterior Distributions. *The Annals of Statistics*, v. 22, n. 4, p. 1701-1728, 1994.
- WINGO, D. R. Computing maximum-likelihood parameter estimates of the generalized gamma distribution by numerical root isolation. *IEEE Trans. Rel.*, v. 36, p. 586-590, 1987.
- ZELLNER, A. Maximal Data Information Prior Distributions, In *New Methods in the applications of Bayesian Methods*. A. Aykac e C. Brumat (eds). North- Holland, Amsterdam, 1977.
- ZELLNER, A. Maximal Data Information Prior Distributions. *Basic Issues in Econometrics*, U. of Chicago Press, 1984.

Apêndices

APÊNDICE A – Demonstração de que as densidades a posteriori são impróprias

Lema 1: Seja $h(x)$ uma função contínua e maior que 0 em $(a, b]$ tal que

$$\lim_{x \rightarrow a^+} h(x) > 0.$$

Então existe $L > 0$ tal que $h(x) > L, \forall x \in (a, b]$.

Lema 2: Seja $h(x)$ uma função contínua e maior que 0 em $[b, \infty)$ tal que

$$\lim_{x \rightarrow \infty} h(x) > 0.$$

Então existe $L > 0$ tal que $h(x) > L, \forall x \in [b, \infty)$.

Parte I

Queremos mostrar que

$$d_1 = \int_{\mathcal{A}} \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} = \infty$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$.

Como $\frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} \geq 0$ sempre, pelo teorema de Tonelli (ver Folland, 1999) temos que

$$\begin{aligned} d_1 &= \int_{\mathcal{A}} \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu d\phi d\alpha. \end{aligned}$$

Observe que

$$\begin{aligned} &\int_0^\infty \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu = \\ &\frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \int_0^\infty \mu^{n\alpha\phi-1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu. \end{aligned} \tag{A.1}$$

Utilizando-se o resultado (5.4) temos que

$$\frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \int_0^\infty \mu^{n\alpha\phi-1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu = \alpha^{n-2} \frac{\Gamma(n\phi)}{\phi \Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}}. \quad (\text{A.2})$$

Agora queremos integrar (A.2) em relação à ϕ , mas, como é difícil integrar expressões contendo $\Gamma(\phi)$, queremos obter uma expressão que possa substituí-la de forma que possamos avaliar a integral. Faremos isto a seguir.

Pela fórmula de reflexão de Euler, temos que

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}, \quad (\text{A.3})$$

de onde

$$\Gamma(z) = \frac{\pi}{\sin(\pi z)} \frac{1}{\Gamma(1-z)}.$$

Assim, temos que

$$\lim_{z \rightarrow 0^+} z\Gamma(z) = \lim_{z \rightarrow 0^+} \frac{\pi z}{\sin(\pi z)} \frac{1}{\Gamma(1-z)} = 1 \times \lim_{z \rightarrow 0^+} \frac{1}{\Gamma(1-z)} = 1.$$

Deste modo, temos que

$$\lim_{\phi \rightarrow 0^+} \frac{(n\phi\Gamma(n\phi))}{\phi^n \Gamma(\phi)^n} = \lim_{\phi \rightarrow 0^+} \frac{(n\phi\Gamma(n\phi))}{(\phi\Gamma(\phi))^n} = \frac{1}{1^n} = 1. \quad (\text{A.4})$$

Por outro lado, sabemos que $\frac{(n\phi\Gamma(n\phi))}{\phi^n \Gamma(\phi)^n}$ é contínua e positiva em $(0, 1]$, então pelo Lema 1, temos que $\frac{(n\phi\Gamma(n\phi))}{\phi^n \Gamma(\phi)^n} > L_1, \forall \phi \in (0, 1]$, onde L_1 é uma constante maior que 0. Portanto

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n} > L_1 \frac{\phi^n}{n\phi} = \frac{L_1}{n} \phi^{n-1}, \forall \phi \in (0, 1]. \quad (\text{A.5})$$

Assim,

$$\begin{aligned} \int_0^\infty \alpha^{n-2} \frac{\Gamma(n\phi)}{\phi \Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi &\geq \int_0^1 \alpha^{n-2} \frac{L_1}{n} \phi^{n-2} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \\ &= \frac{L_1}{n \prod_{i=1}^n t_i} \alpha^{n-2} \int_0^1 \phi^{n-2} \frac{(\prod_{i=1}^n t_i^\alpha)^\phi}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \\ &= c_1 \alpha^{n-2} \int_0^1 \phi^{n-2} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi, \end{aligned}$$

onde $c_1 = \frac{L_1}{n \prod_{i=1}^n t_i}$ é uma constante.

Agora, segundo a desigualdades das médias, como $n > 1$, temos que

$$\sum_{i=1}^n t_i^\alpha \geq n \sqrt[n]{\prod_{i=1}^n t_i^\alpha} > \sqrt[n]{\prod_{i=1}^n t_i^\alpha} \Leftrightarrow \frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} > 1 \Leftrightarrow \log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) > 0. \quad (\text{A.6})$$

Logo, tomando $q(\alpha) = \log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) > 0$, portanto

$$\begin{aligned} \frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} &= \exp \left\{ \log \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right) \right\} = \exp \left\{ -\log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) \right\} \\ &= \exp(-q(\alpha)). \end{aligned} \quad (\text{A.7})$$

Assim,

$$c_1 \alpha^{n-2} \int_0^1 \phi^{n-2} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi = c_1 \alpha^{n-2} \int_0^1 \phi^{n-2} \exp(-n q(\alpha) \phi) d\phi.$$

Utilizando a substituição de variáveis $u = n q(\alpha) \phi$, na integral acima, temos que:

$$\begin{aligned} c_1 \alpha^{n-2} \int_0^1 \phi^{n-2} \exp(-n q(\alpha) \phi) d\phi &= c_1 \alpha^{n-2} \frac{1}{(n q(\alpha))^{n-1}} \int_0^{n q(\alpha)} u^{n-2} \exp(-u) du \\ &= c_1 \alpha^{n-2} \frac{\gamma(n-1, n q(\alpha))}{(n q(\alpha))^{n-1}}. \end{aligned} \quad (\text{A.8})$$

Agora queremos integrar (A.8) em relação à α , mas, novamente, vemos que seria difícil integrar a expressão acima diretamente. Por isso, novamente provaremos alguns resultados que facilitarão a avaliação desta integral.

Sendo t_m o maior entre os valores $t_1, t_2, \dots, t_n$, então temos que:

$$\begin{aligned} \lim_{\alpha \rightarrow \infty} \frac{q(\alpha)}{\alpha} &= \lim_{\alpha \rightarrow \infty} \frac{1}{\alpha} \log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) = \lim_{\alpha \rightarrow \infty} \log \left\{ \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right)^{\frac{1}{\alpha}} \right\} \\ &= \lim_{\alpha \rightarrow \infty} \log \left(\frac{t_m (\sum_{i=1}^n (t_i/t_m)^\alpha)^{\frac{1}{\alpha}}}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right). \end{aligned}$$

Veja que $(t_i/t_m) \leq 1$, de onde $\lim_{\alpha \rightarrow \infty} (t_i/t_m)^\alpha$ existe sendo menor ou igual 1, portanto

$$\lim_{\alpha \rightarrow \infty} \sum_{i=1}^n (t_i/t_m)^\alpha = r \leq n,$$

onde também $r \geq 1$ pois $(t_m/t_m)^\alpha = 1$. Deste modo

$$\lim_{\alpha \rightarrow \infty} \log \left(\frac{t_m (\sum_{i=1}^n (t_i/t_m)^\alpha)^{\frac{1}{\alpha}}}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) = \log \left(\frac{t_m r^0}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) = v < \infty, \quad (\text{A.9})$$

ou seja

$$\lim_{\alpha \rightarrow \infty} \frac{q(\alpha)}{\alpha} < \infty \Leftrightarrow \lim_{\alpha \rightarrow \infty} \frac{\alpha}{q(\alpha)} > 0.$$

Como $\frac{\alpha}{q(\alpha)}$ é contínua e maior que 0 em $[1, \infty)$, pelo Lema 2, temos que

$$\frac{\alpha}{q(\alpha)} > L_2 \Leftrightarrow \frac{1}{q(\alpha)} > \frac{L_2}{\alpha}, \forall \alpha \in [1, \infty). \quad (\text{A.10})$$

sendo L_2 constante maior que 0.

Por outro lado, sabemos que nem todos t_i são iguais, então em especial $t_m > t_j$ para algum j , de onde então

$$t_m = \sqrt[n]{t_m^n} > \sqrt[n]{\prod_{i=1}^n t_i} \Leftrightarrow \frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}} > 1 \Leftrightarrow \log \left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}} \right) = v > 0.$$

Portanto temos que

$$\lim_{\alpha \rightarrow \infty} \frac{q(\alpha)}{\alpha} = v > 0. \quad (\text{A.11})$$

Em especial, segue disso que

$$\lim_{\alpha \rightarrow \infty} q(\alpha) = \lim_{\alpha \rightarrow \infty} \alpha \frac{q(\alpha)}{\alpha} = \infty \times v = \infty. \quad (\text{A.12})$$

É fácil mostrar diretamente da definição de $\gamma(s, x)$ que:

$$\lim_{x \rightarrow \infty} \gamma(s, x) = \Gamma(s). \quad (\text{A.13})$$

Utilizando-se (A.13), temos que $\lim_{\alpha \rightarrow \infty} n q(\alpha) = \infty$, desta forma

$$\lim_{\alpha \rightarrow \infty} \gamma(n-1, n q(\alpha)) = \Gamma(n-1). \quad (\text{A.14})$$

Como $\gamma(n-1, n q(\alpha))$ é contínua e maior que 0 em $[1, \infty)$, novamente pelo Lema 2 segue que

$$\gamma(n-1, n q(\alpha)) > L_3, \forall \alpha \in [1, \infty) \quad (\text{A.15})$$

sendo L_3 uma constante maior que 0.

Por fim temos que

$$\begin{aligned} \int_0^\infty c_1 \alpha^{n-2} \frac{\gamma(n-1, n q(\alpha))}{(n q(\alpha))^{n-1}} d\alpha &\geq \int_1^\infty c_1 \alpha^{n-2} \frac{\gamma(n-1, n q(\alpha))}{(n q(\alpha))^{n-1}} d\alpha \\ &\geq c_1 \int_1^\infty \alpha^{n-2} L_3 \left(\frac{L_2}{n\alpha} \right)^{n-1} d\alpha = \frac{c_1 L_3 L_2^{n-1}}{n^{n-1}} \int_1^\infty \frac{1}{\alpha} d\alpha \quad (\text{A.16}) \\ &= \frac{c_1 L_3 L_2^{n-1}}{n^{n-1}} \times \infty = \infty. \end{aligned}$$

Resumindo os resultados discutidos, temos que,

$$\begin{aligned} d_1 &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu d\phi d\alpha \\ &= \int_0^\infty \int_0^\infty \alpha^{n-2} \frac{\Gamma(n\phi)}{\phi \Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha \geq \int_0^\infty c_1 \alpha^{n-2} \frac{\gamma(n-1, nq(\alpha))}{(nq(\alpha))^{n-1}} d\alpha = \infty \end{aligned}$$

e provamos que

$$d_1 = \int_{\mathcal{A}} \frac{\alpha^{n-1}}{\phi \Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} = \infty .$$

Parte II

Queremos agora demonstrar que

$$d_2 = \int_{\mathcal{A}} \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} = \infty ,$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$.

Como $\frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} \geq 0$ sempre, pelo teorema de Tonelli (ver Folland, 1999) temos que

$$\begin{aligned} d_2 &= \int_{\mathcal{A}} \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu d\phi d\alpha . \end{aligned}$$

Assim, primeiramente observemos que, por meio do resultado (5.4) temos que

$$\begin{aligned} &\int_0^\infty \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu \\ &= \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \int_0^\infty \mu^{n\alpha\phi-1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu \\ &= \alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} . \end{aligned}$$

Agora queremos avaliar a integral deste resultado em relação à ϕ . Para isto, será necessário provarmos algumas desigualdades.

Primeiro mostremos que $\lim_{\phi \rightarrow 0^+} \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} > 0$. Pela fórmula de reflexão da função trigama temos

$$\psi'(\phi) = \frac{\pi^2}{\text{sen}^2(\pi\phi)} - \psi'(1 - \phi). \quad (\text{A.17})$$

Assim, temos que

$$\begin{aligned} \lim_{\phi \rightarrow 0^+} \phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1 &= \lim_{\phi \rightarrow 0^+} \phi^2 \left(\frac{\pi^2}{\text{sen}^2(\pi\phi)} - \psi'(1 - \phi) \right)^2 - \left(\frac{\pi^2}{\text{sen}^2(\pi\phi)} - \psi'(1 - \phi) \right) \\ &\quad - 1 \\ &= \lim_{\phi \rightarrow 0^+} \phi^2 \frac{\pi^4}{\text{sen}^4(\pi\phi)} - 2\phi^2 \frac{\pi^2}{\text{sen}^2(\pi\phi)} \psi'(1 - \phi) + \phi^2 \psi'(1 - \phi)^2 \\ &\quad - \frac{\pi^2}{\text{sen}^2(\pi\phi)} + \psi'(1 - \phi) - 1 \\ &= \lim_{\phi \rightarrow 0^+} \left(\phi^2 \frac{\pi^4}{\text{sen}^4(\pi\phi)} - \frac{\pi^2}{\text{sen}^2(\pi\phi)} \right) - 2\psi'(1) + 0 + \psi'(1) - 1 \\ &= \lim_{\phi \rightarrow 0^+} \left(\frac{\pi^2}{\text{sen}^2(\pi\phi)} \left(\frac{\pi^2}{\text{sen}^2(\pi\phi)} - 1 \right) \right) - \psi'(1) - 1. \end{aligned}$$

Então calculemos o primeiro limite:

$$\begin{aligned} \lim_{\phi \rightarrow 0^+} \left(\frac{\pi^2}{\text{sen}^2(\pi\phi)} \left(\frac{\pi^2}{\text{sen}^2(\pi\phi)} - 1 \right) \right) &= \lim_{\phi \rightarrow 0^+} \pi^2 \frac{(\pi^2 \phi^2 - \text{sen}^2(\pi\phi))}{\text{sen}^4(\pi\phi)} \\ &= \lim_{\phi \rightarrow 0^+} \frac{1}{\pi^2} \frac{(\pi\phi)^4}{\text{sen}^4(\pi\phi)} \frac{(\pi^2 \phi^2 - \text{sen}^2(\pi\phi))}{\phi^4} \\ &= \lim_{\phi \rightarrow 0^+} \frac{1}{\pi^2} \frac{(\pi^2 \phi^2 - \text{sen}^2(\pi\phi))}{\phi^4} \\ &\stackrel{L'h}{=} \lim_{\phi \rightarrow 0^+} \frac{1}{\pi^2} \frac{(2\pi^2 \phi - 2\pi \cos(\pi\phi) \text{sen}(\pi\phi))}{4\phi^3} \\ &\stackrel{L'h}{=} \lim_{\phi \rightarrow 0^+} \frac{1}{\pi^2} \frac{(2\pi^2 + 2\pi^2 \text{sen}^2(\pi\phi) - 2\pi^2 \cos^2(\pi\phi))}{12\phi^2} \\ &= \lim_{\phi \rightarrow 0^+} \frac{(\text{sen}^2(\pi\phi) + \cos^2(\pi\phi) + \text{sen}^2(\pi\phi) - \cos^2(\pi\phi))}{6\phi^2} \\ &= \lim_{\phi \rightarrow 0^+} \frac{\text{sen}^2(\pi\phi)}{3\phi^2} = \frac{\pi^2}{3} \lim_{\phi \rightarrow 0^+} \frac{\text{sen}^2(\pi\phi)}{(\pi\phi)^2} = \frac{\pi^2}{3} > 0. \end{aligned}$$

Como $\psi'(1) = \frac{\pi^2}{6}$, então

$$\lim_{\phi \rightarrow 0^+} \phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1 = \frac{\pi^2}{3} - \frac{\pi^2}{6} - 1 = \frac{\pi^2}{6} - 1 > 0.$$

Desta forma mostramos que

$$\lim_{\phi \rightarrow 0^+} \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} = \sqrt{\frac{\pi^2}{6} - 1} > 0. \quad (\text{A.18})$$

Assim, segue que $\lim_{\phi \rightarrow 0^+} \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} > 0$ e pelo Teorema da Conservação do Sinal existe um $0 < s \leq 1$ tal que $\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} > 0, \forall \phi \in (0, s]$.

Deste modo, como $\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}$ é contínua e maior que 0 em $\phi \in (0, s]$, existe L_4 tal que $\sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} > L_4, \forall \phi \in (0, s]$.

Logo, temos que

$$\begin{aligned} \int_0^\infty \alpha^{n-1} \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi &\geq \\ \int_0^s \alpha^{n-1} \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi &\geq L_4 \alpha^{n-1} \int_0^s \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi. \end{aligned}$$

Mas por (A.5), temos que, para $0 \leq \phi \leq s \leq 1$,

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n} > \frac{L_1}{n} \phi^{n-1}.$$

Desta forma, por (A.7), temos que

$$\begin{aligned} L_4 \alpha^{n-1} \int_0^s \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi &\geq \frac{L_4 L_1}{n \prod_{i=1}^n t_i} \alpha^{n-1} \int_0^s \phi^{n-1} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \\ &= c_2 \alpha^{n-1} \int_0^s \phi^{n-1} \exp(-n q(\alpha) \phi) d\phi = c_2 \alpha^{n-1} \frac{\gamma(n, sn q(\alpha))}{(n q(\alpha))^n}. \end{aligned}$$

onde $c_2 = \frac{L_4 L_1}{n \prod_{i=1}^n t_i}$.

Agora queremos avaliar a integral do resultado acima em relação à α . Para isto, será necessário provarmos algumas desigualdades.

Utilizando-se o resultado (A.13), como já vimos em (A.12) que $\lim_{\alpha \rightarrow \infty} q(\alpha) = \infty$, segue que

$$\lim_{\alpha \rightarrow \infty} \gamma(n, sn q(\alpha)) = \Gamma(n) \quad (\text{A.19})$$

Sendo também que $\gamma(n, sn q(\alpha))$ é contínua e maior que 0 em $[1, \infty)$, pelo Lema 2 temos que

$$\gamma(n, sn q(\alpha)) > L_5, \forall \alpha \in [1, \infty) \quad (\text{A.20})$$

em que L_5 é uma constante maior que 0.

Utilizando (A.20) e (A.10), temos que

$$\begin{aligned} \int_0^\infty c_2 \alpha^{n-1} \frac{\gamma(n, sn q(\alpha))}{(n q(\alpha))^n} d\alpha &\geq \int_1^\infty c_2 \alpha^{n-1} \frac{\gamma(n, sn q(\alpha))}{(n q(\alpha))^n} d\alpha \geq \int_1^\infty c_2 \alpha^{n-1} L_5 \left(\frac{L_2}{n\alpha}\right)^n d\alpha \\ &= \frac{c_2 L_5 L_2^n}{n^n} \int_1^\infty \frac{1}{\alpha} d\alpha = \frac{c_2 L_5 L_2^n}{n^n} \times \infty = \infty \end{aligned}$$

Resumindo,

$$\begin{aligned}
d_2 &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu d\phi d\alpha \\
&\geq \int_0^\infty \int_0^\infty \alpha^{n-1} \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha \\
&\geq \int_0^\infty c_2 \alpha^{n-1} \frac{\gamma(n, \text{sn q}(\alpha))}{(n \text{ q}(\alpha))^n} d\alpha = \infty.
\end{aligned}$$

Desta forma, provamos que

$$d_2 = \int_{\mathcal{A}} \frac{\alpha^n \sqrt{\phi^2 \psi'(\phi)^2 - \psi'(\phi) - 1}}{\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} = \infty,$$

Parte III

Queremos agora demonstrar que

$$d_3 = \int_{\mathcal{A}} \frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} d\boldsymbol{\theta} = \infty.$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$.

Como $\frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \geq 0$ sempre, pelo teorema de Tonelli (ver Folland, 1999) temos que

$$\begin{aligned}
d_3 &= \int_{\mathcal{A}} \frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} d\boldsymbol{\theta} \\
&= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} d\mu d\phi d\alpha.
\end{aligned}$$

Assim, primeiramente, veja que

$$\begin{aligned}
&\int_0^\infty \frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} d\mu = \\
&\frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left(\psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right) \int_0^\infty \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu.
\end{aligned} \tag{A.21}$$

Utilizando-se o resultado (5.3) e após algumas manipulações algébricas, obtemos que

$$\int_0^\infty \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu = \frac{\Gamma(n\phi + \frac{2}{\alpha})}{\alpha (\sum_{i=1}^n t_i^\alpha)^{n\phi + \frac{2}{\alpha}}}. \tag{A.22}$$

Substituindo-se a solução da integral (A.22) em (A.21), obtemos a expressão

$$\frac{\alpha^n \Gamma\left(n\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)^{n+1} \left(\sum_{i=1}^n t_i^\alpha\right)^{n\phi + \frac{2}{\alpha}}} \prod_{i=1}^n t_i^{\alpha\phi-1} \exp\left(\psi(\phi) \left(\phi - \frac{1}{\alpha}\right) - \phi\right). \quad (\text{A.23})$$

Agora queremos avaliar a integral do resultado acima em relação à ϕ . Para isto, será necessário provarmos algumas desigualdades.

Consideraremos aqui $1/2 \leq \alpha \leq 1$. Pela fórmula de reflexão da função digama temos que:

$$\psi(\phi) = \psi(1 - \phi) - \pi \cot(\pi\phi) = \psi(1 - \phi) - \frac{\pi \cos(\pi\phi)}{\sin(\pi\phi)} \quad (\text{A.24})$$

Assim, temos,

$$\begin{aligned} \lim_{\phi \rightarrow 0^+} \frac{\psi(\phi)}{-\frac{1}{\phi}} &= \lim_{\phi \rightarrow 0^+} \frac{\psi(1 - \phi) - \frac{\pi \cos(\pi\phi)}{\sin(\pi\phi)}}{-\frac{1}{\phi}} = \lim_{\phi \rightarrow 0^+} \left(-\phi\psi(1 - \phi) + \frac{\pi\phi}{\sin(\pi\phi)} \cos(\pi\phi) \right) \\ &= 0 \times (-\psi(1)) + 1 \times 1 = 1 \end{aligned}$$

Assim, temos que

$$\lim_{\phi \rightarrow 0^+} \frac{\psi(\phi) \left(\phi - \frac{1}{\alpha}\right) - \phi}{\frac{1}{\phi}} = \lim_{\phi \rightarrow 0^+} -\frac{\psi(\phi)}{-\frac{1}{\phi}} \left(\phi - \frac{1}{\alpha}\right) - \phi^2 = -1 \times \left(0 - \frac{1}{\alpha}\right) - 0 = \frac{1}{\alpha} > 0$$

Como o limite é positivo, pelo Teorema da Conservação do Sinal (ver Lima, 2011), existe $b_\alpha > 0$ tal que, $\frac{\psi(\phi) \left(\phi - \frac{1}{\alpha}\right) - \phi}{\frac{1}{\phi}} > 0$ para todo $\phi \in (0, b_\alpha]$ (veja que b_α só depende de α).

Como também $\frac{\psi(\phi) \left(\phi - \frac{1}{\alpha}\right) - \phi}{\frac{1}{\phi}}$ é contínua em $(0, b_\alpha]$, pelo Lema 1, existe uma $B_\alpha > 0$, constante em relação à ϕ , tal que

$$\frac{\psi(\phi) \left(\phi - \frac{1}{\alpha}\right) - \phi}{\frac{1}{\phi}} > B_\alpha \Leftrightarrow \psi(\phi) \left(\phi - \frac{1}{\alpha}\right) - \phi > B_\alpha \frac{1}{\phi}, \forall \phi \in (0, b_\alpha].$$

Por outro lado, considere o limite

$$\lim_{\phi \rightarrow 0^+} \frac{e^{B_\alpha \frac{1}{\phi}}}{\frac{1}{\phi^{n+1}}}.$$

Fazendo a substituição $u = B_\alpha \frac{1}{\phi} \Leftrightarrow \phi = B_\alpha \frac{1}{u}$ acima, segue que

$$\lim_{\phi \rightarrow 0^+} \frac{e^{B_\alpha \frac{1}{\phi}}}{\frac{1}{\phi^{n+1}}} = \lim_{u \rightarrow \infty} B_\alpha^{n+1} \frac{e^u}{u^{n+1}} = B_\alpha^{n+1} \times \infty = \infty > 0.$$

Como $\frac{e^{B_\alpha \frac{1}{\phi}}}{\frac{1}{\phi^{n+1}}}$ é contínuo e maior que 0 em $(0, b_\alpha]$, então segue pelo Lema 1 que $\frac{e^{B_\alpha \frac{1}{\phi}}}{\frac{1}{\phi^{n+1}}} > C_\alpha \Leftrightarrow e^{B_\alpha \frac{1}{\phi}} > C_\alpha \frac{1}{\phi^{n+1}}$, para alguma $C_\alpha > 0$, constante em relação a ϕ .

Assim, temos

$$\begin{aligned} & \alpha^n \frac{\Gamma\left(n\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)^{n+1}} \frac{\prod_{i=1}^n t_i^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi+\frac{2}{\alpha}}} \exp\left(\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi\right) \geq \\ & \frac{\alpha^n}{(\prod_{i=1}^n t_i) (\sum_{i=1}^n t_i^\alpha)^{\frac{2}{\alpha}}} \frac{\Gamma\left(n\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)^{n+1}} \frac{\prod_{i=1}^n t_i^{\alpha\phi}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} \exp\left(B_\alpha \frac{1}{\phi}\right) \geq \\ & \frac{\alpha^n C_\alpha}{(\prod_{i=1}^n t_i) (\sum_{i=1}^n t_i^\alpha)^{\frac{2}{\alpha}}} \frac{\Gamma\left(n\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)^{n+1}} \frac{\prod_{i=1}^n t_i^{\alpha\phi}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} \frac{1}{\phi^{n+1}}. \end{aligned}$$

Pode-se demonstrar facilmente que a função $\Gamma(s)$ é crescente para $s \geq 2$.

Por outro lado, como $\alpha \leq 1$, temos que $\frac{2}{\alpha} \geq 2$, de onde $n\phi + \frac{2}{\alpha} \geq n\phi + 2 \geq 2, \forall \phi > 0$. Assim segue que $\Gamma\left(n\phi + \frac{2}{\alpha}\right) \geq \Gamma(n\phi + 2) = (n\phi + 1)n\phi\Gamma(n\phi) \geq n\phi\Gamma(n\phi)$.

Agora, como $\frac{(\prod_{i=1}^n t_i)^{\alpha\phi}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}}$ é contínuo e maior que 0 em $[0, b_\alpha]$, pelo Lema 1 temos que, $\frac{(\prod_{i=1}^n t_i)^{\alpha\phi}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} \geq D_\alpha$ para alguma $D_\alpha > 0$, constante em relação a ϕ .

Destes fatos e de (A.5) segue que

$$\begin{aligned} & \frac{\alpha^n C_\alpha}{(\prod_{i=1}^n t_i) (\sum_{i=1}^n t_i^\alpha)^{\frac{2}{\alpha}}} \frac{\Gamma\left(n\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)^{n+1}} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} \frac{1}{\phi^{n+1}} \geq \frac{\alpha^n C_\alpha D_\alpha}{(\prod_{i=1}^n t_i) (\sum_{i=1}^n t_i^\alpha)^{\frac{2}{\alpha}}} \frac{n\phi\Gamma(n\phi)}{\Gamma(\phi)^{n+1}} \frac{1}{\phi^{n+1}} \\ & \geq \frac{\alpha^n C_\alpha D_\alpha}{(\prod_{i=1}^n t_i) (\sum_{i=1}^n t_i^\alpha)^{\frac{2}{\alpha}}} n\phi^n \frac{L_1}{n} \frac{1}{\phi^{n+1}} \\ & = c_3 \frac{1}{\phi^n} \end{aligned}$$

onde $c_3 = \frac{\alpha^n C_\alpha D_\alpha L_1}{(\prod_{i=1}^n t_i) (\sum_{i=1}^n t_i^\alpha)^{\frac{2}{\alpha}}}$ é constante em relação a ϕ .

Assim, para $\frac{1}{2} \leq \alpha \leq 1$, temos que

$$\begin{aligned} & \int_0^\infty \frac{\alpha^n \Gamma\left(n\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)^{n+1} (\sum_{i=1}^n t_i^\alpha)^{n\phi+\frac{2}{\alpha}}} \prod_{i=1}^n t_i^{\alpha\phi-1} \exp\left(\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi\right) d\phi \geq \\ & \int_0^{b_\alpha} \frac{\alpha^n \Gamma\left(n\phi + \frac{2}{\alpha}\right)}{\Gamma(\phi)^{n+1} (\sum_{i=1}^n t_i^\alpha)^{n\phi+\frac{2}{\alpha}}} \prod_{i=1}^n t_i^{\alpha\phi-1} \exp\left(\psi(\phi)\left(\phi - \frac{1}{\alpha}\right) - \phi\right) d\phi \geq \\ & c_3 \int_0^{b_\alpha} \frac{1}{\phi} d\phi = c_3 \times \infty = \infty. \end{aligned}$$

Portanto

$$\begin{aligned}
 d_3 &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} d\mu d\phi d\alpha \\
 &= \int_0^\infty \int_0^\infty \frac{\alpha^n \Gamma \left(n\phi + \frac{2}{\alpha} \right)}{\Gamma(\phi)^{n+1} (\sum_{i=1}^n t_i^\alpha)^{n\phi + \frac{2}{\alpha}}} \prod_{i=1}^n t_i^{\alpha\phi-1} \exp \left(\psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right) d\phi d\alpha \\
 &\geq \int_{1/2}^1 \int_0^\infty \frac{\alpha^n \Gamma \left(n\phi + \frac{2}{\alpha} \right)}{\Gamma(\phi)^{n+1} (\sum_{i=1}^n t_i^\alpha)^{n\phi + \frac{2}{\alpha}}} \prod_{i=1}^n t_i^{\alpha\phi-1} \exp \left(\psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right) d\phi d\alpha \\
 &= \int_{1/2}^1 \infty d\alpha = \infty.
 \end{aligned}$$

Observe que tem sentido falar em $\int_{1/2}^1 \infty d\alpha$ pois estamos trabalhando com integrais de Lebesgue, e aqui ∞ representa a função Lebesgue mensurável $f : [0, \infty) \rightarrow [0, \infty]$, dada por $f(\alpha) = \infty$. Veja também que integrais impróprias estão contempladas no teorema de Tonelli descrito acima.

Segue então que

$$d_3 = \int_{\mathcal{A}} \frac{\alpha^{n+1}}{\Gamma(\phi)^{n+1}} \mu^{n\alpha\phi+1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha + \psi(\phi) \left(\phi - \frac{1}{\alpha} \right) - \phi \right\} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} d\theta = \infty.$$

APÊNDICE B – Demonstração de que a densidade a posteriori 4.20 é própria

Lema 3: Seja $h(x)$ uma função contínua e maior que 0 em $(a, b]$ tal que

$$\lim_{x \rightarrow a^+} h(x) < \infty.$$

Então $h(x)$ é limitado superiormente em $(a, b]$, isto é, existe $L > 0$ tal que $h(x) < L, \forall x \in (a, b]$.

Lema 4: Seja $h(x)$ uma função contínua e maior que 0 em $[b, \infty)$ tal que

$$\lim_{x \rightarrow \infty} h(x) < \infty.$$

Então $h(x)$ é limitado superiormente em $(a, b]$, isto é, existe $L > 0$ tal que $h(x) < L, \forall x \in [b, \infty)$.

Queremos demonstrar que:

$$d_4 = \int_{\mathcal{A}} \frac{\alpha^{n-\frac{1}{2}-\alpha/(1+\alpha)}}{\phi\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} < \infty$$

em que $\boldsymbol{\theta} = (\phi, \mu, \alpha)$ e $\mathcal{A} = \{(0, \infty) \times (0, \infty) \times (0, \infty)\}$ é o espaço paramétrico de $\boldsymbol{\theta}$.

Como $\frac{\alpha^{n-\frac{1}{2}-\alpha/(1+\alpha)}}{\phi\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} \geq 0$ sempre, pelo teorema de Tonelli (ver Folland, 1999) temos que

$$\begin{aligned} d_4 &= \int_{\mathcal{A}} \frac{\alpha^{n-\frac{1}{2}-\alpha/(1+\alpha)}}{\phi\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\boldsymbol{\theta} \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\alpha^{n-\frac{1}{2}-\alpha/(1+\alpha)}}{\phi\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu d\phi d\alpha. \end{aligned}$$

Assim, primeiramente por (5.3), veja que

$$\begin{aligned} &\int_0^\infty \frac{\alpha^{n-\frac{1}{2}-\alpha/(1+\alpha)}}{\phi\Gamma(\phi)^n} \mu^{n\alpha\phi-1} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu = \\ &\frac{\alpha^{n-\frac{1}{2}-\alpha/(1+\alpha)}}{\phi\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \int_0^\infty \mu^{n\alpha\phi-1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n t_i^\alpha \right\} d\mu = \\ &\frac{\alpha^{n-\frac{3}{2}-\alpha/(1+\alpha)}}{\phi\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}}. \end{aligned}$$

Assim, segue que

$$d_4 = \int_0^\infty \int_0^\infty \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha = s_1 + s_2 + s_3 + s_4. \quad (\text{B.1})$$

Onde

$$\begin{aligned} s_1 &= \int_0^1 \int_0^1 \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha, \\ s_2 &= \int_1^\infty \int_0^1 \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha, \\ s_3 &= \int_0^1 \int_1^\infty \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha, \\ s_4 &= \int_1^\infty \int_1^\infty \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha. \end{aligned}$$

Assim, para mostrarmos que $d_4 < \infty$, basta mostrarmos que $s_1 < \infty$, $s_2 < \infty$, $s_3 < \infty$ e que $s_4 < \infty$.

Consideremos inicialmente o caso s_1 :

Claramente $\frac{n\phi\Gamma(n\phi)}{(\phi\Gamma(\phi))^n}$ é contínua em $(0, 1]$. Juntando este fato ao limite obtido em (A.4) temos que $\frac{n\phi\Gamma(n\phi)}{(\phi\Gamma(\phi))^n}$ satisfaz o Lema 3, isto é existe $L_1 > 0$ tal que $\frac{n\phi\Gamma(n\phi)}{(\phi\Gamma(\phi))^n} < L_1, \forall \phi \in (0, 1]$. Desta forma,

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n} < L_1 \frac{\phi^n}{n\phi} = \frac{L_1}{n} \phi^{n-1}, \forall \phi \in (0, 1]. \quad (\text{B.2})$$

Lembremos também, por (A.6) que, dado $q(\alpha) = \log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right)$, então $q(\alpha) > 0$, e, por (A.7) temos:

$$\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} = \exp(-q(\alpha)). \quad (\text{B.3})$$

Assim, segue por (B.2 e B.3) que

$$\begin{aligned} \int_0^1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\Gamma(n\phi)}{\phi \Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi &< \int_0^1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{1}{\prod_{i=1}^n t_i} \frac{L_1}{n} \phi^{n-2} \frac{(\prod_{i=1}^n t_i^\alpha)^\phi}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi \\ &= \frac{L_1}{n (\prod_{i=1}^n t_i)} \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \int_0^1 \phi^{n-2} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi \\ &= \frac{L_1}{n (\prod_{i=1}^n t_i)} \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \int_0^1 \phi^{n-2} \exp(-n\phi q(\alpha)) d\phi \end{aligned}$$

Tomando $c_1 = \frac{L_1}{n(\prod_{i=1}^n t_i)}$ e utilizando a mesma integração realizada em (A.8), segue que

$$\begin{aligned} \int_0^1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\Gamma(n\phi)}{\phi\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi &< \frac{L_1}{n(\prod_{i=1}^n t_i)} \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \int_0^1 \phi^{n-2} \exp(-n\phi q(\alpha)) d\phi \\ &= c_1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\gamma(n-1, nq(\alpha))}{(nq(\alpha))^{n-1}}. \end{aligned} \quad (\text{B.4})$$

Agora queremos avaliar a integral da expressão anterior em relação à α variando de 0 à 1. Para isto, primeiramente observemos que:

$$\lim_{\alpha \rightarrow 0^+} q(\alpha) = \lim_{\alpha \rightarrow 0^+} \log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) = \log \left(\frac{n}{1} \right) = \log(n) > 0. \quad (\text{B.5})$$

Assim, também:

$$\lim_{\alpha \rightarrow 0^+} \frac{\gamma(n-1, nq(\alpha))}{(nq(\alpha))^{n-1}} = \frac{\gamma(n-1, n\log(n))}{(n\log(n))^{n-1}} > 0. \quad (\text{B.6})$$

Assim, como $\frac{\gamma(n-1, nq(\alpha))}{(nq(\alpha))^{n-1}}$ é contínuo em $(0, 1]$, então pelo Lema 3 temos que existe $L_2 > 0$ tal que $\frac{\gamma(n-1, nq(\alpha))}{(nq(\alpha))^{n-1}} < L_2, \forall \alpha \in (0, 1]$.

Agora, é fácil ver que, para $\alpha > 0$, $n - \frac{3}{2} - \frac{\alpha}{(1+\alpha)}$ é decrescente em função de α , pois $-\frac{\alpha}{(1+\alpha)} = -1 + \frac{1}{1+\alpha}$. Disto segue que, para $\alpha > 0$ temos que $n - \frac{3}{2} - \frac{\alpha}{(1+\alpha)} < n - \frac{3}{2}$. Logo,

$$n - \frac{3}{2} - \frac{\alpha}{(1+\alpha)} < n - \frac{3}{2}, \text{ para } \alpha > 0. \quad (\text{B.7})$$

Segue então que:

$$\int_0^1 c_1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\gamma(n-1, nq(\alpha))}{(nq(\alpha))^{n-1}} d\alpha < c_1 L_2 \int_0^1 \alpha^{n-\frac{3}{2}} d\alpha < \infty, \quad (\text{B.8})$$

pois $n - \frac{3}{2} > 0$.

Logo,

$$\begin{aligned} s_1 &= \int_0^1 \int_0^1 \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha \\ &< \int_0^1 c_1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\gamma(n-1, nq(\alpha))}{(nq(\alpha))^{n-1}} d\alpha < \infty. \end{aligned}$$

Consideremos agora o caso s_2 :

Por (B.4) temos que

$$\int_0^1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\Gamma(n\phi)}{\phi\Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi < c_1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\gamma(n-1, n\mathbf{q}(\alpha))}{(n\mathbf{q}(\alpha))^{n-1}}.$$

Agora queremos avaliar a integral da expressão anterior em relação à α de 1 à ∞ . Para isto, primeiramente observemos que, por (A.11) temos:

$$\lim_{\alpha \rightarrow \infty} \frac{\mathbf{q}(\alpha)}{\alpha} = v > 0.$$

Logo,

$$\lim_{\alpha \rightarrow \infty} \frac{\alpha}{\mathbf{q}(\alpha)} = \frac{1}{v} < \infty.$$

Assim, como $\frac{\alpha}{\mathbf{q}(\alpha)}$ é contínuo em $[1, \infty)$ segue pelo Lema 4 que existe $L_3 > 0$ tal que

$$\frac{\alpha}{\mathbf{q}(\alpha)} < L_3 \Leftrightarrow \frac{1}{\mathbf{q}(\alpha)} < \frac{L_3}{\alpha}, \forall \alpha \in [1, \infty).$$

Agora, por (A.14) temos que:

$$\lim_{\alpha \rightarrow \infty} \gamma(n-1, n\mathbf{q}(\alpha)) = \Gamma(n-1) < \infty.$$

Assim, como $\gamma(n-1, n\mathbf{q}(\alpha))$ é contínuo em $[1, \infty)$ segue pelo Lema 4 que existe $L_4 > 0$ tal que $\gamma(n-1, n\mathbf{q}(\alpha)) < L_4$.

Por outro lado, temos que

$$\lim_{\alpha \rightarrow \infty} \frac{a^{-\frac{\alpha}{(1+\alpha)}}}{a^{-1}} = \lim_{\alpha \rightarrow \infty} a^{\frac{1}{(1+\alpha)}} = \lim_{\alpha \rightarrow \infty} a^{\frac{1}{\alpha} \left(\frac{1}{\alpha+1} \right)} = \lim_{\alpha \rightarrow \infty} \left(a^{\frac{1}{\alpha}} \right)^{\left(\frac{1}{\alpha+1} \right)} = 1^1 = 1.$$

Assim, como $\frac{a^{-\frac{\alpha}{(1+\alpha)}}}{a^{-1}}$ é contínua em $[1, \infty)$, pelo Lema 4, temos que existe $L_5 > 0$ tal que

$$\frac{a^{-\frac{\alpha}{(1+\alpha)}}}{a^{-1}} < L_5 \Leftrightarrow a^{-\frac{\alpha}{(1+\alpha)}} < a^{-1} L_5 \Leftrightarrow a^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} < a^{n-\frac{5}{2}} L_5, \forall \alpha \in [1, \infty). \quad (\text{B.9})$$

Temos então que:

$$\begin{aligned} \int_1^\infty c_1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\gamma(n-1, n\mathbf{q}(\alpha))}{(n\mathbf{q}(\alpha))^{n-1}} d\alpha &< c_1 \int_1^\infty L_5 \alpha^{n-\frac{5}{2}} L_4 \left(\frac{L_3}{n\alpha} \right)^{n-1} d\alpha \\ &< \frac{c_1 L_5 L_4 L_3^{n-1}}{n^{n-1}} \int_1^\infty \frac{1}{\alpha^{\frac{3}{2}}} d\alpha < \infty. \end{aligned}$$

Logo

$$\begin{aligned} s_2 &= \int_1^\infty \int_0^1 \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha \\ &< \int_1^\infty c_1 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\gamma(n-1, n\mathbf{q}(\alpha))}{(n\mathbf{q}(\alpha))^{n-1}} d\alpha < \infty. \end{aligned}$$

Consideremos agora o caso s_3 :

Sabemos pela aproximação de Stirling que:

$$\lim_{z \rightarrow \infty} \frac{\Gamma(z)}{\sqrt{2\pi} z^{z-\frac{1}{2}} e^{-z}} = 1$$

de onde

$$\lim_{z \rightarrow \infty} \frac{\Gamma(z)}{z^{z-\frac{1}{2}} e^{-z}} = \sqrt{2\pi}.$$

Assim temos que:

$$\lim_{\phi \rightarrow \infty} \frac{\frac{\Gamma(n\phi)}{(n\phi)^{n\phi-\frac{1}{2}} e^{-n\phi}}}{\left(\frac{\Gamma(\phi)}{\phi^{\phi-\frac{1}{2}} e^{-\phi}} \right)^n} = \frac{\sqrt{2\pi}}{\sqrt{2\pi}^n} = \sqrt{2\pi}^{1-n}.$$

Mas, simplificando a expressão, obtemos

$$\frac{\frac{\Gamma(n\phi)}{(n\phi)^{n\phi-\frac{1}{2}} e^{-n\phi}}}{\left(\frac{\Gamma(\phi)}{\phi^{\phi-\frac{1}{2}} e^{-\phi}} \right)^n} = n^{\frac{1}{2}} n^{-n\phi} \phi^{-\frac{n-1}{2}} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n}.$$

Assim segue que

$$\lim_{\phi \rightarrow \infty} n^{-n\phi} \phi^{-\frac{n-1}{2}} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} = n^{-\frac{1}{2}} \sqrt{2\pi}^{1-n} < \infty.$$

E, como $n^{-n\phi} \phi^{-\frac{n-1}{2}} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n}$ é contínua em $[1, \infty)$, segue que pelo Lema 4 que existe $L_6 > 0$ tal que

$$n^{-n\phi} \phi^{-\frac{n-1}{2}} \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} < L_6 \Leftrightarrow \frac{\Gamma(n\phi)}{\Gamma(\phi)^n} < n^{n\phi} \phi^{\frac{n-1}{2}} L_6, \forall \phi \in [1, \infty).$$

Assim, temos que:

$$\begin{aligned}
 & \int_1^\infty \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\Gamma(n\phi)}{\phi \Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi < \\
 & \frac{L_6}{\prod_{i=1}^n t_i} \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \int_1^\infty \phi^{\frac{n-3}{2}} n^{n\phi} \frac{(\prod_{i=1}^n t_i^\alpha)^\phi}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi = \\
 & c_2 \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \int_1^\infty \phi^{\frac{n-3}{2}} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\frac{1}{n} \sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi.
 \end{aligned} \tag{B.10}$$

onde $c_2 = \frac{L_6}{\prod_{i=1}^n t_i}$ é uma constante.

Como por hipótese nem todos t_i 's são iguais, temos que nem todos t_i^α 's são iguais, logo segue da desigualdade estrita das médias que:

$$\frac{1}{n} \sum_{i=1}^n t_i^\alpha > \sqrt[n]{\prod_{i=1}^n t_i^\alpha} \Leftrightarrow \frac{\frac{1}{n} \sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} > 1 \Leftrightarrow \log \left(\frac{\frac{1}{n} \sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) > 0. \tag{B.11}$$

Logo, tomando $p(\alpha) = \log \left(\frac{\frac{1}{n} \sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) > 0$, temos então:

$$\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\frac{1}{n} \sum_{i=1}^n t_i^\alpha} = \exp \left(\log \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\frac{1}{n} \sum_{i=1}^n t_i^\alpha} \right) \right) = \exp \left(\log \left(-\frac{\frac{1}{n} \sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) \right) = \exp(-p(\alpha)).$$

Assim segue que:

$$\begin{aligned}
 c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \int_1^\infty \phi^{\frac{n-3}{2}} \left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\frac{1}{n} \sum_{i=1}^n t_i^\alpha} \right)^{n\phi} d\phi &= c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \int_1^\infty \phi^{\frac{n-1}{2}-1} \exp(-n p(\alpha)\phi) d\phi \\
 &= c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \frac{\Gamma(\frac{n-1}{2}, n p(\alpha))}{(n p(\alpha))^{\frac{n-1}{2}}}.
 \end{aligned} \tag{B.12}$$

Agora queremos avaliar a integral da expressão anterior em relação à α variando de 1 à ∞ .

Primeiramente, tome, para cada $1 \leq i \leq n$, $u_i = \frac{t_i}{\sqrt[n]{\prod_{i=1}^n t_i}}$, assim pode-se verificar que:

$$q(\alpha) = \log \left(\frac{\frac{1}{n} \sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) = \log \left(\frac{1}{n} \sum_{i=1}^n u_i^\alpha \right).$$

E que:

$$\sum_{i=1}^n \log(u_i) = \log \left(\frac{\prod_{i=1}^n t_i}{\prod_{i=1}^n t_i} \right) = \log(1) = 0.$$

Assim, segue que:

$$\begin{aligned} \lim_{\alpha \rightarrow 0^+} \frac{p(\alpha)}{\alpha^2} &= \frac{\log\left(\frac{1}{n} \sum_{i=1}^n u_i^\alpha\right)}{\alpha^2} \stackrel{L'h}{=} \lim_{\alpha \rightarrow 0^+} \frac{\frac{\sum_{i=1}^n \log(u_i) u_i^\alpha}{\sum_{i=1}^n u_i^\alpha}}{2\alpha} \\ &\stackrel{L'h}{=} \frac{1}{2} \lim_{\alpha \rightarrow 0^+} \frac{(\sum_{i=1}^n \log(u_i)^2 u_i^\alpha)(\sum_{i=1}^n u_i^\alpha) - (\sum_{i=1}^n \log(u_i) u_i^\alpha)^2}{(\sum_{i=1}^n u_i^\alpha)^2} \\ &= \frac{1}{2} \frac{n \sum_{i=1}^n \log(u_i)^2 - (\sum_{i=1}^n \log(u_i))^2}{n^2} = \frac{1}{2} \frac{\sum_{i=1}^n \log(u_i)^2}{n} > 0, \end{aligned}$$

pois para termos $\frac{1}{2} \frac{\sum_{i=1}^n \log(u_i)^2}{n} = 0$ teríamos de ter $\log(u_i) = 0 \Leftrightarrow u_i = 1, \forall i$, o que nunca acontece, pois por hipótese nem todos os t_i 's são iguais e logo nem todos os u_i 's são iguais.

Assim segue que:

$$\lim_{\alpha \rightarrow 0^+} \frac{p(\alpha)}{\alpha^2} > 0 \Leftrightarrow \lim_{\alpha \rightarrow 0^+} \frac{\alpha^2}{p(\alpha)} < \infty.$$

Assim, como $\frac{\alpha^2}{p(\alpha)}$ é contínua em $(0, 1]$ segue pelo Lema 3 que existe L_7 tal que $\frac{\alpha^2}{p(\alpha)} < L_7 \Leftrightarrow \frac{1}{p(\alpha)} < L_7 \frac{1}{\alpha^2}, \forall \alpha \in (0, 1]$.

Segue também que

$$\lim_{\alpha \rightarrow 0^+} p(\alpha) = \lim_{\alpha \rightarrow 0^+} \alpha^2 \frac{p(\alpha)}{\alpha^2} = 0 \times \frac{1}{2} \frac{\sum_{i=1}^n \log(u_i)^2}{n} = 0.$$

Por outro lado, por definição, é fácil provar que:

$$\lim_{x \rightarrow 0^+} \Gamma(s, x) = \Gamma(s).$$

Assim, segue que:

$$\lim_{\alpha \rightarrow 0^+} \Gamma\left(\frac{n-1}{2}, n p(\alpha)\right) = \Gamma\left(\frac{n-1}{2}\right) < \infty.$$

Assim, como $\Gamma\left(\frac{n-1}{2}, n p(\alpha)\right)$ é contínua em $(0, 1]$, pelo Lema 3 existe L_8 tal que $\Gamma\left(\frac{n-1}{2}, n p(\alpha)\right) < L_8, \forall \alpha \in (0, 1]$.

Por outro lado, observe que:

$$\lim_{\alpha \rightarrow 0^+} \alpha^{-\frac{\alpha}{(1+\alpha)}} = \lim_{\alpha \rightarrow 0^+} \left(\frac{1}{\alpha^\alpha}\right)^{\frac{1}{(1+\alpha)}} = \left(\frac{1}{1}\right)^1 = 1.$$

Assim, como $\alpha^{-\frac{\alpha}{(1+\alpha)}}$ é contínua em $(0, 1]$ então pelo Lema 3 existe um $L_9 > 0$ tal que $\alpha^{-\frac{\alpha}{(1+\alpha)}} < L_9, \forall \alpha \in (0, 1]$.

Segue então:

$$\begin{aligned} \int_0^1 c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \frac{\Gamma(\frac{n-1}{2}, n p(\alpha))}{(n p(\alpha))^{\frac{n-1}{2}}} d\alpha &< c_2 \int_0^1 L_9 \alpha^{n-\frac{3}{2}} L_8 \left(\frac{L_7}{n \alpha^2} \right)^{\frac{n-1}{2}} d\alpha \\ &< \frac{c_2 L_9 L_8 L_7^{\frac{n-1}{2}}}{n^{\frac{n-1}{2}}} \int_0^1 \alpha^{-\frac{1}{2}} d\alpha < \infty, \end{aligned} \quad (\text{B.13})$$

pois $\int_0^1 \alpha^{-\frac{1}{2}} d\alpha < \infty$

Logo

$$\begin{aligned} s_3 &= \int_0^1 \int_1^\infty \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha, \\ &< \int_0^1 c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \frac{\Gamma(\frac{n-1}{2}, n p(\alpha))}{(n p(\alpha))^{\frac{n-1}{2}}} d\alpha < \infty. \end{aligned}$$

Consideramos agora o caso s_4 :

Por meio dos resultados obtidos em (B.10 e B.12) temos que

$$\int_1^\infty \alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \frac{\Gamma(n\phi)}{\phi \Gamma(\phi)^n} \frac{(\prod_{i=1}^n t_i)^{\alpha\phi-1}}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi < c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \frac{\Gamma(\frac{n-1}{2}, n p(\alpha))}{(n p(\alpha))^{\frac{n-1}{2}}}, \quad (\text{B.14})$$

onde $p(\alpha) = \log \left(\frac{\frac{1}{n} \sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) > 0$.

Agora queremos avaliar a integral da expressão (B.14) em relação à α variando de 1 à ∞ . Por (A.11) sabemos que:

$$\lim_{\alpha \rightarrow \infty} \frac{q(\alpha)}{\alpha} = v > 0.$$

onde $q(\alpha) = \log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) > 0$.

$$\text{Como } p(\alpha) = \log \left(\frac{\frac{1}{n} \sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) = \log \left(\frac{1}{n} \right) + \log \left(\frac{\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}} \right) = \log \left(\frac{1}{n} \right) + q(\alpha),$$

então segue que:

$$\lim_{\alpha \rightarrow \infty} \frac{p(\alpha)}{\alpha} = \lim_{\alpha \rightarrow \infty} \frac{\log \left(\frac{1}{n} \right) + q(\alpha)}{\alpha} = 0 + v = v > 0 \Leftrightarrow \lim_{\alpha \rightarrow \infty} \frac{\alpha}{p(\alpha)} = \frac{1}{v} < \infty.$$

Assim, como $\frac{\alpha}{p(\alpha)}$ é contínua em $[1, \infty)$, pelo Lema 4, existe $L_{10} > 0$ tal que

$$\frac{\alpha}{p(\alpha)} < L_{10} \Leftrightarrow \frac{1}{p(\alpha)} < L_{10} \frac{1}{\alpha}, \forall \alpha \in [1, \infty).$$

Também, de (A.12), temos que:

$$\lim_{\alpha \rightarrow \infty} p(\alpha) = \lim_{\alpha \rightarrow \infty} \left(\log \left(\frac{1}{n} \right) + q(\alpha) \right) = \infty. \quad (\text{B.15})$$

É um limite conhecido, que pode ser provado diretamente da definição através da regra de L'hôpital, que:

$$\lim_{x \rightarrow \infty} \frac{\Gamma(s, x)}{x^{s-1}e^{-x}} = 1.$$

De onde:

$$\lim_{x \rightarrow \infty} \frac{\Gamma(s, x)}{x^{s+1-n}} = \lim_{x \rightarrow \infty} \frac{x^{n-2}}{e^x} \frac{\Gamma(s, x)}{x^{s-1}e^{-x}} = 0 \times 1 = 0.$$

Assim, em especial, como $\lim_{\alpha \rightarrow \infty} p(\alpha) = \infty$, temos

$$\lim_{\alpha \rightarrow \infty} \frac{\Gamma\left(\frac{n-1}{2}, np(\alpha)\right)}{(np(\alpha))^{\frac{n-1}{2}+1-n}} = 0 < \infty.$$

Deste modo, como $\frac{\Gamma\left(\frac{n-1}{2}, np(\alpha)\right)}{(np(\alpha))^{\frac{n-1}{2}+1-n}}$ é contínuo em $[1, \infty)$ então pelo Lema 4, temos que existe L_{11} tal que

$$\frac{\Gamma\left(\frac{n-1}{2}, np(\alpha)\right)}{(np(\alpha))^{\frac{n-1}{2}+1-n}} < L_{11} \Leftrightarrow \frac{\Gamma\left(\frac{n-1}{2}, np(\alpha)\right)}{(np(\alpha))^{\frac{n-1}{2}}} < L_{11} \frac{1}{(np(\alpha))^{n-1}}.$$

Também, por (B.9) temos que:

$$\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}} \leq L_5 \alpha^{n-\frac{5}{2}}, \text{ para } \alpha \geq 1.$$

Assim, temos que:

$$\begin{aligned} \int_1^\infty c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \frac{\Gamma\left(\frac{n-1}{2}, np(\alpha)\right)}{(np(\alpha))^{\frac{n-1}{2}}} d\alpha &< c_2 \int_1^\infty L_5 a^{n-\frac{5}{2}} L_{11} \frac{1}{(np(\alpha))^{n-1}} d\alpha \\ &< c_2 L_5 L_{11} \int_1^\infty \alpha^{n-\frac{5}{2}} \left(\frac{L_{10}}{n\alpha}\right)^{n-1} d\alpha \\ &= \frac{c_2 L_5 L_{11} L_{10}^{n-1}}{n^{n-1}} \int_1^\infty \alpha^{-\frac{3}{2}} d\alpha < \infty, \end{aligned}$$

pois $\int_1^\infty \alpha^{-\frac{3}{2}} d\alpha < \infty$.

Deste modo:

$$\begin{aligned} s_4 &= \int_1^\infty \int_1^\infty \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha \\ &< \int_1^\infty c_2 a^{n-\frac{3}{2}-\frac{\alpha}{1+\alpha}} \frac{\Gamma\left(\frac{n-1}{2}, np(\alpha)\right)}{(np(\alpha))^{\frac{n-1}{2}}} d\alpha < \infty. \end{aligned}$$

Portanto, temos que: $s_1 < \infty, s_2 < \infty, s_3 < \infty$ e $s_4 < \infty$, de modo que:

$$d_4 = \int_0^\infty \int_0^\infty \frac{\alpha^{n-\frac{3}{2}-\frac{\alpha}{(1+\alpha)}}}{\phi \Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\alpha\phi-1} \right\} \frac{\Gamma(n\phi)}{(\sum_{i=1}^n t_i^\alpha)^{n\phi}} d\phi d\alpha = s_1 + s_2 + s_3 + s_4 < \infty$$

como queríamos demonstrar.

APÊNDICE C – Outros métodos para se obter valores iniciais

Neste anexo, dois métodos foram desenvolvidos para se obter bons valores iniciais para serem utilizados nos métodos de estimação da distribuição gama generalizada, considerando-se observações censuradas. Os métodos estão disponíveis em anexo, pois os mesmos não retornaram bons resultados, mostrando-se inadequados para solucionar o problema proposto.

Método 1

Discutimos na seção 4.3.1 que ao utilizarmos observações censuradas, as equações de verossimilhança (3.13, 3.14 e 3.15) não podem ser simplificadas similarmente, como foi feito nas equações de verossimilhança (3.2, 3.3 e 3.4). Desta forma, as equações obtidas considerando-se observações censuradas são muito instáveis e as estimativas obtidas de ϕ , μ e α podem depender dos valores iniciais, utilizados no método iterativo.

Valores preditos para as observações censuradas podem ser aproximados utilizando-se uma distribuição exponencial (5.1). Outra abordagem será obter os valores preditos para as observações censuradas utilizando-se a distribuição Weibull. Uma forma de se obtê-los será descrita através dos seguintes resultados:

Shawky e Bakoban (2009) mostram que se X é uma variável aleatória absolutamente contínua e se podemos escrever $F(x)$ como

$$F(x) = [a\phi(x) + b]^c, \quad (\text{C.1})$$

então

$$E[X|X = x^*] = \frac{1}{c+1} \left[c\phi(x^*) - \frac{a}{b} \right]. \quad (\text{C.2})$$

Se considerarmos $a = -1$, $b = 1$, $c = 1$ e $\exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right)$ tem-se que

$$F(x) = \left[(-1) \exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right) + 1 \right]^1 = \left[1 - \exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right) \right], \quad (\text{C.3})$$

que é f.d.a. da distribuição Weibull. Então podemos encontrar a esperança condicional utilizando-se

$$E[X|X = x^*] = \frac{1}{2} \left[\exp\left(-\left(\frac{x^*}{\beta}\right)^\alpha\right) + 1 \right]. \quad (\text{C.4})$$

Para obter os valores de α e β utilizam-se os estimadores de máxima verossimilhança $\hat{\alpha}$ e $\hat{\beta}$ baseado nos dados completos. Desta forma, utilizando-se a equação (C.4), obtemos valores aproximados para as observações censuradas t^* .

Utilizando-se as observações completas e as preditas, através das equações (3.7, 3.8 e 3.9), encontramos os valores iniciais $\tilde{\phi}$, $\tilde{\mu}$ e $\tilde{\alpha}$. Estes valores obtidos podem ser utilizados para inicialização dos métodos iterativos, fazendo com que, ao consideramos dados censurados, obtenhamos as estimativas para ϕ , μ e α .

Método 2

Seja $T_1, \dots, T_n$ uma amostra aleatória com $T \sim \text{GG}(\alpha, \mu, \phi)$, considerando observações com censura tipo I, a função de verossimilhança é dada por

$$L(\boldsymbol{\theta}; \mathbf{t}) = \frac{\alpha^d \mu^{d\alpha\phi}}{\Gamma(\phi)^n} \left\{ \prod_{i=1}^n t_i^{\delta_i \alpha \phi - 1} \right\} \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\} \prod_{i=1}^n (\Gamma[\phi, (\mu t_i)^\alpha])^{1-\delta_i}. \quad (\text{C.5})$$

A função gama incompleta superior pode ser aproximada (veja por exemplo, Abramowitz and Stegun, 1970) por

$$\Gamma(s, x) = x^s e^{-x} \sum_{k=0}^{\infty} \frac{\Gamma(s) x^{-k}}{\Gamma(s-k)}, |z| \rightarrow \infty. \quad (\text{C.6})$$

Desta forma, obtemos uma aproximação de $\Gamma[\phi, (\mu t_i)^\alpha]$ através de

$$\Gamma[\phi, (\mu t_i)^\alpha] \cong (\mu t_i)^\alpha e^{(\mu t_i)^\alpha} \left[1 + \frac{\Gamma(\phi)}{\Gamma(\phi-1)(\mu t_i)^\alpha} + \frac{\Gamma(\phi)}{\Gamma(\phi-2)(\mu t_i)^{2\alpha}} + \dots \right]. \quad (\text{C.7})$$

Como estamos considerando estimativas preliminares de ϕ , μ e α , iremos utilizar o primeiro elemento da aproximação.

$$\Gamma[\phi, (\mu t_i)^\alpha] \cong (\mu t_i)^\alpha e^{(\mu t_i)^\alpha}. \quad (\text{C.8})$$

Através da aproximação (C.8), a função de verossimilhança (C.5) passa a ser dada por

$$\begin{aligned} L(\boldsymbol{\theta}; \mathbf{t}) &\cong \frac{\alpha^d \mu^{n\alpha(\phi-1)+d\alpha}}{\Gamma(\phi)^n} \exp \left\{ (\alpha\phi-1) \sum_{i=1}^n \delta_i \log t_i + \alpha(\phi-1) \sum_{i=1}^n (1-\delta_i) \log t_i \right\} \times \\ &\times \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\}. \end{aligned} \quad (\text{C.9})$$

A distribuição a posteriori conjunta para ϕ , μ e α , utilizando-se a distribuição a priori obtida pela regra de Jeffreys, é proporcional ao produto da função de verossimilhança

(C.9) e da distribuição a priori (4.9), resultando em

$$p_R(\phi, \mu, \alpha | \mathbf{t}, \boldsymbol{\delta}) \propto \frac{\alpha^{d-1} \mu^{n\alpha(\phi-1)+d\alpha-1}}{\phi \Gamma(\phi)^n} \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\} \times \exp \left\{ (\alpha\phi - 1) \sum_{i=1}^n \delta_i \log t_i + \alpha(\phi - 1) \sum_{i=1}^n (1 - \delta_i) \log t_i \right\}. \quad (\text{C.10})$$

A distribuição a posteriori conjunta para ϕ e α é obtida através de

$$p_R(\phi, \mu, \alpha | \mathbf{t}, \boldsymbol{\delta}) \propto \frac{\alpha^{d-1}}{\phi \Gamma(\phi)^n} \exp \left\{ (\alpha\phi - 1) \sum_{i=1}^n \delta_i \log t_i + \alpha(\phi - 1) \sum_{i=1}^n (1 - \delta_i) \log t_i \right\} \times \int_0^\infty \mu^{n\alpha(\phi-1)+d\alpha-1} \exp \left\{ -\mu^\alpha \sum_{i=1}^n \delta_i t_i^\alpha \right\} d\mu. \quad (\text{C.11})$$

Utilizando-se o resultado (5.3) e após algumas manipulações algébricas obtemos que

$$p_R(\phi, \alpha | \mathbf{t}, \boldsymbol{\delta}) \propto \frac{\alpha^{d-2} \Gamma(n(\phi - 1) + d)}{\phi \Gamma(\phi)^n (\sum_{i=1}^n t_i^\alpha)^{n(\phi-1)+d}} \times \exp \left\{ (\alpha\phi - 1) \sum_{i=1}^n \delta_i \log t_i + \alpha(\phi - 1) \sum_{i=1}^n (1 - \delta_i) \log t_i \right\}, \quad (\text{C.12})$$

em que $d = \sum_{i=1}^n \delta_i$.

O logaritmo da distribuição a posteriori conjunta de α e ϕ é dado por

$$\begin{aligned} \log(p_R(\alpha, \phi | \mathbf{t}, \boldsymbol{\delta})) &= (d - 2) \log(\alpha) + \log(\Gamma(n(\phi - 1) + d)) - n \log(\Gamma(\phi)) \\ &\quad - (n(\phi - 1) + d) \log \left(\sum_{i=1}^n t_i^\alpha \right) + (\alpha\phi - 1) \sum_{i=1}^n \delta_i \log t_i \\ &\quad + \alpha(\phi - 1) \sum_{i=1}^n (1 - \delta_i) \log t_i - \log(\phi). \end{aligned} \quad (\text{C.13})$$

As derivadas parciais de primeira ordem do $\log(p_R(\phi, \alpha | \mathbf{t}, \boldsymbol{\delta}))$, com respeito a α e ϕ , são dadas respectivamente por

$$\begin{aligned} \frac{\partial \log(p_R(\phi, \alpha | \mathbf{t}, \boldsymbol{\delta}))}{\partial \alpha} &= \frac{d - 2}{\alpha} - \frac{(n(\phi - 1) + d) \sum_{i=1}^n t_i^\alpha \log t_i}{\sum_{i=1}^n t_i^\alpha} + \phi \sum_{i=1}^n \delta_i \log t_i \\ &\quad + (\phi - 1) \sum_{i=1}^n (1 - \delta_i) \log t_i \end{aligned} \quad (\text{C.14})$$

$$\begin{aligned} \frac{\partial \log(p_R(\alpha, \phi | \mathbf{t}))}{\partial \phi} &= n\psi(n(\phi - 1) + d) - \frac{1}{\phi} - n\psi(\phi) - n \log \left(\sum_{i=1}^n t_i^\alpha \right) \\ &\quad + \alpha \sum_{i=1}^n \delta_i \log(t_i) + \alpha \sum_{i=1}^n (1 - \delta_i) \log(t_i), \end{aligned} \quad (\text{C.15})$$

Por meio de algumas manipulações algébricas a equação (C.14) pode ser reescrita através de

$$\frac{d-2}{\alpha} - (n(\phi-1) + d)\Lambda + \phi \sum_{i=1}^n \log t_i + \sum_{i=1}^n (1 - \delta_i) \log t_i \quad (\text{C.16})$$

em que $\Lambda = \frac{\sum_{i=1}^n t_i^\alpha \log t_i}{\sum_{i=1}^n t_i^\alpha}$.

As modas $\tilde{\phi}$ e $\tilde{\alpha}$ da distribuição a posteriori conjunta $p_R(\alpha, \phi | \mathbf{t}, \boldsymbol{\delta})$ são obtidas resolvendo-se as seguintes equações:

$$\frac{\partial \log(p_R(\alpha, \phi | \mathbf{t}, \boldsymbol{\delta}))}{\partial \alpha} = 0 \quad \text{e} \quad \frac{\partial \log(p_R(\alpha, \phi | \mathbf{t}, \boldsymbol{\delta}))}{\partial \phi} = 0.$$

Resolvendo $\partial \log(p_M(\alpha, \phi | \mathbf{t}) / \partial \alpha$, obtemos a equação

$$\tilde{\phi} = \frac{\frac{d-2}{\tilde{\alpha}} - \sum_{i=1}^n (1 - \delta_i) \log t_i + (n-d)\Lambda}{n\Lambda - \sum_{i=1}^n \log t_i}. \quad (\text{C.17})$$

Resolvendo $\partial \log(p_M(\alpha, \phi | \mathbf{t}) / \partial \phi$, obtemos $\tilde{\alpha}$ substituindo-se ϕ por $\tilde{\phi}$ e resolvendo-se a seguinte equação em α , dada por

$$\begin{aligned} h(\alpha) = & n\psi(n(\tilde{\phi}-1) + d) - \frac{1}{\tilde{\phi}} - n\psi(\tilde{\phi}) - n \log \left(\sum_{i=1}^n t_i^{\tilde{\alpha}} \right) \\ & + \tilde{\alpha} \sum_{i=1}^n \delta_i \log(t_i) + \tilde{\alpha} \sum_{i=1}^n (1 - \delta_i) \log(t_i). \end{aligned} \quad (\text{C.18})$$

Utilizando-se métodos iterativos como Newton-Raphson obtemos os valores de $\tilde{\alpha}$. Substituindo-se $\tilde{\alpha}$ na equação (5.8), encontra-se $\tilde{\phi}$. O valor inicial de μ pode ser obtido resolvendo-se a equação

$$\frac{\partial \log(p_R(\phi, \mu, \alpha | \mathbf{t}))}{\partial \mu} = \frac{(n\tilde{\alpha}(\tilde{\phi}-1) + d\tilde{\alpha})}{\tilde{\mu}} - \tilde{\alpha}\tilde{\mu}^{\tilde{\alpha}-1} \sum_{i=1}^n t_i^{\tilde{\alpha}} = 0$$

e após algumas manipulação algébricas,

$$\tilde{\mu} = \left(\frac{n(\tilde{\phi}-1) + d}{\sum_{i=1}^n t_i^{\tilde{\alpha}}} \right)^{\frac{1}{\tilde{\alpha}}}. \quad (\text{C.19})$$

No entanto o método não retornou boas estimativas.

APÊNDICE D – Programas Utilizados

Análise Clássica Considerando Dados Completos

Algoritmo desenvolvido no R para calcular as estimativas das médias e desvios-padrão dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ .

```
rm(list=ls(all=TRUE))
library(maxLik)      #biblioteca para o comando maxLik
require(rootSolve)   #biblioteca para o comando multiroot
require(VGAM)        #biblioteca para gerar valores da gama generalizada
require(micEcon)

##### Declarando as variaveis e alguns parametros #####

pfi <- 2              # Parametro fi para gerar valores na GG
pmi <- 1              # Parametro mi para gerar valores na GG
palfa<- 0.5           # Parametro alfa para gerar valores na GG
n <- 100              # Tamanho da amostra
B <- 1000             # Qtd de estimativas que serão calculados

set.seed(2013)
o<-0                  # Contador de iterações
ite<-0                # Contador de iterações

##### tamanho das matrizes que armazenarao estimativas#####

emv <- matrix(nrow=B,ncol=3)
raizes <- matrix(nrow=B,ncol=3)
pc1<-rep(0,times=B)
pc2<-rep(0,times=B)
pc3<-rep(0,times=B)

##### Função para encontrar alfa #####
fdigamma<-function(theta2){
  F<-NULL
  F<-log(theta2)- (1/(2*theta2))- (1/(12*(theta2^2))) + (1/(120*(theta2^4))) - (1/(252*(theta2^6)))
  return(F)
}

sistema<-function (alfa)
{
  F<-NULL
  fi<-NULL
  u<-NULL
  u<- (1/alfa)*(n)/(sum((t^alfa)*log(t))-(sum(t^alfa)*sum((log(t)))/n))
  fi<-(1/alfa)*((sum(t^alfa))/(sum((t^alfa)*log(t))-(sum(t^alfa)*sum((log(t)))/n)))
  F<-digamma(fi)-log(u) -(alfa*sum((log(t)))/n)
  return(F)
}
```

```

F<-function(x)
{
  theta<-x[1]
  mi<-x[2]
  alfa<-x[3]
  F<-matrix(nrow=3)
  F[1]<- -n*digamma(theta)+n*alfa*log(mi)+alfa*sum(log(t))
  F[2]<- (mi^alfa)*sum(t^alfa) - (n*theta)
  F[3]<- (mi^alfa)*sum((t^alfa)*log(mi*t)) -(n/alfa)-(n*theta*log(mi))-theta*sum(log(t))
  return(F)
}

fisher<-function(fi,mi,alfa) {
F<-matrix(nrow=3,ncol=3)
F[1,1]<- trigamma(fi)
F[1,2]<- alfa/mi
F[1,3]<- -digamma(fi)/alfa
F[2,1]<- alfa/mi
F[2,2]<- fi*(alfa^2)/(mi^2)
F[2,3]<- -(1+fi*digamma(fi))/mi
F[3,1]<- -digamma(fi)/alfa
F[3,2]<- -(1+fi*digamma(fi))/mi
F[3,3]<- (1+2*digamma(fi)+fi*trigamma(fi)+(fi*(digamma(fi)^2)))/(alfa^2)
return(F)
}

##### Log da Verossimilhança#####
loglike <- function(theta)
{
  fi <- theta[1]
  mi <- theta[2]
  alfa <- theta[3]
  c<-(mi^alfa)*sum((t^alfa))
  l<- n*log(alfa)-n*log(gamma(fi))+(n*alfa*fi*log(mi))+(alfa*fi-1)*sum(log(t))- c
  return(l)
}

##### Calculando #####
### Comando try() faz com que mesmo que o programa encontre erro ele não pare de executar ###
while(o<B)
{
  t<-rgengamma(n, scale = (1/pmi), d = palfa, k = pfi) ##Gera valores de uma dist. gama generalizada
  aux<-multirroot(f=sistema, start=c(2)) #resolve a equação (12) do artigo
  if(is.real(aux$root)& aux$root>0.1 & aux$root<30 & aux$root!="NaN")
  {
    alfa<-aux$root
    fi<-(1/alfa)*((sum(t^alfa))/(sum((t^alfa)*log(t))-(sum(t^alfa)*sum((log(t)))/n)))
    u<- (1/alfa)*(n)/(sum((t^alfa)*log(t))-(sum(t^alfa)*sum((log(t)))/n))
    mi<-u^(1/alfa)
    out=c(fi,mi,alfa)
    FM<- try(fisher(out[1],out[2],out[3]))
    IVFM<-try(solve(n*FM))
    varfi<-try(IVFM[1,1])
    varmi<-try(IVFM[2,2])
    varalfa<-try(IVFM[3,3])
    if(mi<30 & fi<30 & varfi>0 & varmi>0 & varalfa>0){
      o<-o+1
    }
  }
}

```

```

#Calcula a Probabilidade de Cobertura
c1<-qnorm(0.975,0,1)
iifi<-out[1]-c1*(varfi^0.5) ; isfi<-out[1]+c1*(varfi^0.5) ;
iimi<-out[2]-c1*(varmi^0.5) ; ismi<-out[2]+c1*(varmi^0.5) ;
ial<-out[3]-c1*(varalfa^0.5) ; isal<-out[3]+c1*(varalfa^0.5) ;
if(iifi<=pfi & isfi>= pfi) pc1[o]<-1
if(iimi<=pmi & ismi>= pmi) pc2[o]<-1
if(ial<=palfa & isal>= palfa) pc3[o]<-1
emv[o,]<- out          #salva os valores iniciais
raizes[o,]<-F(out)
cat(o,"      ",ite+1, "\n")
} }
ite<-ite+1
}
raizes1<-rep(0,B)
for (i in 1:B)
for (j in 1:3)
raizes1[i]<-raizes[i,j]+raizes1[i]
##### Imprimi #####
mediaemv<-c(mean(emv[,1]),sd(emv[,1]),mean(emv[,2]),sd(emv[,2]),mean(emv[,3]),sd(emv[,3]))
round(mediaemv,3)

probcob<-c(sum(pc1)/B,sum(pc2)/B,sum(pc3)/B)
round(probcob,4)*100

```

Análise Clássica Considerando Dados Censurados

Algoritmo desenvolvido no R para calcular as estimativas das médias e desvios-padrão dos valores iniciais $\theta^{(0)} = (\hat{\phi}, \hat{\mu}, \hat{\alpha})$ e dos EMV's encontrados para 1000 amostras de tamanho $n = (50, 100, 200)$, com diferentes valores de θ e observações censuradas.

```

library(maxLik)      #biblioteca para o comando maxLik
require(rootSolve)   #biblioteca para o comando multiroot
require(VGAM)        #biblioteca para gerar valores da gama generalizada
require(micEcon)
require(MASS)

##### Declarando as variaveis e alguns parametros #####
old = options(digits=10)
pfi <- 4             # Parametro fi para gerar valores na GG
pmi <- 2             # Parametro mi para gerar valores na GG
palfa<- 2           # Parametro alfa para gerar valores na GG
n <- 200             # Tamanho da amostra
B <- 1000            # Qtd de estimativas que serão calculados
cen<-0.2

set.seed(2013)
o<-0                # Contador de iterações
ite<-0              # Contador de iterações
pc1<-rep(0,times=B)
pc2<-rep(0,times=B)
pc3<-rep(0,times=B)

##### tamanho das matrizes que armazenarao estimativas#####

```

```

emv <- matrix(nrow=B,ncol=3)
valores<-matrix(nrow=B,ncol=3)

##### Função para encontrar alfa #####
fdigamma<-function(theta2){
F<-NULL
F<-log(theta2)- (1/(2*theta2))- (1/(12*(theta2^2))) + (1/(120*(theta2^4))) - (1/(252*(theta2^6)))
return(F)
}

sistema<-function (alfa)
{
F<-NULL
fi<-NULL
u<-NULL
u<- (1/alfa)*(n1)/(sum((w^alfa)*log(w))-(sum(w^alfa)*sum((log(w)))/n1))
fi<-(1/alfa)*((sum(w^alfa))/(sum((w^alfa)*log(w))-(sum(w^alfa)*sum((log(w)))/n1)))
F<-digamma(fi)-log(u) -(alfa*sum((log(w)))/n1)
return(F)
}

##### Log da Verossimilhança#####
loglike <- function(theta)
{
fi <- theta[1]
mi <- theta[2]
alfa <- theta[3]
l<- d*log(alfa)+(d*alfa*fi)*log(mi)-n*log(gamma(fi))+(alfa*fi-1)*sum(delta*log(t))
-(mi^alfa)*sum(delta*(t^alfa))+sum((1-delta)*log(pgamma((mi*t)^alfa,fi,lower=FALSE)*gamma(fi)))
return(l)
}

##### Calculando #####
### Comando try() faz com que mesmo que o programa encontre erro ele não pare de executar ###
while(o<B)
{
t<-rgengamma(n, scale = (1/pmi), d = palfa, k = pfi)
delta<-rep(1,n)
#Adiciona censura tipo I
y<-sample(1:n,n*cen)
for (j in 1:length(y)){
t[y[j]]<-runif(1,0, t[y[j]])
delta[y[j]]<-0
}
t1<-t
# Faz a previsão utilizando weibull para os dados censurados
w<-NULL
c1<-0
for (i in 1:n){
if( delta[i]==1) {
c1<-c1+1
w[c1]<-t[i]
} }
n1<-length(w)
d<-sum(delta)
aux<-multiroot(f=sistema, start=c(3)) #resolve a equação (12) do artigo
if(is.real(aux$root)& aux$root>0.1 & aux$root<20 & aux$root!="NaN")

```

```

{
  alfa<-aux$root
  u<- (1/alfa)*(n1)/(sum((w^alfa)*log(w))-(sum(w^alfa)*sum((log(w)))/n1))
  fi<-(1/alfa)*((sum(w^alfa))/(sum((w^alfa)*log(w))-(sum(w^alfa)*sum((log(w)))/n1)))
  mi<-u^(1/alfa)
  estimador=c(fi,mi,alfa)
  t<-t1
  est=c(fi,mi,alfa)
  res <- try(maxNR(loglike, start=c(est[1],est[2],est[3]),iterlim = 10))
  out<- try(res$estimate)
  varfi<-try(solve(-res$hessian)[1,1])
  varmi<-try(solve(-res$hessian)[2,2])
  varalfa<-try(solve(-res$hessian)[3,3])
  #verificar se o método conseguiu encontrar o EMV, em alguns casos ele não conseguia e o programa parava.
  if(is.real(out[1]) & varfi>0 & varmi>0 & varalfa>0)
  if(out[1]>0 & out[2]>0 & out[3]>0 & mi<30 & fi<30 & out[1]!="NaN" & out[2]!="NaN" & out[3]!="NaN" )
    o<-o+1
    #Calcula a Probabilidade de Cobertura
    c1<-qnorm(0.975,0,1)
    iifi<-out[1]-c1*(varfi^0.5) ; isfi<-out[1]+c1*(varfi^0.5) ;
    iimi<-out[2]-c1*(varmi^0.5) ; ismi<-out[2]+c1*(varmi^0.5) ;
    ial<-out[3]-c1*(varalfa^0.5) ; isal<-out[3]+c1*(varalfa^0.5) ;
    if(iifi<=pfi & isfi>= pfi) pc1[o]<-1
    if(iimi<=pmi & ismi>= pmi) pc2[o]<-1
    if(ial<=palfa & isal>= palfa) pc3[o]<-1
    #Fim do calculo da probabilidade de cobertura
    emv[o,] <- out
    valores[o,]<- est          #salva os valores iniciais
    cat(o,"      ",ite+1, "\n")
  } }
  ite<-ite+1
}

##### Imprimi #####
mediavalores<-c(mean(valores[,1]),sd(valores[,1]),mean(valores[,2]),sd(valores[,2]),
mean(valores[,3]),sd(valores[,3]))
round(mediavalores,3)

mediaemv<-c(mean(emv[,1]),sd(emv[,1]),mean(emv[,2]),sd(emv[,2]),mean(emv[,3]),sd(emv[,3]))
round(mediaemv,3)

probcob<-c(sum(pc1)/B,sum(pc2)/B,sum(pc3)/B)
round(probcob,4)*100

```

Análise Bayesiana Considerando Dados Completos e Prioris Empíricas

Algoritmo desenvolvido no R para calcular as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori gama com $\sigma_{\theta}^2 = (1, 1, 1)$, $\sigma_{\theta}^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_{\theta}^2 = (10, 10, 10)$.

```
rm(list=ls(all=TRUE))
```

```

require(R2OpenBUGS) #biblioteca que faz a ligação entre o OPENBUGS e o R
require(rootSolve)  #biblioteca para o comando multiroot
require(VGAM)        #biblioteca para gerar valores da gama generalizada
require(coda)         #Pacote para testar a convergencia dos MCMC

#Para utilizar o R2OpenBUGS é necessario utilizar uma versão atualizada
#do R (2.15.3) e do OPENBUGS

B<- 1000      ### Numero de Estimativas

### Que serão utilizar no OPENBUGS
cadeia<-600
thining<-15
Burnin<-100
indice<-cadeia-Burnin

#Declarando as variaveis
pfi <- 0.4      # Parametro fi para gerar valores na GG
pmi <- 1.5      # Parametro mi para gerar valores na GG
palfa<- 5       # Parametro alfa para gerar valores na GG
n <- 50         # Tamanho da amostra
vp<-10         # Variancia das prioris
#Declarando vetores e contadores
emcmc <- matrix(nrow=B,ncol=3)
valores<-matrix(nrow=B,ncol=3)
old = options(digits=10)
o<-1
ite<-1
pc1<-rep(0,times=B)
pc2<-rep(0,times=B)
pc3<-rep(0,times=B)

mediaprior<- function(media,variancia) {
  auxbeta<-media/variancia
  auxalpha<-media*auxbeta
  return(auxalpha)
}

varprior<- function(media,variancia) {
  auxbeta<-media/variancia
  return(auxbeta)
}

fdigamma<-function(theta2){
  F<-NULL
  F<-log(theta2)- (1/(2*theta2))- (1/(12*(theta2^2))) + (1/(120*(theta2^4))) - (1/(252*(theta2^6)))
  return(F)
}

sistema<-function (alfa)
{
  F<-NULL
  fi<-NULL
  fi<-((n-(3/2)+(alfa/(1+alfa)))*(1/alfa) -(1/((1+alfa)^2))*log(alfa))/
  (((n*sum((t^alfa)*log(t))/sum(t^alfa))-sum(log(t))))
  F<- n*fdigamma(n*fi)+alfa*sum(log(t))-(1/fi)-n*fdigamma(fi)-n*log(sum(t^alfa))
  return(F)
}

```



```

}

set.seed(2013)

## Declando as distribuições a priori e posteriori no OPENBUGS ##
modelo<-function () {

for (i in 1:N) {

x[i] ~ dggamma(phi,mu,alpha)
}
phi ~ dgamma(phim,phiv)
mu ~ dgamma(mim,miv)
alpha~ dgamma(alpham,alphav)

}

while(o<=B)
{
#####-----#####

x<-rgengamma(n, scale = (1/pmi), d = palfa, k = pfi)
t<-x
aux<-multroot(f=sistema, start=c(4))
if(is.real(aux$root) & aux$root>0.1 & aux$root<30 & aux$root!="NaN")
{
alfa<-aux$root
fi<-((n-(3/2)+(alfa/(1+alfa)))*(1/alfa) -(1/((1+alfa)^2))*log(alfa))/
((n*sum((t^alfa)*log(t))/sum(t^alfa))-sum(log(t)))
mi<-((n*fi*alfa-1)/(alfa*sum(t^alfa)))^(1/alfa)
estimador=c(fi,mi,alfa)
phim<-mediaprior(fi,vp) ; phiv<-varprior(fi,vp)
mim<-mediaprior(mi,vp) ; miv<-varprior(mi,vp)
alpham<-mediaprior(alfa,vp) ; alphav<-varprior(alfa,vp)
#Declarando valores iniciais do OpenBUGS
inits <- function(){
list(phi = fi, mu = mi, alpha = alfa)
}
#####-----#####

#Declarando o vetor que conterá a amostra e as médias das distribuições a priori
N<-n
data<-list("N", "x","phim","phiv","mim","miv","alpham","alphav")
#####-----#####

resultados<-bugs(data, inits, model.file = modelo,parameters = c("phi", "mu", "alpha"),
debug=FALSE ,n.burnin=Burnin,n.thin=thining,n.chains = 1, n.iter = cadeia,codaPkg = T)
codaobject <- read.bugs(resultados)

vphi<-NULL
vmi<-NULL
valfa<-NULL

for (i in 1:indice) {
vphi[i]<- codaobject[i,4][[1]]
vmi[i]<- codaobject[i,3][[1]]
valfa[i]<- codaobject[i,1][[1]]
}
}

```

```

}

prpi<-quantile(vphi, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
prps<-quantile(vphi, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)
prmi<-quantile(vmi, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
prms<-quantile(vmi, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)
prai<-quantile(valfa, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
pras<-quantile(valfa, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)

ge1<-abs(geweke.diag(vphi)$z[1])
ge2<-abs(geweke.diag(vmi)$z[1])
ge3<-abs(geweke.diag(valfa)$z[1])

if (ge1<1.96 & ge2<1.96 & ge3<1.96 & fi<30 & mi<30) { emcmc[o,1] <- mean(vphi);
  emcmc[o,2]<- mean(vmi); emcmc[o,3]<- mean(valfa); valores[o,]<- estimador;
  {
    if(prpi<=pfi & prps>= pfi) pc1[o]<-1
    if(prmi<=pmi & prms>= pmi) pc2[o]<-1
    if(prai<=palfa & pras>= palfa) pc3[o]<-1 };
    cat(o,"      ",ite,"      ",round(sum(pc1)/o,3),"      ",round(sum(pc2)/o,3),
"      ",round(sum(pc3)/o,3),"\\n"); o<-(o+1);
  } }
  ite<-ite+1
}

mediavalores<-c(mean(valores[,1]),sd(valores[,1]),mean(valores[,2]),
sd(valores[,2]),mean(valores[,3]),sd(valores[,3]))
round(mediavalores,3)

estimativas<-c(mean(emcmc[,1]),sd(emcmc[,1]),mean(emcmc[,2]),
sd(emcmc[,2]),mean(emcmc[,3]),sd(emcmc[,3]))
round(estimativas,3)

probcob<-c(sum(pc1)/B,sum(pc2)/B,sum(pc3)/B)
round(probcob,4)*100

```

Para se utilizar outras distribuições a priori, devemos apenas trocar o modelo utilizado no OPENBUGS, dados a seguir

```

modelo<-function () {
for (i in 1:N) {
x[i] ~ dgamma(phi,mu,alpha)
}
phi ~ dgamma(fi,1)
mu ~ dgamma(mi,1)
alpha~ dgamma(alfa,1)
}

modelo<-function () {
for (i in 1:N) {
x[i] ~ dgamma(phi,mu,alpha)
}
phi ~ dunif(a1,b1)
mu ~ dunif(a2,b2)
alpha~ dunif(a3,b3)
}

```

```
}
```

Análise Bayesiana Considerando Dados Completos e Priors Não-Informativas

Algoritmo desenvolvido no R para calcular as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se as densidades a posteriori obtidas por meio da prioriobtida pela Regra de Jeffreys Modificada.

```
require(VGAM)          #biblioteca para gerar valores da gama generalizada
require(rootSolve)     #biblioteca para o comando multroot
require(msm)
require(coda)

set.seed(2013)
B<- 1000
#Valores que armazenarão informações sobre as iterações
o<-1
ite<-1
emcmc <- matrix(nrow=B,ncol=3)
valores<-matrix(nrow=B,ncol=3)
# supply the number of iterations
R<-8500
n<-50
pfi <- 0.5             # Parametro fi para gerar valores na GG
pmi <- 0.5             # Parametro mi para gerar valores na GG
palfa<- 3              # Parametro alfa para gerar valores na GG
old = options(digits=20)
pc1<-rep(0,times=B)
pc2<-rep(0,times=B)
pc3<-rep(0,times=B)

#Condicional de alfa | phi , t

posterior1 <- function (a,fi,mi) {
  likeh<- (n-0.5-(a/(1+a)))*log(a)+(n*a*fi-1)*log(mi)+(a*fi-1)*sum(log(X))-(mi^a)*sum(X^a)
  posterior= likeh
  return(posterior) }

#Condicional de phi | alfa , t

posterior3 <- function (a,fi,mi) {
  posterior<- (n*a*fi-1)*log(mi)+(a*fi-1)*sum(log(X))-log(fi)-n*lgamma(fi)
  return(posterior) }

#Aproximação util para função digamma

fdigamma<-function(theta2){
  F<-NULL
  F<-log(theta2)- (1/(2*theta2))- (1/(12*(theta2^2))) + (1/(120*(theta2^4))) - (1/(252*(theta2^6)))
  return(F)
}
```

```

# Sistema para Encontrar o valor de alfa

sistema<-function (alfa)
{
  F<-NULL
  fi<-NULL
  fi<-((n-(3/2)+(alfa/(1+alfa)))*(1/alfa) -(1/((1+alfa)^2))*log(alfa))/
  (((n*sum((t^alfa)*log(t))/sum(t^alfa))-sum(log(t))))
  F<- n*fdigamma(n*fi)+alfa*sum(log(t))-(1/fi)-n*fdigamma(fi)-n*log(sum(t^alfa))
  return(F)
}

#Laço de repetição

while(o<=B)
{
  # dimension of parameters
  alpha<-length(R+1)
  mi<-length(R+1)
  k<-length(R+1)

  #Gerando valores da Gamma Generalizada
  X<-rgengamma(n, scale = (1/pmi), d = palfa, k = pfi)

  #Encontra os Valores iniciais que serão utilizados nas distribuições auxiliares
  t<-X
  aux<-multiroot(f=sistema, start=c(4))
  if(is.real(aux$root)& aux$root>0.1 & aux$root<30 & aux$root!="NaN")
  {
    inialpha<-aux$root
    inifi<-((n-(3/2)+(inialpha/(1+inialpha)))*(1/inialpha) -(1/((1+inialpha)^2))*log(inialpha))
    /(((n*sum((t^inialpha)*log(t))/sum(t^inialpha))-sum(log(t))))
    inimi<-((n*inifi*inialpha-1)/(inialpha*sum(t^inialpha)))^(1/inialpha)

    estimador=c(inifi,inimi,inialpha)

    sdfi<-0.5
    sdmi<-0.5
    sdalpha<-1
    #Valores Iniciais
    alpha[1]<-inialpha
    mi[1]<-inimi
    k[1]<-inifi

    #Contador de valores aceitos

    c1<-rep(0,times=R)
    c2<-rep(0,times=R)
    c3<-rep(0,times=R)

    ## Realizando o M-H hibrido
    a1<-0 ; a2<-0 ; a3<-0 ; i<-1

    while (i<=R) {
      if(i<1) i<-1
      prop1<-rtnorm(1,inialpha,sdalpha,lower=0, upper=Inf)
      ratio1<-posterior1(prop1,k[i],mi[i])-posterior1(alpha[i],k[i],mi[i])+log(dtnorm(alpha[i],

```

```

mean=inialpha, sd=sdalpha, lower=0, upper=Inf, log = FALSE))-log(dtnorm(prop1,
mean=inialpha, sd=sdalpha, lower=0, upper=Inf, log = FALSE))
alpha1<-min(1,exp(ratio1))
u1<-runif(1)

if (u1<alpha1 & alpha1!="NaN" & alpha1!="Inf" & alpha1!="-Inf") {alpha[i+1]<-prop1 ;
c1[i]<-0 ; a1<-0} else { alpha[i+1]<-alpha[i] ;
a1<-a1+1; c1[i]<-1 }

prop3<-rtnorm(1,inifi,sdfi,lower=0, upper=Inf)
ratio3<-posterior3(alpha[i+1],prop3,mi[i])-posterior3(alpha[i+1],k[i],mi[i])+
log(dtnorm(k[i], mean=inifi, sd=sdfi, lower=0, upper=Inf, log = FALSE))-log(
dtnorm(prop3, mean=inifi, sd=sdfi, lower=0, upper=Inf, log = FALSE))
alpha3<-min(1,exp(ratio3))
u3<-runif(1)

if (u3<alpha3 & alpha3!="NaN" & alpha3!="Inf" & alpha3!="-Inf") {k[i+1]<-prop3 ;
c3[i]<-0 ; a2<-0} else { k[i+1]<-k[i] ;
a2<-a2+1; c3[i]<-1 }

ppmi<-(sum(t^alpha[i+1])^(1/alpha[i+1]))
ppfi<-n*k[i+1]
prop2<-rtnorm(1,inimi,sdmi,lower=0, upper=Inf)
ratio2<-dgengamma(prop2, scale = (1/ppmi), d = alpha[i+1], k = ppfi, log =TRUE)-
dgengamma(mi[i], scale = (1/ppmi), d = alpha[i+1], k = ppfi, log =TRUE)
+dtorm(mi[i], mean=inimi, sd=sdmi, lower=0, upper=Inf, log = TRUE)-
dtorm(prop2, mean=inimi, sd=sdmi, lower=0, upper=Inf, log = TRUE)
alpha2<-min(1,exp(ratio2))
u2<-runif(1)

if (u2<alpha2 & alpha2!="NaN" & alpha2!="Inf" & alpha2!="-Inf") {mi[i+1]<-prop2 ;
c2[i]<-0 ; a3<-0} else { mi[i+1]<-mi[i] ;
a3<-a3+1 ; c2[i]<-1 }

if(a1==40 | a2==40 | a3==40 ) {i<-i-50; a1=0 ; a2=0 ;a3=0}
i<-i+1
}
nac<-length(c3)
vphi<-k[seq(1000,R,15)]
vmi<- mi[seq(1000,R,15)]
valfa<-alpha[seq(1000,R,15)]
old = options(digits=5)
ace1<-(1-sum(c3)/nac)
ace2<- (1-sum(c2)/nac)
ace3<- (1-sum(c1)/nac)
atc1<- mean(acf(vphi,plot=F)$acf)
atc2<- mean(acf(vmi,plot=F)$acf)
atc3<- mean(acf(valfa,plot=F)$acf)
ge1<-abs(geweke.diag(vphi)$z[1])
ge2<-abs(geweke.diag(vmi)$z[1])
ge3<-abs(geweke.diag(valfa)$z[1])
prpi<-quantile(vphi, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
prps<-quantile(vphi, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)
prmi<-quantile(vmi, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
prms<-quantile(vmi, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)
prai<-quantile(valfa, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
pras<-quantile(valfa, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)

```

```

if (ace1>=0.10 & ace1<=0.5 & ace2>=0.10 & ace2<=0.5 & ace3>=0.10 & ace3<=0.5 &
inifi<30 & inimi<30 & ge1<1.96 & ge2<1.96 & ge3<1.96 ) {
emcmc[o,1] <- median(vphi); emcmc[o,2]<- median(vmi); emcmc[o,3]<- median(valfa);
valores[o,]<- estimador;
{
  if(prpi<=pfi & prps>= pfi) pc1[o]<-1
  if(prmi<=pmi & prms>= pmi) pc2[o]<-1
  if(prai<=palfa & pras>= palfa) pc3[o]<-1
};
cat(o,"      ",ite,"      ",round(sum(pc1)/o,3),"      ",round(sum(pc2)/o,3),"      ",
round(sum(pc3)/o,3),"\\n");o<-(o+1);
}
cat(o-1,"      ",ite,"      ",round(sum(pc1)/(o-1),3),"      ",round(sum(pc2)/(o-1),3),
"      ",round(sum(pc3)/(o-1),3),"      ",round(ace1,3),"      ",round(ace2,3),"      ",
round(ace3,3),"\\n");
} ite<-ite+1
}
mediavalores<-c(mean(valores[,1]),sd(valores[,1]),mean(valores[,2]),sd(valores[,2]),
mean(valores[,3]),sd(valores[,3]))
round(mediavalores,3)

estimativas<-c(mean(emcmc[,1]),sd(emcmc[,1]),mean(emcmc[,2]),sd(emcmc[,2]),
mean(emcmc[,3]),sd(emcmc[,3]))
round(estimativas,3)

probcob<-c(sum(pc1)/B,sum(pc2)/B,sum(pc3)/B)
round(probcob,4)*100

```

Análise Bayesiana Considerando Dados Censurados e Prioris Empíricas

Algoritmo desenvolvido no R para calcular as médias e os desvios-padrão dos valores iniciais e das estimativas das médias a posteriori encontradas para 1000 amostras simuladas de tamanho $n = (50, 100, 200)$, com diferentes valores de θ , utilizando-se a distribuição a priori gama com $\sigma_\theta^2 = (1, 1, 1)$, $\sigma_\theta^2 = (\tilde{\phi}, \tilde{\mu}, \tilde{\alpha})$ e $\sigma_\theta^2 = (10, 10, 10)$ e considerando dados censurados.

```

rm(list=ls(all=TRUE))
require(R2OpenBUGS) #biblioteca que faz a ligação entre o OPENBUGS e o R
require(rootSolve)  #biblioteca para o comando multiroot
require(VGAM)        #biblioteca para gerar valores da gama generalizada
require(coda)        #Pacote para testar a convergencia dos MCMC
require(MASS)

#Para utilizar o R2OpenBUGS é necessario utilizar uma versão atualizada
#do R (2.15.3) e do OPENBUGS

B<- 1000    ### Numero de Estimativas
### Que serão utilizar no OPENBUGS
cadeia<-600
thining<-15
Burnin<-100

```

```

indice<-cadeia-Burnin
#Declarando as variaveis
pfi <- 0.5          # Parametro fi para gerar valores na GG
pmi <- 0.5          # Parametro mi para gerar valores na GG
palfa<- 3           # Parametro alfa para gerar valores na GG
n <- 200            # Tamanho da amostra
cen<-0.2
vp<-10             # Variancia das prioris
#Declarando vetores e contadores
emcmc <- matrix(nrow=B,ncol=3)
valores<-matrix(nrow=B,ncol=3)
old = options(digits=10)
o<-1
ite<-1
pc1<-rep(0,times=B)
pc2<-rep(0,times=B)
pc3<-rep(0,times=B)

mediaprior<- function(media,variancia) {
auxbeta<-media/variancia
auxalpha<-media*auxbeta
return(auxalpha)
}

varprior<- function(media,variancia) {
auxbeta<-media/variancia
return(auxbeta)
}

fdigamma<-function(theta2){
F<-NULL
F<-log(theta2)- (1/(2*theta2))- (1/(12*(theta2^2))) + (1/(120*(theta2^4))) -
(1/(252*(theta2^6)))
return(F) }

sistema<-function (alfa)
{
F<-NULL
fi<-NULL
fi<-((n1-(3/2)+(alfa/(1+alfa)))*(1/alfa) -(1/((1+alfa)^2))*log(alfa))/
(((n1*sum((w^alfa)*log(w))/sum(w^alfa))-sum(log(w))))
F<- n1*fdigamma(n1*fi)+alfa*sum(log(w))-(1/fi)-n1*fdigamma(fi)-n1*log(sum(w^alfa))
return(F) }
set.seed(2013)

model.file <- file.path(.path.package("R2OpenBUGS"), "model", "modelo.txt")
## Declando as distribuições a priori e posteriori no OPENBUGS ##

while(o<=B)
{
#####
t<-rgengamma(n, scale = (1/pmi), d = palfa, k = pfi)
w<-t
delta<-rep(0,n)
y<-sample(1:n,n*cen)
for (j in 1:length(y)){
delta[y[j]]<-runif(1,0,t[y[j]])
}
}

```

```

w[y[j]]<-delta[y[j]]
t[y[j]]<- NA
}
w<-NULL
c1<-0
for (i in 1:n){
  if(!is.na(t[i])) {
    c1<-c1+1
    w[c1]<-t[i]
  } }
n1<-length(w)

x<-t
aux<-multiroot(f=sistema, start=c(4))
if(is.real(aux$root) & aux$root>0.1 & aux$root<30 & aux$root!="NaN")
{
  alfa<-aux$root
  fi<-((n1-(3/2)+(alfa/(1+alfa)))*(1/alfa) -(1/((1+alfa)^2))*log(alfa))/
  ((n1*sum((w^alfa)*log(w))/sum(w^alfa))-sum(log(w)))
  mi<-((n1*fi*alfa-1)/(alfa*sum(w^alfa)))^(1/alfa)
  estimador=c(fi,mi,alfa)
  phim<-mediaprior(fi,vp) ; phiv<-varprior(fi,vp)
  mim<-mediaprior(mi,vp) ; miv<-varprior(mi,vp)
  alphas<-mediaprior(alfa,vp) ; alphav<-varprior(alfa,vp)
  #Declarando valores iniciais do OpenBUGS
  inits <- function(){
    list(phi = fi, mu = mi, alpha = alfa)
  }
  #####-----#####
  #Declarando o vetor que conterá a amostra e as médias das distribuições a priori
  N<-n
  data<-list("N", "x","delta","phim","phiv","mim","miv","alpham","alphav")
  #####-----#####
  try(resultados<-bugs(data, inits, model.file,parameters = c("phi", "mu", "alpha"),
  debug=FALSE ,n.burnin=Burnin,n.thin=thining,n.chains = 1, n.iter = cadeia,
  codaPkg = T))
  try(codaobject <- read.bugs(resultados))
  if(is.real(codaobject[1,4][[1]]) & length(codaobject[,4][[1]])==500) {
    vphi<-NULL
    vmi<-NULL
    valfa<-NULL

    for (i in 1:indice) {
      vphi[i]<- codaobject[i,4][[1]]
      vmi[i]<- codaobject[i,3][[1]]
      valfa[i]<- codaobject[i,1][[1]]
    }
    if(var(vphi)!=0 & var(vmi)!=0 & var(valfa)!=0) {

prpi<-quantile(vphi, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
prps<-quantile(vphi, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)
prmi<-quantile(vmi, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
prms<-quantile(vmi, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)
prai<-quantile(valfa, probs = 0.025, na.rm = FALSE,names = FALSE,type = 7)
pras<-quantile(valfa, probs = 0.975, na.rm = FALSE,names = FALSE,type = 7)
ge1<-abs(geweke.diag(vphi)$z[1])
ge2<-abs(geweke.diag(vmi)$z[1])

```



```

ge3<-abs(geweke.diag(valfa)$z[1])

if (ge1<1.96 & ge2<1.96 & ge3<1.96 & fi<30 & mi<30) { emcmc[o,1] <- mean(vphi);
  emcmc[o,2]<- mean(vmi); emcmc[o,3]<- mean(valfa); valores[o,]<- estimador;
  {
    if(prpi<=pfi & prps>= pfi) pc1[o]<-1
    if(prmi<=pmi & prms>= pmi) pc2[o]<-1
    if(prai<=palfa & pras>= palfa) pc3[o]<-1
  };
  cat(o,"      ",ite,"      ",round(sum(pc1)/o,3),"      ",round(sum(pc2)/o,3),
"      ",round(sum(pc3)/o,3),"\\n"); o<-(o+1);
  } } } }
ite<-ite+1
}
mediavalores<-c(mean(valores[,1]),sd(valores[,1]),mean(valores[,2]),sd(valores[,2]),
mean(valores[,3]),sd(valores[,3]))
round(mediavalores,3)

estimativas<-c(mean(emcmc[,1]),sd(emcmc[,1]),mean(emcmc[,2]),sd(emcmc[,2]),
mean(emcmc[,3]),sd(emcmc[,3]))
round(estimativas,3)

probcob<-c(sum(pc1)/B,sum(pc2)/B,sum(pc3)/B)
round(probcob,4)*100

```

O arquivo contendo as informações do modelo do OPENBUGS deve ser salvo em um arquivo de texto (txt) na pasta 'C:\Meus Documentos\Documentos\R\win-library\2.15\R2OpenBUGS'. Contendo um dos modelos disponíveis a seguir:

```

modelo<-function () {
for (i in 1:N) {
x[i] ~ dggamma(phi,mu,alpha)I(delta[i],)
}
phi ~ dgamma(phim,phiv)
mu ~ dgamma(mim,miv)
alpha~ dgamma(alpham,alphav)
}

modelo<-function () {
for (i in 1:N) {
x[i] ~ dggamma(phi,mu,alpha)I(delta[i],)
}
phi ~ dgamma(fi,1)
mu ~ dgamma(mi,1)
alpha~ dgamma(alfa,1)
}

modelo<-function () {
for (i in 1:N) {
x[i] ~ dggamma(phi,mu,alpha)I(delta[i],)
}
phi ~ dunif(a1,b1)
mu ~ dunif(a2,b2)
}

```

```
alpha~ dunif(a3,b3)
}
```

APÊNDICE E – Bayesian Inference for Two-Parameter Gamma Distribution Assuming Different Noninformative Priors

A distribuição gama é um modelo flexível que possibilita obter bons ajustes a diversos problemas.

Diferentemente da distribuição gama generalizada, inferências podem ser facilmente obtidas para a distribuição gama utilizando as duas abordagens (clássica e Bayesiana).

Com o objetivo de estudar um modelo mais simples obtido da distribuição gama generalizada. Neste artigo foram investigados os efeitos de diferentes distribuições a priori não-informativas nas estimativas a posteriori da distribuição gama. Para isto serão apresentadas inferências baseadas em dados simulados. Uma aplicação com dados de confiabilidade também será apresentada.

O artigo foi publicado pela Revista Colombiana de Estadística com o título "Bayesian Inference for Two-Parameter Gamma Distribution Assuming Different Noninformative Priors" e pode ser acessado através do link:

<http://www.emis.de/journals/RCE/V36/v36n2a09MoalaEtAl2.pdf>

Bayesian Inference for Two-Parameter Gamma Distribution Assuming Different Noninformative Priors

Inferencia Bayesiana para la distribución Gamma de dos parámetros
asumiendo diferentes a prioris no informativas

FERNANDO ANTONIO MOALA^{1,a}, PEDRO LUIZ RAMOS^{1,b},
JORGE ALBERTO ACHCAR^{2,c}

¹DEPARTAMENTO DE ESTATÍSTICA, FACULTAD DE CIENCIA Y TECNOLOGÍA, UNIVERSIDADE
ESTADUAL PAULISTA, PRESIDENTE PRUDENTE, BRASIL

²DEPARTAMENTO DE MEDICINA SOCIAL, FACULTAD DE MEDICINA DE RIBEIRÃO PRETO,
UNIVERSIDADE DE SÃO PAULO, RIBEIRÃO PRETO, BRASIL

Abstract

In this paper distinct prior distributions are derived in a Bayesian inference of the two-parameters Gamma distribution. Noninformative priors, such as Jeffreys, reference, MDIP, Tibshirani and an innovative prior based on the copula approach are investigated. We show that the maximal data information prior provides in an improper posterior density and that the different choices of the parameter of interest lead to different reference priors in this case. Based on the simulated data sets, the Bayesian estimates and credible intervals for the unknown parameters are computed and the performance of the prior distributions are evaluated. The Bayesian analysis is conducted using the Markov Chain Monte Carlo (MCMC) methods to generate samples from the posterior distributions under the above priors.

Key words: Gamma distribution, noninformative prior, copula, conjugate, Jeffreys prior, reference, MDIP, orthogonal, MCMC.

Resumen

En este artículo diferentes distribuciones a priori son derivadas en una inferencia Bayesiana de la distribución Gamma de dos parámetros. A prioris no informativas tales como las de Jeffrey, de referencia, MDIP, Tibshirani y una priori innovativa basada en la alternativa por cópulas son investigadas. Se muestra que una a priori de información de datos maximales conlleva a una a

^aProfessor. E-mail: femoala@fct.unesp.br

^bStudent. E-mail: pedrolramos@hotmail.com

^cProfessor. E-mail: achcar@fmrp.usp.br

posteriori impropia y que las diferentes escogencias del parámetro de interés permiten diferentes a prioris de referencia en este caso. Datos simulados permiten calcular las estimaciones Bayesianas e intervalos de credibilidad para los parámetros desconocidos así como la evaluación del desempeño de las distribuciones a priori evaluadas. El análisis Bayesiano se desarrolla usando métodos MCMC (Markov Chain Monte Carlo) para generar las muestras de la distribución a posteriori bajo las a priori consideradas.

Palabras clave: a prioris de Jeffrey, a prioris no informativas, conjugada, cópulas, distribución Gamma, MCMC, MDIP, ortogonal, referencia.

1. Introduction

The Gamma distribution is widely used in reliability analysis and life testing (see for example, Lawless 1982) and it is a good alternative to the popular Weibull distribution. It is a flexible distribution that commonly offers a good fit to any variable such as in environmental, meteorology, climatology, and other physical situations.

Let X be representing the lifetime of a component with a Gamma distribution, denoted by $\Gamma(\alpha, \beta)$ and given by

$$f(x | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}, \text{ for all } x > 0 \quad (1)$$

where $\alpha > 0$ and $\beta > 0$ are unknown shape and scale parameters, respectively.

There are many papers considering Bayesian inference for the estimation of the Gamma parameters. Son & Oh (2006) assume vague priors for the parameters to the estimation of parameters using Gibbs sampling. Apolloni & Bassis (2009) compute the joint probability distribution of the parameters without assuming any prior. They propose a numerical algorithm based on an approximate analytical expression of the probability distribution. Pradhan & Kundu (2011) assume that the scale parameter has a Gamma prior and the shape parameter has any log-concave prior and they are independently distributed. However, most of these papers have in common the use of proper priors and the assumption of independence a priori of the parameters. Although this is not a problem and have been much used in the literature we, would like to propose a noninformative prior for the Gamma parameters which incorporates the dependence structure of parameters. Some of priors proposed in the literature are Jeffreys (1967), MDIP (Zellner 1977, Zellner 1984, Zellner 1990, Tibshirani 1989), and reference prior (Bernardo 1979). Moala (2010) provides a comparison of these priors to estimate the Weibull parameters.

Therefore, the main aim of this paper is to present different noninformative priors for a Bayesian estimation of the two-parameter Gamma distribution. We also propose a bivariate prior distribution derived from copula functions (see for example, Nelsen 1999, Trivedi & Zimmer 2005a, Trivedi & Zimmer 2005b) in order to construct a prior distribution to capture the dependence structure between the parameters α and β .

We investigate the performance of the prior distributions through a simulation study using a small data set. Accurate inference for the parameters of the Gamma is obtained using MCMC (Markov Chain Monte Carlo) methods.

2. Maximum Likelihood Estimation

Let $X_1, \dots, X_n$ be a complete sample from (1) then the likelihood function is

$$L(\alpha, \beta \mid \mathbf{x}) = \frac{\beta^{n\alpha}}{[\Gamma(\alpha)]^n} \left(\prod_{i=1}^n x_i^{\alpha-1} \right) \exp \left\{ -\beta \sum_{i=1}^n x_i \right\} \quad (2)$$

for $\alpha > 0$ and $\beta > 0$.

Considering $\frac{\partial}{\partial \alpha} \log L$ and $\frac{\partial}{\partial \beta} \log L$ equal to 0 and after some algebraic manipulations we get the likelihood equations given by

$$\hat{\beta} = \frac{\hat{\alpha}}{\bar{X}} \quad \text{and} \quad \log \hat{\alpha} - \psi(\hat{\alpha}) = \log \left(\frac{\bar{X}}{\hat{X}} \right) \quad (3)$$

where $\psi(k) = \frac{\partial}{\partial k} \log \Gamma(k) = \frac{\Gamma'(k)}{\Gamma(k)}$ (see Lawless 1982) is the diGamma function, $\bar{X} = \frac{\sum_{i=1}^n x_i}{n}$ and $\hat{X} = \left(\prod_{i=1}^n x_i \right)^{1/n}$. The solutions for these equations provide the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$ for the parameters of the Gamma distribution (1). As closed form solution is not possible to evaluate (3), numerical techniques must be used. The Fisher information matrix is given by

$$I(\alpha, \beta) = \begin{bmatrix} \psi'(\alpha) & -\frac{1}{\beta} \\ -\frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{bmatrix} \quad (4)$$

where $\psi'(\alpha)$ is the derivative of $\psi(\alpha)$ called as triGamma function.

For large samples, approximated confidence intervals can be constructed for the parameters α and β through normal marginal distributions given by

$$\hat{\alpha} \sim N(\alpha, \sigma_1^2) \quad \text{and} \quad \hat{\beta} \sim N(0, \sigma_2^2), \quad \text{for } n \rightarrow \infty \quad (5)$$

where $\sigma_1^2 = \text{var}(\hat{\alpha}) = \frac{\hat{\alpha}}{\hat{\alpha}\psi'(\hat{\alpha})-1}$ and $\sigma_2^2 = \text{var}(\hat{\beta}) = \frac{\hat{\beta}^2\psi'(\hat{\alpha})}{\hat{\alpha}\psi'(\hat{\alpha})-1}$. In this case, the approximated 100(1- Γ)% confidence intervals for each parameter α and β are given by

$$\hat{\alpha} - z_{\frac{\Gamma}{2}}\sigma_1 < \alpha < \hat{\alpha} + z_{\frac{\Gamma}{2}}\sigma_1 \quad \text{and} \quad \hat{\beta} - z_{\frac{\Gamma}{2}}\sigma_2 < \beta < \hat{\beta} + z_{\frac{\Gamma}{2}}\sigma_2 \quad (6)$$

respectively.

3. Jeffrey's Prior

A well-known weak prior to represent a situation with little information about the parameters was proposed by Jeffreys (1967). This prior denoted by $\pi_J(\alpha, \beta)$ is derived from the Fisher information matrix $I(\alpha, \lambda)$ given in (4) as

$$\pi_J(\alpha, \beta) \propto \sqrt{\det I(\alpha, \beta)} \quad (7)$$

Jeffrey's prior is widely used due to its invariance property under one-to-one transformations of parameters although there has been an ongoing discussion about whether the multivariate form prior is appropriate.

Thus, from (4) and (7) the Jeffreys prior for (α, β) parameters is given by:

$$\pi_J(\alpha, \beta) \propto \frac{\sqrt{\alpha \psi'(\alpha) - 1}}{\beta} \quad (8)$$

4. Maximal Data Information Prior (MDIP)

It is of interest that the data gives more information about the parameter than the information on the prior density; otherwise, there would not be justification for the realization of the experiment. Thus, we wish a prior distribution $\pi(\phi)$ that provides a gain in the information supplied by data in which the largest possible relative to the prior information of the parameter, that is, which maximize the information on the data. With this idea Zellner (1977), Zellner (1984), Zellner (1990) and Min & Zellner (1993) derived a prior which maximize the average information in the data density relative to that one in the prior. Let

$$H(\phi) = \int_{R_x} f(x | \phi) \ln f(x | \phi) dx, x \in R_x \quad (9)$$

be the negative entropy of $f(x | \phi)$, the measure of the information in $f(x | \phi)$ and R_x the range of density $f(x | \phi)$. Thus, the following functional criterion is employed in the MDIP approach:

$$G[\pi(\phi)] = \int_a^b H(\phi) \pi(\phi) d\phi - \int_a^b \pi(\phi) \ln \pi(\phi) d\phi \quad (10)$$

which is the prior average information in the data density minus the information in the prior density. $G[\pi(\phi)]$ is maximized by selection of $\pi(\phi)$ subject to $\int_a^b \pi(\phi) d\phi = 1$. The solution is then a proper prior given by

$$\pi(\phi) = k \exp\{H(\phi)\} \quad a \leq \phi \leq b \quad (11)$$

where $k^{-1} = \int_a^b \exp\{H(\phi)\} d\phi$ is the normalizing constant.

Therefore, the MDIP is a prior that leads to an emphasis on the information in the data density or likelihood function. That is, its information is weak in comparison with data information.

Zellner (1977), Zellner (1984), Zellner (1990) shows several interesting properties of MDIP and additional conditions that can also be imposed to the approach reflection given initial information. However, the MDIP has restrictive invariance properties.

Theorem 1. Suppose that we do not have much prior information available about α and β . Under this condition, the prior distribution MDIP, denoted by $\pi_Z(\alpha, \beta)$, for the parameters (α, β) of the Gamma density (1) is given by:

$$\pi_Z(\alpha, \beta) \propto \frac{\beta}{\Gamma(\alpha)} \exp\left\{(\alpha - 1)\psi(\alpha) - \alpha\right\} \quad (12)$$

Proof. Firstly, we have to evaluate the measure information $H(\alpha, \beta)$ for the Gamma density which is given by

$$H(\alpha, \beta) = \int_0^\infty \ln\left(\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}\right) f(x | \alpha, \beta) dx \quad (13)$$

and after some algebra, the result is

$$H(\alpha, \beta) = \alpha \ln \beta - \ln \Gamma(\alpha) + (\alpha - 1) \int_0^\infty \ln(x) f(x | \alpha, \beta) dx - \beta E(X) \quad (14)$$

with $E(X) = \frac{\alpha}{\beta}$. □

Since the integral functions $\int_0^\infty u^{\alpha-1} e^{-u} du = \Gamma(\alpha)$ and $\int_0^\infty u^{\alpha-1} \log(u) e^{-u} du = \Gamma'(\alpha)$, the function (14) involving these integrals can be expressed as

$$H(\alpha, \beta) = -\ln \Gamma(\alpha) + \ln \beta + (\alpha - 1)\psi(\alpha) - \alpha \quad (15)$$

Therefore, the MDIP prior for the parameters α and β is given by

$$\pi_Z(\alpha, \beta) \propto \frac{\beta}{\Gamma(\alpha)} \exp\left\{(\alpha - 1)\psi(\alpha) - \alpha\right\} \quad (16)$$

However, the corresponding joint posterior density is not proper, but surprisingly, the prior density given by

$$\pi_Z(\alpha, \beta) \propto \frac{\beta}{\Gamma(\alpha)} \exp\left\{(\alpha - 1)\frac{\psi(\alpha)}{\Gamma(\alpha)} - \alpha\right\} \quad (17)$$

yields a proper posterior density. Thus, we will use (17) as MDIP prior in the numerical illustration in Section 8.

5. Reference Prior

Another well-known class of noninformative priors is the reference prior, first described by Bernardo (1979) and further developed by Berger & Bernardo (1992).

The idea is to derive a prior $\pi(\phi)$ that maximizes the expected posterior information about the parameters provided by independent replications of an experiment relative to the information in the prior. A natural measure of the expected information about ϕ provided by data $\mathbf{x}$ is given by

$$I(\phi) = E_{\mathbf{x}}[K(p(\phi | \mathbf{x}), \pi(\phi))] \quad (18)$$

where

$$K(p(\phi | \mathbf{x}), \pi(\phi)) = \int_{\Phi} p(\phi | \mathbf{x}) \log \frac{p(\phi | \mathbf{x})}{\pi(\phi)} d\phi \quad (19)$$

is the Kullback-Leibler distance. So, the reference prior is defined as the prior $\pi(\phi)$ that maximizes the expected Kullback-Leibler distance between the posterior distribution $p(\phi | \mathbf{x})$ and the prior distribution $\pi(\phi)$, taken over the experimental data.

The prior density $\pi(\phi)$ which maximizes the functional (19) is found through calculus of variation and, the solution is not explicit. However, when the posterior $p(\phi | \mathbf{x})$ is asymptotically normal, this approach leads to Jeffreys prior for a single parameter situation. If on the other hand, we are interested in one of the parameters, being the remaining parameters nuisances, the situation is quite different, and the appropriated reference prior is not a multivariate Jeffrey prior. Bernardo (1979) argues that when nuisance parameters are present the reference prior should depend on which parameter(s) are considered to be of primary interest. The reference prior in this case is derived as follows. We will present here the two-parameters case in details. For the multiparameter case, see Berger & Bernardo (1992).

Let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ be the whole parameter, θ_1 being the parameter of interest and θ_2 the nuisance parameter. The algorithm is as follows:

Step 1: Determine $\pi_2(\theta_2 | \theta_1)$, the conditional reference prior for θ_2 assuming that θ_1 is known, is given by,

$$\pi_2(\theta_2 | \theta_1) = \sqrt{I_{22}(\theta_1, \theta_2)} \quad (20)$$

where $I_{22}(\theta_1, \theta_2)$ is the (2,2)-entry of the Fisher Information Matrix.

Step 2: Normalize $\pi_2(\theta_2 | \theta_1)$.

If $\pi_2(\theta_2 | \theta_1)$ is improper, choose a sequence of subsets $\Omega_1 \subseteq \Omega_2 \subseteq \dots \rightarrow \Omega$ on which $\pi_2(\theta_2 | \theta_1)$ is proper. Define the normalizing constant and the proper prior $p_m(\theta_2 | \theta_1)$ respectively as

$$c_m(\theta_1) = \frac{1}{\int_{\Omega_m} \pi_2(\theta_2 | \theta_1) d\theta_2} \quad (21)$$

and

$$p_m(\theta_2 | \theta_1) = c_m(\theta_1) \pi_2(\theta_2 | \theta_1) 1_{\Omega_m}(\theta_2), \quad (22)$$

Step 3: Find the marginal reference prior $\pi_m(\theta_1)$ for θ_1 the reference prior for the experiment found by marginalizing out with respect to $p_m(\theta_2 | \theta_1)$. We obtain

$$\pi_m(\theta_1) \propto \exp \left\{ \frac{1}{2} \int_{\Omega_m} p_m(\theta_2 | \theta_1) \log \left\| \frac{\det I(\theta_1, \theta_2)}{I_{22}(\theta_1, \theta_2)} \right\| d\theta_2 \right\} \quad (23)$$

Step 4: Compute the reference prior $\pi_{\theta_1}(\theta_1, \theta_2)$ when θ_1 is the parameter of interest

$$\pi_{\theta_1}(\theta_1, \theta_2) = \lim_{m \rightarrow \infty} \left(\frac{c_m(\theta_1) \pi_m(\theta_1)}{c_m(\theta_1^*) \pi_m(\theta_1^*)} \right) \pi(\theta_2 | \theta_1) \quad (24)$$

where θ_1^* is any fixed point with positive density for all π_m .

We will derive the reference prior for the parameters of the Gamma distribution given in (1), where α will be considered as the parameter of interest and β the nuisance parameter.

Theorem 2. *The reference prior for the parameters of the Gamma distribution given in (1), where α will be considered as the parameter of interest and β the nuisance parameter, is given by:*

$$\pi_{\alpha}(\alpha, \beta) = \frac{1}{\beta} \sqrt{\frac{\alpha \psi'(\alpha) - 1}{\alpha}} \quad (25)$$

If β is the parameter of interest and α the nuisance, thus the prior is

$$\pi_{\beta}(\alpha, \beta) \propto \frac{\sqrt{\psi'(\alpha)}}{\beta} \quad (26)$$

Proof. By the approach proposed by Berger & Bernardo (1992), we find the reference prior for the nuisance parameter β , conditionally on the parameter of interest α , given by

$$\pi(\beta | \alpha) = \sqrt{I_{\beta\beta}(\alpha, \beta)} \propto \frac{1}{\beta} \quad (27)$$

where $I_{\beta\beta}(\alpha, \beta)$ is the (2,2)-entry of the Fisher Information Matrix given in (4). $\square$

As in Moala (2010), a natural sequence of compact sets for (α, β) is $(l_{1n}, l_{2n}) \times (q_{1n}, q_{2n})$, so that $l_{1i}, q_{1i} \rightarrow 0$ and $l_{2i}, q_{2i} \rightarrow \infty$ when $i \rightarrow \infty$. Therefore, the normalizing constant is given by,

$$c_i(\alpha) = \frac{1}{\int_{q_{1i}}^{q_{2i}} \frac{1}{\beta} d\beta} = \frac{1}{\log q_{2i} - \log q_{1i}}. \quad (28)$$

Now from (23), the marginal reference prior for α is given by

$$\pi_i(\alpha) = \exp \left\{ \frac{1}{2} \int_{q_{1i}}^{q_{2i}} c_i(\alpha) \frac{1}{\beta} \log \left\| \frac{\frac{\alpha \psi'(\alpha) - 1}{\beta^2}}{\frac{\alpha}{\beta^2}} \right\| d\beta \right\} \quad (29)$$

which after some mathematical arrangement, we have

$$\pi_i(\alpha) = \sqrt{\frac{\alpha \psi'(\alpha) - 1}{\alpha}} \exp \left\{ \frac{1}{2} c_i(\alpha) \int_{q_{1i}}^{q_{2i}} \frac{1}{\beta} d\beta \right\} \quad (30)$$

Therefore, the resulting marginal reference prior for α is given by

$$\pi_i(\alpha) \propto \sqrt{\frac{\alpha \psi'(\alpha) - 1}{\alpha}} \quad (31)$$

and the global reference prior for (α, β) with parameter of interest α is given by,

$$\pi_{\alpha}(\alpha, \beta) = \lim_{i \rightarrow \infty} \left(\frac{c_i(\alpha) \pi_i(\alpha)}{c_i(\alpha^*) \pi_i(\alpha^*)} \right) \pi(\beta | \alpha) \propto \frac{1}{\beta} \sqrt{\frac{\alpha \psi'(\alpha) - 1}{\alpha}} \quad (32)$$

considering $\alpha^* = 1$

Similarly we obtain the reference prior considering β as the parameter of interest and α as nuisance. In this case, the prior is

$$\pi_{\beta}(\alpha, \beta) \propto \frac{\sqrt{\psi'(\alpha)}}{\beta} \quad (33)$$

6. Tibishirani's Prior

Given a vector parameter ϕ , Tibshirani (1989) developed an alternative method to derive a noninformative prior $\pi(\delta)$ for the parameter of interest $\delta = t(\phi)$ so that the credible interval for δ has coverage error $O(n^{-1})$ in the frequentist sense. This means that the difference between the posterior and frequentist confidence interval should be small. To achieve that, Tibshirani (1989) proposed to reparametrize the model in terms of the orthogonal parameters (δ, λ) (see Cox & Reid 1987) where δ is the parameter of interest and λ is the orthogonal nuisance parameter. In this way, the approach specifies the weak prior to be any prior of the form

$$\pi(\delta, \lambda) = g(\lambda) \sqrt{I_{\delta\delta}(\delta, \lambda)} \quad (34)$$

where $g(\lambda) > 0$ is an arbitrary function and $I_{\delta\delta}(\delta, \lambda)$ is the “delta” entry of the Fisher Information Matrix.

Theorem 3. *The Tibishirani's prior distribution $\pi_T(\alpha, \beta)$ for the parameters (α, β) of the Gamma distribution given in (1) by considering α as the parameter of interest and β the nuisance parameter is given by:*

$$\pi_T(\alpha, \beta) \propto \frac{1}{\beta} \sqrt{\frac{\alpha \psi'(\alpha) - 1}{\alpha}} \quad (35)$$

Proof. For the Gamma model (1), we will propose an orthogonal reparametrization (δ, λ) where $\delta = \alpha$ is the parameter of interest and λ is the nuisance parameter to be evaluated. The orthogonal parameter λ is obtained by solving the differential equation:

$$I_{\beta\beta} \frac{\partial \beta}{\partial \alpha} = -I_{\alpha\beta} \quad (36)$$

□

From (4) and (36) we have

$$\frac{\alpha}{\beta^2} \frac{\partial \beta}{\partial \alpha} = \frac{1}{\beta} \quad (37)$$

Separating the variables, (37) becomes the following,

$$\frac{1}{\beta} \partial \beta = \frac{1}{\alpha} \partial \alpha \quad (38)$$

Integrating both sides we get,

$$\log \beta = \log \alpha + h(\lambda) \quad (39)$$

where $h(\lambda)$ is an arbitrary function of λ .

By choosing $h(\lambda) = \log \lambda$, we obtained the solution to (36), the nuisance parameter λ orthogonal to δ ,

$$\lambda = \frac{\beta}{\alpha} \quad (40)$$

Thus, the information matrix for the orthogonal parameters is given by

$$I(\delta, \lambda) = \begin{bmatrix} \psi'(\delta) - \frac{1}{\delta} & 0 \\ 0 & \frac{\delta}{\lambda^2} \end{bmatrix} \quad (41)$$

From (34) and (41), the corresponding prior for (δ, λ) is given by

$$\pi_{\delta}(\delta, \lambda) \propto g(\lambda) \sqrt{\frac{\delta \psi'(\delta) - 1}{\delta}} \quad (42)$$

where $g(\lambda)$ is an arbitrary function.

Due to a lack of uniqueness in the choice of the orthogonal parametrization, then the class of orthogonal parameters is of the form $g(\lambda)$, where $g(\cdot)$ is any reparametrization. This non-uniqueness is reflected by the function $g(\cdot)$ corresponding to (26). One possibility, in the single nuisance parameter case, is to require that (δ, λ) also satisfies Stein's condition (see Tibshirani 1989) for λ with p taken as the nuisance parameter. Under this condition we obtain

$$\pi_{\lambda}(\delta, \lambda) \propto g^*(\delta) \frac{\sqrt{\delta}}{\lambda} \quad (43)$$

Now, requiring $g(\lambda) \sqrt{\frac{\delta \psi'(\delta) - 1}{\delta}} = g^*(\delta) \frac{\sqrt{\delta}}{\lambda}$ we have that

$$\pi_T(\delta, \lambda) \propto \frac{1}{\lambda} \sqrt{\frac{\delta \psi'(\delta) - 1}{\delta}} \quad (44)$$

Thus, from (40), the prior expressed in terms of the (α, β) parametrization is given by

$$\pi_T(\alpha, \beta) \propto \frac{1}{\beta} \sqrt{\frac{\alpha \psi'(\alpha) - 1}{\alpha}} \quad (45)$$

Note that this prior coincides with reference prior (25) considering α as the parameter of interest.

7. Copula Prior

In this section we derive a bivariate prior distribution from copula functions (see for example, Nelsen 1999, Trivedi & Zimmer 2005*a*, Trivedi & Zimmer 2005*b*) in order to construct a prior distribution to capture the dependence structure between the parameters α and β . Copulas can be used to correlate two or more random variables and they provide great flexibility to fit known marginal densities.

A special case is given by the Farlie-Gumbel-Morgenstern copula which is suitable to model weak dependences (see Morgenstern 1956) with corresponding bivariate prior distribution for α and β given ρ ,

$$\pi(\alpha, \beta \mid \rho) = f_1(\alpha)f_2(\beta) + \rho f_1(\alpha)f_2(\beta)[1 - 2F_1(\alpha)][1 - 2F_2(\beta)], \quad (46)$$

where $f_1(\alpha)$ and $f_2(\beta)$ are the marginal densities for the random quantities α and β ; $F_1(\alpha)$ and $F_2(\beta)$ are the corresponding marginal distribution functions for α and β , and $-1 \leq \rho \leq 1$.

Observe that if $\rho = 0$, we have independence between α and β .

Different choices could be considered as marginal distributions for α and β as Gamma, exponential, Weibull or uniform distributions.

In this paper, we will assume Gamma marginal distribution $\Gamma(a_1, b_1)$ and $\Gamma(a_2, b_2)$ for α and β , respectively, with known hyperparameters a_1, a_2, b_1 and b_2 . Thus,

$$\begin{aligned} \pi(\alpha, \beta \mid a_1, a_2, b_1, b_2, \rho) \propto \alpha^{a_1-1} \beta^{a_2-1} \exp\{-b_1\alpha - b_2\beta\} \times \\ \left[1 + \rho(1 - 2I(a_1, b_1\alpha))(1 - 2I(a_2, b_2\beta))\right] \end{aligned} \quad (47)$$

where $I(k, x) = \frac{1}{\Gamma(k)} \int_0^x u^{k-1} e^{-u} du$ is the incomplete Gamma function.

Assuming the prior (47), the joint posterior distribution for α, β and ρ is given by,

$$\begin{aligned} p(\alpha, \beta, \rho \mid x) \propto \frac{\beta^{n\alpha}}{[\Gamma(\alpha)]^n} \left(\prod_{i=1}^n x_i^{\alpha-1} \right) \\ \exp\left\{-\beta \sum_{i=1}^n x_i\right\} \pi(\alpha, \beta \mid a_1, a_2, b_1, b_2, \rho) \pi(\rho) \end{aligned} \quad (48)$$

where $\pi(\rho)$ is a prior distribution for ρ .

In general, many different priors can be used for ρ ; one possibility is to consider an uniform prior distribution for ρ over the interval $[-1, 1]$.

8. Numerical Illustration

8.1. Simulation Study

In this section, we investigate the performance of the proposed prior distributions through a simulation study with samples of size $n = 5$, $n = 10$ and $n = 30$ generated from the Gamma distribution with parameters $\alpha = 2$ and $\beta = 3$.

As we do not have an analytic form for marginal posterior distributions we need to appeal to the MCMC algorithm to obtain the marginal posterior distributions and hence to extract characteristics of parameters such as Bayes estimators and credible intervals. The chain is run for 10,000 iterations with a burn-in period of 1,000. Details of the implementation of the MCMC algorithm used in this paper are given below.

- i) choose starting values α_0 and β_0 .
- ii) at step $i + 1$, we draw a new value α_{i+1} conditional on the current α_i from the Gamma distribution $\Gamma(\alpha_i/c, c)$;
- iii) the candidate α_{i+1} will be accepted with a probability given by the Metropolis ratio

$$u(\alpha_i, \alpha_{i+1}) = \min \left\{ 1, \frac{\Gamma(\alpha_i/c, c)p(\alpha_{i+1}, \beta_i | \mathbf{x})}{\Gamma(\alpha_{i+1}/c, c)p(\alpha_i, \beta_i | \mathbf{x})} \right\}$$

- iv) sample the new value β_{i+1} from the Gamma distribution $\Gamma(\beta_i/d, d)$;
- v) the candidate β_{i+1} will be accepted with a probability given by the Metropolis ratio

$$u(\beta_i, \beta_{i+1}) = \min \left\{ 1, \frac{\Gamma(\beta_i/d, d)p(\alpha_{i+1}, \beta_{i+1} | \mathbf{x})}{\Gamma(\beta_{i+1}/d, d)p(\alpha_{i+1}, \beta_i | \mathbf{x})} \right\}$$

The proposal distribution parameters c and d were chosen to obtain a good mixing of the chains and the convergence of the MCMC samples of parameters are assessed using the criteria given by Raftery and Lewis (1992). More details of MCMC in order to construct these chains see, for example, Smith & Roberts (1993), Gelfand & Smith (1990), Gilks, Clayton, Spiegelhalter, Best, McNeil, Sharples & Kirby (1993).

We examine the performance of the priors by computing point estimates for parameters α and β based on 1,000 simulated samples and then we averaged the estimates of the parameters, obtain the variances and the coverage probability of 95% confidence intervals. Table 1 shows the point estimates for α and its respective variances given between parenthesis. Table 2 shows the same summaries for β .

The results of our numerical studies show that there is little difference between the point estimates for both parameters α and β . However, the MDIP prior produces a much smaller variance than using the other assumed priors. The uniform prior and MLE estimate produce bad estimations with large variances showing

TABLE 1: Summaries for parameter α .

$\alpha = 2$	Jeffreys	MDIP	Tibshirani	Reference	Copula	Uniform	MLE
$n = 5$	2.2529 (2.1640)	2.1894 (0.5112)	2.3666 (3.0297)	2.3602 (2.4756)	2.0909 (2.1987)	3.3191 (3.1577)	3.2850 (7.5372)
$n = 10$	2.5227 (1.3638)	2.2253 (0.3855)	2.4138 (1.3849)	2.4761 (1.3301)	2.3068 (1.2052)	2.9769 (1.4658)	2.7013 (1.8308)
$n = 30$	2.1259 (0.2712)	2.1744 (0.1910)	2.0606 (0.2651)	2.1079 (0.2728)	2.0369 (0.2571)	2.2504 (0.2829)	2.2253 (0.3138)

TABLE 2: Summaries for parameter β .

$\beta = 3$	Jeffreys	MDIP	Tibshirani	Reference	Copula	Uniform	MLE
$n = 5$	2.9136 (4.2292)	3.2680 (1.6890)	3.2161 (6.1384)	3.1058 (4.7787)	2.7486 (4.2447)	4.3419 (5.6351)	5.2673 (23.1881)
$n = 10$	3.8577 (3.8086)	3.7727 (1.5872)	3.6705 (3.8110)	3.7649 (3.6667)	3.6112 (3.5599)	4.5186 (3.7803)	4.2960 (5.4296)
$n = 30$	3.2328 (0.7950)	3.4255 (0.6292)	3.1475 (0.7861)	3.1798 (0.7856)	3.0805 (0.7453)	3.4633 (0.8335)	3.4005 (0.9163)

that, despite being widely used in the literature, they are not suitable for the Gamma distribution. As expected, the performance of these priors improves when the sample size increases.

Frequentist property of coverage probabilities for the parameters α and β have also been studied to compare the priors and MLE. Table 3 summarizes the simulated coverage probabilities of 95% confidence intervals. For the three sample sizes considered here, the intervals of MDIP prior produce an over-coverage for small sample sizes while, the intervals of uniform prior and MLE seem to have an under-coverage for some cases. Coverage probabilities are very close to the nominal value when n increases.

TABLE 3: Frequentist coverage probability of the 95% confidence intervals for α and β .

$\alpha = 2$	Jeffreys	MDIP	Tibshirani	Reference	Copula	Uniform	MLE
$n = 5$	96.30%	99.60%	97.20%	96.10%	95.30%	95.70%	95.60%
$n = 10$	96.40%	99.50%	94.90%	95.00%	95.80%	90.60%	95.30%
$n = 30$	96.10%	96.20%	98.10%	95.80%	96.80%	95.50%	95.00%
$\beta = 3$	Jeffreys	MDIP	Tibshirani	Reference	Copula	Uniform	MLE
$n = 5$	96.60%	99.60%	97.50%	97.00%	96.30%	99.90%	94.30%
$n = 10$	98.10%	98.00%	96.00%	96.70%	97.80%	93.90%	95.70%
$n = 30$	97.30%	95.80%	96.90%	96.10%	97.40%	94.90%	96.80%

8.2. Rainfall Data Example

Data in Table 4 represent the average monthly rainfall obtained from the Information System for Management of Water Resources of the State of São Paulo, including a period of 56 years from 1947 to 2003, by considering the month of November.

Let us assume a Gamma distribution with density (1) to analyse the data.

TABLE 4: Historical rainfall averages over last 56 years in State of São Paulo.

0.2,3.5,2.8,3.7,8.7,6.9,7.4,0.8,4.8,2.5,2.9,3.1,4.0,5.0,3.8,3.5,5.4,3.3,2.9,
1.7,7.3,2.9,4.6,1.1,1.4,3.9,6.2,4.1,10.8,3.8,7.3,1.8,6.7,3.5,3.2,5.2,2.8,5.2,
5.4,2.2,9.9,2.1,4.7,5.5,2.6,4.1,5.4,5.5,2.1,1.9,8.8,1.3,24.1,5.4,6.2,2.9

Table 5 presents the posterior means assuming the different prior distributions and maximum likelihood estimates (MLE) for the parameters α and β .

TABLE 5: Posterior means for parameters α and β of rainfall data.

	Uniform	Jeffreys	Ref- β	MDIP	Tibshirani	Copula	MLE
α	2.493	2.387	2.393	2.659	2.357	2.380	2.395
β	0.543	0.516	0.517	0.641	0.510	0.515	0.518

From Table 5, we observe similar inference results assuming the different prior distributions for α and β , except for MDIP prior as observed in the simulation study introduced in the example presented in section 8.1.

The 95% posterior credible intervals obtained using the different priors for the parameters are displayed in Table 6. The MLE intervals for the parameters α and β are given respectively by (1.56; 3.22) and (0.31; 0.72).

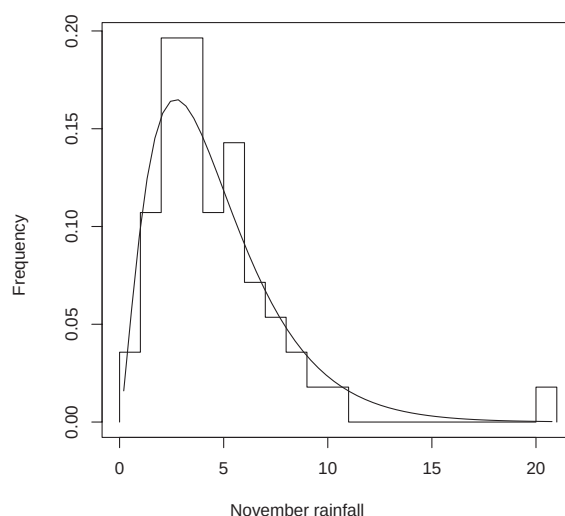


FIGURE 1: Histogram and fitted Gamma distribution for rainfall data.

TABLE 6: 95% posterior intervals for the parameters α and β of rainfall data.

	Uniform	Jeffreys	Ref- β	MDIP	Tibshirani	Copula
α	(1.71; 3.43)	(1.63; 3.29)	(1.64; 3.28)	(1.91; 3.52)	(1.60; 3.25)	(1.60; 3.34)
β	(0.35; 0.76)	(0.33; 0.73)	(0.34; 0.73)	(0.44; 0.87)	(0.33; 0.73)	(0.32; 0.75)

Figure 2 shows the marginal posterior densities for both parameters α and β . We can see that the MDIP prior leads to a posterior slightly more sharply peaked for both parameters, while the other priors are quite similar, agreeing with simulated data with sample size $n = 30$.

To determine the appropriate prior distribution to be used with the rainfall data fitted by the Gamma distribution, some selection criteria can be examined. These include information-based criteria (AIC, BIC and DIC) given in the Table 7 for each prior distribution.

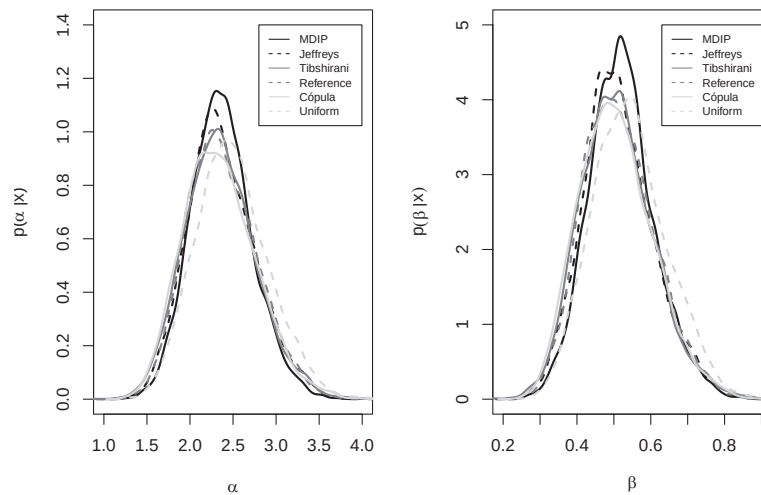


FIGURE 2: Plot of marginal posterior densities for the parameters α and β of rainfall data.

TABLE 7: Information-based model selection criteria (AIC, BIC and DIC) for rainfall data.

Prior	AIC	BIC	DIC
Jeffreys	272.213	268.162	267.827
MDIP	272.247	268.196	267.502
Ref- β	272.212	268.162	267.922
Tibshirani	272.219	268.169	268.068
Copula	272.222	268.171	268.197
Uniform	272.266	268.215	267.935

From the results of Table 7 and Figure 2 we observe that the choice of the prior distributions for parameters α and β has a negligible effect on the posterior distribution, surely due to the large amount of data in this study.

8.3. Reliability Data Example

In this example, we consider a lifetime data set related to an electrical insulator subjected to constant stress and strain introduced by Lawless (1982). The dataset does not have censored values and represent the lifetime (in minutes) to failure: 0.96, 4.15, 0.19, 0.78, 8.01, 31.75, 7.35, 6.50, 8.27, 33.91, 32.52, 16.03, 4.85, 2.78, 4.67, 1.31, 12.06, 36.71 and 72.89. Let us denote this data as “Lawless data”. We assume a Gamma distribution with density (1) to analyse the data.

The maximum likelihood estimators and the Bayesian summaries for α and β , considering the different prior distributions are given in Table 8. Table 9 shows the 95% posterior intervals for α and β . The estimated marginal posterior distributions for the parameters are shown in Figure 3.

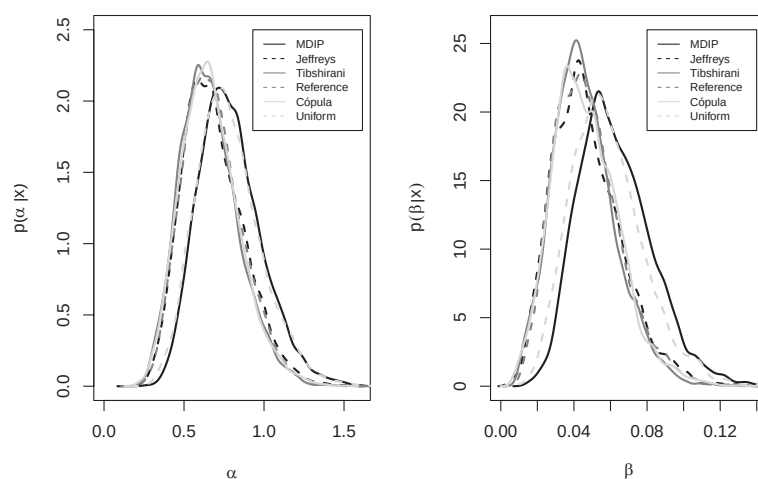


FIGURE 3: Plot of marginal posterior densities for the parameters α and β for Lawless data.

Tables 8 and 9 present the posterior statistics and 95% confidence intervals for both parameters resulting from the proposed priors. Again the performance of the MDIP prior clashes from the others.

TABLE 8: Posterior means for parameters α and β (Lawless data).

	Uniform	Jeffreys	Ref- β	MDIP	Tibshirani	Copula	MLE
α	0.779	0.686	0.681	0.789	0.660	0.666	0.690
β	0.058	0.047	0.047	0.063	0.046	0.047	0.048

TABLE 9: 95% posterior intervals for the parameters α and β (Lawless data).

	Uniform	Jeffreys	Ref- β	MDIP	Tibshirani	Copula
α	(0.433, 1.229)	(0.371, 1.111)	(0.374, 1.087)	(0.459, 1.232)	(0.350, 1.061)	(0.353, 1.080)
β	(0.025, 0.105)	(0.017, 0.091)	(0.018, 0.090)	(0.031, 0.108)	(0.017, 0.085)	(0.017, 0.087)

Table 10 shows the AIC, BIC and DIC values for all priors under investigation, with similar results as presented in Table 7 are obtained in this comparison which shows no differences using the different assumed priors.

TABLE 10: Information-based model selection criteria (AIC, BIC and DIC) for (Lawless data).

Prior	AIC	BIC	DIC
Jeffreys	143.125	141.236	141.409
MDIP	143.642	141.753	141.082
Ref- β	143.126	141.237	141.168
Tibshirani	143.148	141.259	141.247
Copula	143.138	141.249	141.471
Uniform	143.401	141.512	141.119

9. Conclusion and Discussion

The large number of noninformative priors can cause difficulties in the choosing one, especially when these priors does not produce similar results. Thus, in this paper, we presented a Bayesian analysis using a variety of prior distributions for the estimation of the parameters of the Gamma distribution.

We have shown that the use of the maximal data information process proposed by Zellner (1977), Zellner (1984), Zellner (1990) yields an improper posterior distribution for the parameters α and β . In this way, we proposed a “modified” MDIP prior analytically similar to the original one but with proper posterior. We also shown that the reference prior provides nonuniqueness of prior due to the choice of the parameter of interest, although the simulation shows the same performance. We have shown that the Tibshirani prior applied to the parameters of the Gamma distribution is equal to the reference prior when α is the parameter of interest.

Besides, a simulation study to check the impact of the use of different noninformative priors in the posterior distributions was also carried out. From this study we can conclude that it is necessary to carefully choose a prior for the parameters of the Gamma distribution when there is not enough data.

As expected, a moderated large sample size is need to achieve the desirable accuracy. In this case, the choice of the priors become irrelevant. However, the disagreement is substantial for small sample sizes.

Our simulation study indicates that the class of priors: Jeffreys, Reference, Tibshirani and Copula, had the same performance while the Uniform prior had worse performance. On the other hand, the “modified” MDIP prior produced the best estimations for α and β . Thus, the simulation study showed that the effect of the prior distributions can be substantial in the estimation of parameters and therefore the modified MDIP prior should be the recommended noninformative prior for the estimation of parameters of the Gamma distribution.

[Recibido: enero de 2013 — Aceptado: septiembre de 2013]

References

- Apolloni, B. & Bassis, S. (2009), 'Algorithmic inference of two-parameter gamma distribution', *Communications in Statistics - Simulation and Computation* **38**(9), 1950–1968.
- Berger, J. & Bernardo, J. M. (1992), On the development of the reference prior method, Fourth Valencia International Meeting on Bayesian Statistics, Spain.
- Bernardo, J. M. (1979), 'Reference posterior distributions for Bayesian inference', *Journal of the Royal Statistical Society* **41**(2), 113–147.
- Cox, D. R. & Reid, N. (1987), 'Parameter orthogonality and approximate conditional inference (with discussion)', *Journal of the Royal Statistical Society, Series B* **49**, 1–39.
- Gelfand, A. E. & Smith, F. M. (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association* **85**, 398–409.
- Gilks, W., Clayton, D., Spiegelhalter, D., Best, N., McNeil, A., Sharples, L. & Kirby, A. (1993), 'Modeling complexity: Applications of Gibbs sampling in medicine', *Journal of the Royal Statistical Society, Series B* **55**(1), 39–52.
- Jeffreys, S. H. (1967), *Theory of Probability*, 3 edn, Oxford University Press, London.
- Lawless, J. (1982), *Statistical Models and Methods for Lifetime Data*, John Wiley, New York.
- Min, C.-k. & Zellner, A. (1993), Bayesian Analysis, Model Selection and Prediction, in 'Physics and Probability: Essays in honor of Edwin T Jaynes', Cambridge University Press, pp. 195–206.
- Moala, F. (2010), 'Bayesian analysis for the Weibull parameters by using noninformative prior distributions', *Advances and Applications in Statistics* (14), 117–143.
- Morgenstern, D. (1956), 'Einfache beispiele sw edimensionaler vertielung', *Mit Mathematics Statistics* **8**, 234–235.
- Nelsen, R. B. (1999), *An Introduction to Copulas*, Springer Verlag, New York.
- Pradhan, B. & Kundu, D. (2011), 'Bayes estimation and prediction of the two-parameter Gamma distribution', *Journal of Statistical Computation and Simulation* **81**(9), 1187–1198.
- Smith, A. & Roberts, G. (1993), 'Bayesian computation via the Gibbs sampler and related Markov Chain Monte Carlo Methods', *Journal of the Royal Statistical Society: Series B* **55**, 3–24.

- Son, Y. & Oh, M. (2006), 'Bayesian estimation of the two-parameter Gamma distribution', *Communications in Statistics – Simulation and Computation* **35**, 285–293.
- Tibshirani, R. (1989), 'Noninformative priors for one parameters of many', *Biometrika* **76**, 604–608.
- Trivedi, P. K. & Zimmer, D. M. (2005a), *Copula Modelling*, New Publishers, New York.
- Trivedi, P. K. & Zimmer, D. M. (2005b), 'Copula modelling: An introduction to practitioners', *Foundations and Trends in Econometrics* .
- Zellner, A. (1977), Maximal data information prior distributions, in A. Aykac & C. Brumat, eds, 'In New Methods in the Applications of Bayesian Methods', North-Holland, Amsterdam.
- Zellner, A. (1984), *Maximal Data Information Prior Distributions, Basic Issues in Econometrics*, The University of Chicago Press, Chicago, USA.
- Zellner, A. (1990), Bayesian methods and entropy in economics and econometrics, in W. J. Grandy & L. Schick, eds, 'Maximum Entropy and Bayesian Methods', Dordrecht, Netherlands: Kluwer Academic Publishers, pp. 17–31.