



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

Faculdade de Ciências e Tecnologia - Câmpus de Presidente Prudente

Jonathan Perez Conde

Elicitação de Priori derivada da Aproximação de Laplace

Presidente Prudente, São Paulo
2024

Jonathan Perez Conde

Elicitação de Priori derivada da Aproximação de Laplace

Relatório Final de Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Estatística da FCT/UNESP para aproveitamento na disciplina Trabalho de Conclusão de Curso.
Orientador: Prof. Dr. Fernando Antonio Moala

Presidente Prudente, São Paulo
2024

C745e

Conde, Jonathan Perez

Elicitação de Priori Derivada da Aproximação de Laplace /
Jonathan Perez Conde. -- Presidente Prudente, 2024
36 p. : tabs.

Trabalho de conclusão de curso (Bacharelado - Estatística) -
Universidade Estadual Paulista (UNESP), Faculdade de Ciências e
Tecnologia, Presidente Prudente
Orientador: Fernando Antonio Moala

1. Bayesiana. 2. Priori Informativa. 3. Aproximação de Laplace. 4.
Distribuição Gumbel. I. Título.

TERMO DE APROVAÇÃO

Jonathan Perez Conde

ELICITAÇÃO DE PRIORI DERIVADA DA APROXIMAÇÃO DE LAPLACE

Relatório de Final de Trabalho de Conclusão de Curso aprovado como requisito para obtenção de créditos na disciplina Trabalho de Conclusão do curso de graduação em Estatística da Faculdade de Ciências e Tecnologia da Unesp, pela seguinte banca examinadora:

Orientador: _____

Prof. Dr. Fernando Antonio Moala
Departamento de Estatística

Prof. Dr. Mário Hissamitsu Tarumoto
Departamento de Estatística

Prof. Dr. Sérgio Minoru Oikawa
Departamento de Estatística

Presidente Prudente, 06 de dezembro de 2024.

Resumo

Este trabalho apresentou a aplicação de uma metodologia para estimar os hiperparâmetros de uma distribuição a priori, elicitada com base em um sistema de quantis derivado do método de Laplace. Amplamente utilizado na estatística Bayesiana antes do MCMC, o método de Laplace era empregado para calcular distribuições marginais a posteriori. Neste trabalho, o método de Laplace foi utilizado para ajustar os hiperparâmetros da distribuição a priori de uma distribuição Gumbel, enquanto o algoritmo MCMC foi aplicado para obter amostras das distribuições marginais a posteriori. Três distribuições a priori foram consideradas: a de Jeffreys, derivada da informação de Fisher; a distribuição Flat, caracterizada por sua grande variância; e a distribuição elicitada, desenvolvida a partir da metodologia apresentada. Os resultados das distribuições posteriores, obtidos por meio do MCMC, evidenciaram que a priori elicitada resultou em estimadores com menor viés e menor erro quadrático médio (EQM) em relação às demais. Embora o processo de elicitación tenha demandado o uso de métodos numéricos avançados, dada a complexidade do sistema de quantis, os resultados demonstram que a abordagem baseada no método de Laplace é eficiente e apresenta grande potencial para a elicitación de distribuições a priori na estatística Bayesiana.

Palavras-chave: Bayesiana; priori informativa; aproximação de Laplace; distribuição Gumbel.

Abstract

This study presented the application of a methodology to estimate the hyperparameters of a prior distribution, elicited based on a quantile system derived from Laplace's method. Widely used in Bayesian statistics prior to the advent of MCMC, Laplace's method was employed to compute posterior marginal distributions. In this work, Laplace's method was utilized to adjust the hyperparameters of the prior distribution of a Gumbel distribution, while the MCMC algorithm was applied to obtain samples from the posterior marginal distributions. Three prior distributions were considered: Jeffreys', derived from Fisher information; the Flat distribution, characterized by its large variance; and the elicited distribution, developed based on the proposed methodology. The results of the posterior distributions obtained through MCMC showed that the elicited prior produced estimators with lower bias and mean squared error (MSE) compared to the others. Although the elicitation process required the use of advanced numerical methods due to the complexity of the quantile system, the results demonstrate that the approach based on Laplace's method is efficient and holds significant potential for the elicitation of prior distributions in Bayesian statistics.

Keywords: Bayesian; informative prior; Laplace method; Gumbel distribution.

Sumário

1	Introdução	8
2	Distribuição Gumbel	10
2.1	Propriedades	11
2.2	Estimadores de máxima verossimilhança	11
2.3	Geração de amostras aleatórias da distribuição Gumbel	12
3	Análise Bayesiana	13
3.1	Priori de Jeffreys	14
3.2	Produto de prioris independentes	14
3.2.1	Priori Flat	15
3.2.2	Priori Elicitada I	15
3.2.3	Priori Elicitada II	16
4	Fundamentos da priori de Laplace	17
4.1	Elicitação	17
4.2	Metodologia de Elicitação Usando Aproximação de Laplace	18
4.3	Aplicação a Distribuição de Gumbel	19
5	Aplicação dos modelos	23
5.1	Método de Monte Carlo via Cadeia de Markov	23
5.1.1	Algoritmo para gerar ϕ e φ	23
5.2	Distribuições de Proababilidade a Posteriori	25
5.3	Dados simulados	26
6	Conclusão	29
6.1	Propostas futuras	29
	Apêndice A Tabelas de Resultados	31

Apêndice B	Códigos na linguagem R	32
B.1	Geração de Amostras Marginais da Distribuição de Probabilidade Posterior	32
B.2	Resolução de Sistema de Quantis Derivadas da Aproximação de Laplace . .	34

Capítulo 1

Introdução

Eventos raros, como inundações, deslizamentos de terra, furacões, terremotos ou crises financeiras são pouco frequentes, mas apesar da baixa probabilidade de ocorrência desses eventos, o estudo deles tem se intensificado, principalmente devido ao impacto devastador que poderiam causar, caso acontecessem. Um dos principais exemplos é o caso da Holanda, em 1953, quando o rompimento das barragens que protegiam o país contra o avanço do mar resultou na morte de mais de 1.800 pessoas. Após esse desastre, a Holanda adotou a teoria de valores extremos, também conhecida como teoria de eventos raros, para projetar a altura de suas barragens e prevenir novas catástrofes.

Do ponto de vista estatístico, criar modelos para prever eventos raros é um desafio. As metodologias tradicionais, que dependem de grandes amostras para fornecer estimativas precisas, não são adequadas para esse tipo de fenômeno que carece de dados.

É nesse contexto que a metodologia Bayesiana se destaca, oferecendo uma abordagem vantajosa. Ao contrário das metodologias clássicas, que dependem exclusivamente dos dados observados, a inferência Bayesiana também utiliza um conhecimento a priori sobre o problema. Esse conhecimento pode ser fornecido por um especialista da área e, quando combinado com os dados observados, pode levar a estimativas mais robustas e precisas.

Entretanto, o grande desafio é como transformar o conhecimento subjetivo do especialista em uma informação útil para a análise estatística. A eliciação de distribuições a priori é uma área da inferência Bayesiana dedicada a esse processo. Embora existam várias técnicas para isso, o processo de eliciação ainda enfrenta dificuldades devido ao caráter subjetivo das informações fornecidas pelos especialistas.

Uma abordagem prática para superar a dificuldade na eliciação de prioris consiste em obter os quantis fornecidos por especialistas e, a partir deles, determinar os hiperparâmetros das distribuições a priori. Perguntas como "qual é a probabilidade de determinado evento ocorrer?" são mais intuitivas e compreensíveis, facilitando esse processo. No entanto, essa abordagem requer o uso de funções fechadas, o que implica realizar integrais envolvendo a distribuição a priori e função de distribuição de probabilidade que modela o problema. Essas integrações frequentemente apresentam elevada complexidade, tornando difícil obter soluções analíticas simples.

Para contornar esse problema, o método de Laplace surge como uma solução eficaz, fornecendo boas aproximações para integrais complexas. Essa técnica desempenhou

um papel crucial no desenvolvimento da estatística bayesiana, especialmente antes da popularização dos métodos de Cadeia de Markov Monte Carlo (MCMC). Durante esse período, o método de Laplace era amplamente utilizado para resolver integrais múltiplas e derivar distribuições marginais, compensando a falta de ferramentas computacionais avançadas. Com o advento do MCMC, este se tornou a principal ferramenta para obter as amostras das distribuições marginais a posteriori, substituindo, em muitos casos, o método de Laplace.

Neste estudo, utilizaremos o método de Laplace para integrar funções da forma:

$$P(X \leq x_p) = \int_0^\infty \cdots \int_0^\infty \left[\int_0^{x_p} f_X(x|\boldsymbol{\theta}) dx \right] \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}, \quad \boldsymbol{\theta} = (\theta_1, \dots, \theta_n),$$

onde $f_X(x|\boldsymbol{\theta})$ é a distribuição de probabilidade dos dados, $\pi(\boldsymbol{\theta})$ é a distribuição de probabilidade da priori, e $\boldsymbol{\theta}$ representa os parâmetros da priori, ou os hiperparâmetros da distribuição de probabilidade a posteriori $f_X(x|\boldsymbol{\theta})\pi(\boldsymbol{\theta})$.

O método de Laplace já foi aplicado com sucesso em estudos anteriores, como demonstrado em (MOALA; DEY, 2018), que utilizou essa técnica para elicitar os hiperparâmetros de uma priori baseada em duas distribuições gama independentes.

Para implementar as metodologias propostas, será empregada a distribuição de valor extremo do máximo, conhecida como distribuição de Gumbel. Este trabalho está organizado em seis capítulos, conforme descrito a seguir:

- **Capítulo 2:** Apresenta um estudo aprofundado das principais propriedades da distribuição Gumbel. São exploradas a função de distribuição acumulada, esperança, variância, função geradora de momentos, métodos para geração de amostras aleatórias e estimadores de máxima verossimilhança.
- **Capítulo 3:** Inicia a análise Bayesiana, utilizando distribuições a priori não informativas, como a priori de Jeffreys, para estimar os parâmetros ϕ e φ . Além disso, é introduzida a priori de Laplace, cuja metodologia é detalhada no capítulo seguinte.
- **Capítulo 4:** Discute o processo de elicitação e apresenta a metodologia utilizada para determinar os hiperparâmetros da priori de Laplace.
- **Capítulo 5:** Introduz o algoritmo MCMC e realiza uma análise comparativa das distribuições posteriores obtidas a partir das diferentes priors apresentadas no capítulo 3. Adicionalmente, compara-se o desempenho dos estimadores Bayesianos com o estimador de máxima verossimilhança.
- **Capítulo 6:** Por ultimo é apresentado as conclusões gerais sobre a metodologia proposta, destacando os principais resultados e possíveis extensões para estudos futuros.

Capítulo 2

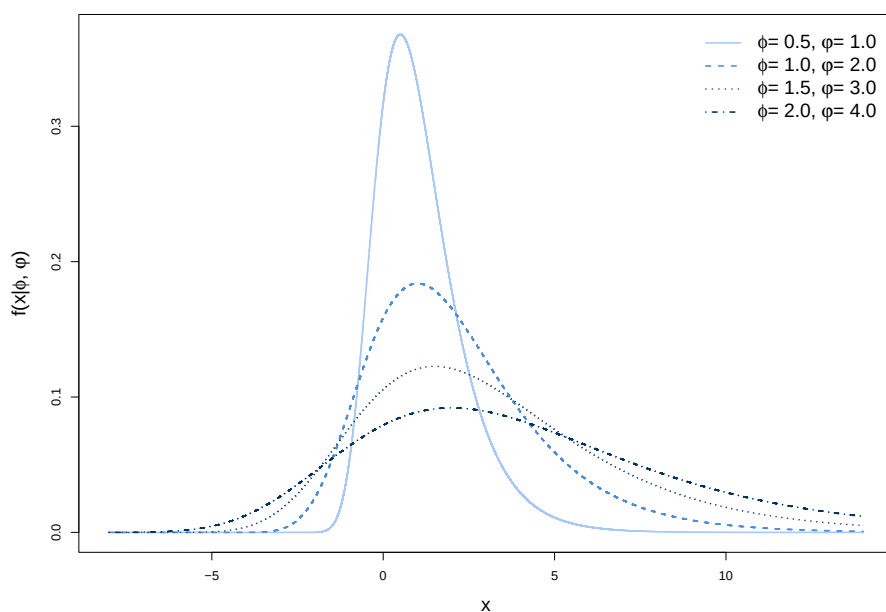
Distribuição Gumbel

A Distribuição de valores extremos do tipo I, também conhecida como distribuição Gumbel, é comumente utilizada para analisar eventos extremos, como chuvas máximas, níveis de enchentes, ventos fortes e temperaturas extremas. Seja X uma variável aleatória com distribuição gumbel, sua densidade é dada por

$$f(x; \phi, \varphi) = \frac{1}{\varphi} \exp \left(-\frac{x - \phi}{\varphi} - \exp \left(-\frac{x - \phi}{\varphi} \right) \right) \quad (2.1)$$

Para $x \in (-\infty, \infty)$, onde $\phi \in (-\infty, \infty)$ é o parâmetro de locação e $\varphi > 0$ o de escala. A Figura 2.1 mostra a distribuição gumbel para diferentes parâmetros.

Figura 2.1: Densidade da Distribuição Gumbel para diferentes valores dos parâmetros ϕ e φ



A seguir apresenta-se as propriedades e características da distribuição de Gumbel.

2.1 Propriedades

P1: A função distribuição acumulada gumbel é dado por:

$$F(x|\phi, \varphi) = \exp \left(- \exp \left(- \frac{x - \phi}{\varphi} \right) \right). \quad (2.2)$$

P2: Esperança

$$E[X] = \phi + \varphi\gamma. \quad (2.3)$$

Onde $\gamma = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n \frac{1}{k} - \ln(n) \right)$ é a constante de Euler-Mascheroni.

P3: Variância

$$Var[X] = \frac{\pi^2}{6} \varphi^2. \quad (2.4)$$

P4: Os momentos de X são encontrados derivando $M_X(t)$ e igualando t a zero.

$$M_X(t) = \Gamma(1 - \varphi t) e^{\phi t}. \quad (2.5)$$

2.2 Estimadores de máxima verossimilhança

O estimador de máxima verossimilhança (EMV) busca os parâmetros que maximizam a probabilidade da amostra aleatória correr.

Considere uma amostra aleatória independente e identicamente distribuída (i.i.d.) $(x_1, x_2, \dots, x_n)$, em que cada elemento da amostra segue uma distribuição de probabilidade $f_X(x | \theta)$. Devido à independência dos elementos da amostra, a distribuição de probabilidade conjunta é dada por:

$$f_X(x_1, x_2, \dots, x_n | \theta) = f_X(x_1 | \theta) f_X(x_2 | \theta) \dots f_X(x_n | \theta) = \prod_{i=1}^n f(x_i | \theta).$$

Ao variar o parâmetro θ , a função conjunta $f_X(x_1, x_2, \dots, x_n | \theta)$ fornece a probabilidade (ou, no caso contínuo, a densidade) de observação da amostra aleatória para cada valor de θ . A função $\mathcal{L}(\theta|x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i | \theta)$ é denominada *função de verossimilhança*. O EMV é obtido ao maximizar essa função em relação ao parâmetro θ .

Na prática, é comum trabalhar com log da função de verossimilhança, em vez de utilizar diretamente a função de verossimilhança. Essa função facilita o processo de derivação, permitindo encontrar o ponto de máximo de forma mais simples. Isso é possível porque o logaritmo é uma função monótona crescente, o que implica que o ponto de máximo da função de verossimilhança será o mesmo da função logaritimizada.

Para a distribuição Gumbel a verossimilhança é

$$\mathcal{L}(\phi, \varphi|x) = \frac{1}{\varphi^n} \exp \left\{ - \frac{n}{\varphi} (\bar{x} - \phi) - \sum_{i=1}^n \exp \left(- \frac{x_i - \phi}{\varphi} \right) \right\} \quad (2.6)$$

onde $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. Obtendo as derivadas de primeira ordem e igualando a zero para encontrar o ponto de máximo, temos que os EMV são dados por

$$\begin{cases} \hat{\phi} &= \hat{\varphi} \log \left[\frac{n}{\sum_{i=1}^n \exp \left(-\frac{x_i}{\hat{\varphi}} \right)} \right] \\ \hat{\varphi} &= \bar{x} - \frac{\sum_{i=1}^n x_i \exp \left(-\frac{x_i}{\hat{\varphi}} \right)}{\sum_{i=1}^n \exp \left(-\frac{x_i}{\hat{\varphi}} \right)} \end{cases} \quad (2.7)$$

Observe que para encontrar os estimadores é necessário recorrer a métodos numéricos como método de newton-raphson.

2.3 Geração de amostras aleatórias da distribuição Gumbel

Muitos problemas práticos não possuem soluções analíticas exatas ou são excessivamente complexos para serem resolvidos de forma direta. A simulação utiliza amostras aleatórias que permitem manipular distribuições de maneira eficiente e prática, facilitando a análise de problemas que, de outra forma, seriam intratáveis ou muito difíceis de abordar por métodos analíticos.

Uma técnica amplamente utilizada para a geração de valores aleatórios é o *método da transformação inversa*. Esse método parte de amostras aleatórias uniformemente distribuídas no intervalo $(0, 1)$ e utiliza a função de distribuição acumulada (FDA) correspondente para gerar valores da distribuição desejada.

Para obter um valor X utilizando essa técnica, é necessário determinar a *inversa* da função de distribuição acumulada. A partir de uma amostra aleatória U , proveniente de uma distribuição uniforme no intervalo $(0, 1)$, aplica-se a inversa da FDA para obter os valores desejados de X .

No caso específico da distribuição Gumbel, o gerador de amostras aleatórias é dado pela seguinte expressão:

$$X = \phi - \varphi \log(-\log U), \quad (2.8)$$

onde U é uma amostra uniforme no intervalo $(0, 1)$

Capítulo 3

Análise Bayesiana

A principal diferença entre a estatística Bayesiana e clássica está na forma como se trata os parâmetros. Na inferência clássica são considerados fixos para a população, enquanto na inferência Bayesiana são assumidas como variáveis aleatórias, pois como não sabemos nada sobre o parâmetro, nada mais justo que tratá-lo como uma variável aleatória. Essa abordagem altera a maneira como as inferências são realizadas, passando a estudar a distribuição $f(\theta|x)$ em vez de $f(x|\theta)$, como é feito na inferência clássica. Em outras palavras, a inferência Bayesiana se concentra na probabilidade dos parâmetros condicionados aos dados observados. Essa mudança de perspectiva é fundamental na inferência Bayesiana, permitindo incorporar o conhecimento prévio as estimativas dos parâmetros com base nos dados observados.

A inferência Bayesiana, fundamentada no Teorema de Bayes, permite atualizar o conhecimento ou a incerteza inicial sobre os parâmetros com base nos dados observados. Esse processo combina o conhecimento prévio, representado pela distribuição a priori, com os dados observados, por meio da função de verossimilhança, para obter a distribuição a posteriori, definida como:

$$p(\boldsymbol{\theta} \mid \mathbf{x}) = \frac{\mathcal{L}(\mathbf{x} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\Theta} \mathcal{L}(\mathbf{x} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \quad (3.1)$$

Onde $\pi(\boldsymbol{\theta})$ é a distribuição a priori e $\mathcal{L}(\mathbf{x} \mid \boldsymbol{\theta})$ a função de verossimilhança. Observe que o denominador é uma integral em relação a θ , temos que a função integrada é uma constante. Dessa forma a distribuição de probabilidade a posteriori é proporcional ao produto da priori pela verossimilhança, conforme indicado a seguir:

$$p(\theta|x) \propto \mathcal{L}(\mathbf{x} \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta}) \quad (3.2)$$

A partir da distribuição a posteriori, é possível calcular os intervalos de credibilidade e estimadores. Um exemplo comum de estimador Bayesiano é dado pela esperança posteriori, calculada como:

$$\hat{\theta} = E[p(\boldsymbol{\theta} \mid \mathbf{x})] \quad (3.3)$$

Na análise Bayesiana, é essencial especificar distribuições a priori para os parâmetros de interesse, sendo que o domínio da distribuição deve ser igual ao do parâmetro.

Existem diversas maneiras de escolher a distribuição a priori, que podem ser classificadas em informativas e não informativas. As distribuições informativas requerem o conhecimento de um especialista no assunto, enquanto as distribuições não informativas, mais comumente usadas, não dependem desse conhecimento prévio. Entre as principais distribuições a priori não informativas, destacam-se as distribuições de Jeffrey e as distribuições Flat, que são caracterizadas por um alto grau de variabilidade.

Neste capítulo serão apresentados algumas priors informativas e não-informativas para a distribuição Gumbel

3.1 Priori de Jeffreys

Uma das priors não informativas mais utilizadas na inferência Bayesiana é a priori de Jeffreys, proposta por (JEFFREYS, 1961). Essa distribuição é obtida a partir da matriz de informação de Fisher e é definida como:

$$\pi(\boldsymbol{\theta}) \propto \sqrt{\det \mathcal{I}(\boldsymbol{\theta})}.$$

Para a distribuição Gumbel, a matriz de informação de Fisher $\mathcal{I}(\phi, \varphi)$ é dada por:

$$\mathcal{I}(\phi, \varphi) = \begin{bmatrix} \frac{1}{\varphi^2} & -\frac{1}{\varphi^2}[1 + \psi(1)] \\ -\frac{1}{\varphi^2}[1 + \psi(1)] & \frac{1}{\varphi^2}[1 + \psi'(2) + \psi^2(2)] \end{bmatrix}$$

onde $\psi(k)$ é a função digama, definida como a derivada logarítmica da função gama, ou seja, $\psi(k) = \frac{d}{dk} \log \Gamma(k)$.

Os elementos da matriz de informação de Fisher são definidos como $\mathcal{I}_{ij}(\theta) = -E \left(\frac{\partial^2 \log f(x|\boldsymbol{\theta})}{\partial \theta_i \partial \theta_j} \right)$.

Portanto, a priori de Jeffreys para os parâmetros ϕ e φ da distribuição Gumbel é dada por:

$$\pi_J(\phi, \varphi) = \frac{1}{\varphi^2}. \quad (3.4)$$

A solução obtida por aplicações direta da regra de jeffreys multiparamétrica costuma ter implicações pouco satisfatórias que são frequentemente ultrapassadas por imposição de independência a priori entre os parâmetros de naturezas diferentes (parâmetros de localização e escala)

3.2 Produto de priors independentes

Como mencionado em 3.1 impor a condição de independência entre as priors resulta em uma abordagem mais eficaz do que a priori de Jeffreys, já que facilita a determinação das priors individuais de cada parâmetro, em comparação com o uso de uma priori conjunta.

Para o caso da distribuição Gumbel serão consideradas as seguintes distribuições a priori:

$$\phi \sim Normal(\mu, \sigma^2) \quad e \quad \varphi \sim Gama(\alpha, \beta) \quad (3.5)$$

E por independência temos $\pi(\phi, \varphi) = \pi_\phi(\phi|\mu, \sigma^2)\pi_\varphi(\varphi|\alpha, \beta)$. Dessa forma obtemos a posteriori que é proporcional ao produto das prioris 3.5 e verossimilhança 2.6

$$p(\phi, \varphi|x) \propto \varphi^{\alpha-1-n} \exp \left\{ -\frac{n}{\varphi}(\bar{x} - \phi) - \sum_{i=1}^n \exp \left(\frac{\phi - x_i}{\varphi} \right) - \beta\varphi - \frac{(\phi - \mu)^2}{2\sigma^2} \right\}. \quad (3.6)$$

Uma observação importante é que cada combinação de parâmetros da distribuição a priori resulta em uma distribuição a posteriori distinta, refletindo diferentes suposições iniciais sobre os parâmetros. Assim, apresentaremos, a seguir, uma distribuição a priori não informativa (Flat) para o caso de prioris independentes, bem como uma distribuição a priori informativa (Elicitada), derivada do método de Laplace, que será detalhado no Capítulo 4.

3.2.1 Priori Flat

Uma forma de representar o estado de "pouca informação" é aumentar expressivamente a variância da distribuição a priori escolhida. Esse tipo de abordagem é conhecida como priori flat, é usada quando não há razão para privilegiar algum valor em particular, dada o estado de ignorância. Ao adotar uma priori flat, o objetivo é minimizar a influência do conhecimento prévio na análise, assegurando que os dados observados desempenhem um papel predominante na construção da distribuição a posteriori. Para alcançar esse objetivo, escolhemos parâmetros que ampliem de forma significativa a variância da distribuição a priori. Assim temos:

$$\pi_F(\phi, \varphi) = \pi_\phi(\phi|0, 1000^2)\pi_\varphi(\varphi|0.001, 0.001) \quad (3.7)$$

3.2.2 Priori Elicitada I

O ajuste dos parâmetros da priori pode resultar em estimadores mais precisos; no entanto, esse processo é delicado, pois valores incorretos podem introduzir viés ou comprometer a análise. Neste contexto, apresentamos a priori elicitada a partir de um sistema de quantis, no qual consideramos dois cenários distintos. A metodologia utilizada para a determinação dos parâmetros da distribuição de probabilidade a priori será descrita detalhadamente no Capítulo 4.

No primeiro cenário, o especialista comete um erro de 10% a menos em relação ao quantil esperado, ou seja, o quantil fornecido pelo especialista apresenta uma discrepância de 10% para menos em relação ao valor verdadeiro. A distribuição de probabilidade a priori associada a esse erro é dada por:

$$\pi_{EI}(\phi, \varphi) = \pi_\phi(\phi|0.91802, 0.4511^2)\pi_\varphi(\varphi|8.3526854, 4.2140159) \quad (3.8)$$

3.2.3 Priori Elicitada II

Em um segundo cenário, o erro cometido pelo especialista é de 10% a menos, mas, desta vez, em relação à probabilidade. Por exemplo, foi solicitado ao especialista que fornecesse a probabilidade correspondente a até 25% dos casos, mas ele acabou fornecendo o quantil correspondente a apenas 15%. Nesse caso, a distribuição de probabilidade a priori é dada por:

$$\pi_{EII}(\phi, \varphi) = \pi_{\phi}(\phi|0.7400815, 1.0029117^2)\pi_{\varphi}(\varphi|6.5492880, 2.5181159) \quad (3.9)$$

A metodologia utilizada para determinar os parâmetros das prioris elicítadas será detalhada no próximo capítulo.

Capítulo 4

Fundamentos da priori de Laplace

Neste capítulo, apresentamos a metodologia para determinar os hiperparâmetros de uma priori, utilizando um sistema de quantis com equações derivadas de soluções aproximadas. Para isso, introduzimos o processo de elicitacão e a integraão por meio da aproximaão de laplace.

4.1 Elicitacão

O processo de elicitacão incorpora conhecimento do especialista nas análises, ampliando a capacidade de realizar inferências estatísticas mais profundas e precisas. Sua aplicabilidade torna-se particularmente notável em cenários com limitações de dados.

Isso não só complementa os dados existentes, promovendo adaptabilidade e enriquecimento analítico através das experiências e percepções dos especialistas, mas também traz desafios significativos. A subjetividade inerente aos julgamentos de especialistas pode introduzir viés e variabilidade, enquanto a transformação de opiniões qualitativas em estimativas quantitativas confiáveis apresentam complexidades. Além disso, em situações com múltiplas perspectivas especializadas, o consenso pode ser difícil de alcançar. Assim, a elicitacão na análise bayesiana representa um equilíbrio delicado entre aproveitar o conhecimento humano profundo e gerenciar as limitações que este conhecimento pode trazer.

Segundo (O'HAGAN, 1998), define-se o *método estrutural* de elicitacão como qualquer técnica que se concentre em elaborar questões específicas relacionadas a um parâmetro θ , que faz parte de um conjunto Θ . A especificidade e complexidade destas questões são determinadas pelo tipo do parâmetro a ser eliciada e pelo conhecimento do especialista envolvido. A efetividade deste método depende de dois fatores principais:

1. Formular perguntas em uma linguagem clara e acessível ao especialista, garantindo que sejam compreensíveis e pertinentes.
2. Identificar e focar nas principais fontes de incerteza associadas ao parâmetro, tratando-as de maneira isolada para uma análise mais precisa.

Uma abordagem prática e simples para elicitacão envolve questionar os quantis do modelo. Por exemplo, em um problema relacionado à volumetria de chuva mensal,

pode-se perguntar: 'Qual é a probabilidade de chover menos de 500 litros este mês?'. Esta técnica facilita o processo ao evitar a necessidade de conhecimento aprofundado em estatística ou probabilidade por parte do especialista. Entretanto, surge o desafio de como integrar esses quantis na distribuição a priori, o que será discutido na próxima seção.

4.2 Metodologia de Elicitação Usando Aproximação de Laplace

Neste capítulo, formularemos o processo de elicitación de parâmetros para um modelo estatístico no contexto bayesiano. Consideremos que a distribuição de probabilidade $f_X(x|\boldsymbol{\theta})$ que modela os dados, onde $\boldsymbol{\theta}$ são os parâmetros do modelo, e $\pi(\boldsymbol{\theta}|\boldsymbol{\alpha})$ a distribuição de probabilidade a priori, onde $\boldsymbol{\alpha}$ são os hiperparâmetros da distribuição de probabilidade a posteriori $p(\boldsymbol{\theta}|x)$. Neste processo de elicitación, considerando a informação de um especialista como fonte de dados, procuramos determinar o quantil da distribuição. O cálculo do quantil pode ser expresso pela integral:

$$P(X \leq x_p) = \int_0^\infty \cdots \int_0^\infty \left[\int_0^{x_p} f_X(x|\boldsymbol{\theta}) dx \right] \pi(\boldsymbol{\theta}|\boldsymbol{\alpha}) d\boldsymbol{\theta}, \quad \boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$$

Uma dificuldade surge quando a integração não possui solução analítica ou é complexa para calcular. Para contornar isso, utilizamos a aproximação de Laplace.

A aproximação de Laplace é aplicada quando não é possível obter uma forma fechada para a integral. Suponha que h seja uma função suave de um parâmetro unidimensional θ , e $-h$ tenha um máximo em $\hat{\theta}$. A abordagem de Laplace aproxima assintoticamente uma integral da forma

$$I = \int \exp(-nh(\theta)) d\theta$$

expandindo h em uma série de Taylor em torno de $\hat{\theta}$ (TIERNEY; KADANE, 1986) O método de Laplace fornece a aproximação

$$\int_0^\infty \cdots \int_0^\infty \exp(-nh(\boldsymbol{\theta})) d\boldsymbol{\theta} \approx \left(\frac{2\pi}{n} \right)^{\frac{m}{2}} \left| \boldsymbol{\Sigma}^{-1} \right|^{1/2} \exp(-nh(\hat{\boldsymbol{\theta}})) \quad (4.1)$$

Onde $\hat{\boldsymbol{\theta}}$ é o ponto de máximo da função h , m a dimensão do parâmetro e $|\boldsymbol{\Sigma}^{-1}|$ é o determinante inversa da matriz hessiana de $h(\hat{\boldsymbol{\theta}})$. A matriz hessiana, por sua vez, corresponde à derivada de segunda ordem da função h . Assim, temos:

$$\begin{aligned}
P(X \leq x_p) &= \int_0^\infty \cdots \int_0^\infty \left[\int_0^{x_p} f_X(x|\boldsymbol{\theta}) dx \right] \pi(\boldsymbol{\theta}|\boldsymbol{\alpha}) d\boldsymbol{\theta} \\
&= \int_0^\infty \cdots \int_0^\infty F_X(x_p|\boldsymbol{\theta}) \pi(\boldsymbol{\theta}|\boldsymbol{\alpha}) d\boldsymbol{\theta} \\
&= \int_0^\infty \cdots \int_0^\infty \exp(-h(\boldsymbol{\theta}|x_p, \boldsymbol{\alpha})) d\boldsymbol{\theta} \\
&\approx (2\pi)^{\frac{m}{2}} \left| \sum^{-1} \right|^{1/2} \exp\left(-h(\hat{\boldsymbol{\theta}}|x_p, \boldsymbol{\alpha})\right)
\end{aligned}$$

onde $h(\hat{\boldsymbol{\theta}}|x_p, \boldsymbol{\alpha}) = -\log(F_X(x_p|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\boldsymbol{\alpha}))$

Para determinar os valores dos hiperparâmetros, $\boldsymbol{\alpha}$, da priori pedimos ao especialista o mesmo número de quantis. Assim, obtemos o seguinte sistema de quantis:

$$\begin{cases} p_1 = (2\pi)^{\frac{m}{2}} \left| \sum^{-1} \right|^{1/2} \exp\left(-h(\hat{\boldsymbol{\theta}}|x_{p_1}, \boldsymbol{\alpha})\right) \\ p_2 = (2\pi)^{\frac{m}{2}} \left| \sum^{-1} \right|^{1/2} \exp\left(-h(\hat{\boldsymbol{\theta}}|x_{p_2}, \boldsymbol{\alpha})\right) \\ \vdots \\ p_n = (2\pi)^{\frac{m}{2}} \left| \sum^{-1} \right|^{1/2} \exp\left(-h(\hat{\boldsymbol{\theta}}|x_{p_n}, \boldsymbol{\alpha})\right) \end{cases} \quad (4.2)$$

Resolvendo este sistema, obtemos os hiperparâmetros, $\boldsymbol{\alpha} = \alpha_1, \alpha_2, \dots, \alpha_n$, da distribuição a priori $\pi(\boldsymbol{\theta}|\boldsymbol{\alpha})$.

4.3 Aplicação a Distribuição de Gumbel

Consideraremos o mesmo problema do capítulo 3, com algumas adaptações para facilitar a compreensão e a notação. Em particular, substituímos os símbolos ϕ e φ pelos mais intuitivos y e z , respectivamente. A formulação do problema pode ser representada pela seguinte integral:

$$P(X < x) = \int_0^\infty \int_{-\infty}^\infty F_X(x|y, z) f_Y(y|\mu, \sigma^2) f_Z(z|\alpha, \beta) dy dz,$$

onde as distribuições envolvidas são: $X \sim \text{Gumbel}(Y, Z)$, $Y \sim \text{Normal}(\mu, \sigma^2)$ e $Z \sim \text{Gamma}(\alpha, \beta)$. Para transformar o domínio de f_Y para $(0, \infty)$ aplicamos a mudança de variável em y descrita a seguir:

$$\begin{aligned}
P(X < x) &= \int_0^\infty \int_{-\infty}^\infty F_X(x|y, z) f_Y(y|\mu, \sigma) f_Z(z|\alpha, \beta) dy dz \\
&= \int_0^\infty \left[\int_{-\infty}^\infty F_X(x|y, z) f_Y(y|\mu, \sigma) dy \right] f_Z(z|\alpha, \beta) dz \\
&= \int_0^\infty E[F_X(x|Y, z)] f_Z(z|\alpha, \beta) dz \\
&= \int_0^\infty E \left[\exp \left(- \exp \left(- \frac{x - Y}{z} \right) \right) \right] f_Z(z|\alpha, \beta) dz.
\end{aligned}$$

Realizando a seguinte transformação da variável Y :

$$J = \exp\left(-\frac{x - Y}{z}\right) \sim \text{Lognormal}\left(\frac{\mu - x}{z}, \frac{\sigma^2}{z^2}\right),$$

Logo:

$$\begin{aligned} P(X < x) &= \int_0^\infty E[\exp(-J)] f_Z(z|\alpha, \beta) dz \\ &= \int_0^\infty \left[\int_0^\infty e^{-j} f_J(j|\mu_j, \sigma_j^2) dj \right] f_Z(z) dz \\ &= \int_0^\infty \int_0^\infty e^{-j} f_J(j|\mu_j, \sigma_j^2) f_Z(z|\alpha, \beta) dj dz. \end{aligned}$$

Manipulando a função dentro integral, temos:

$$\begin{aligned} P(X < x) &= \int_0^\infty \int_0^\infty e^{-j} f_J(j|\mu_j, \sigma_j^2) f_Z(z|\alpha, \beta) dj dz \\ &= \int_0^\infty \int_0^\infty \exp(-j + \log(f_J(j|\mu_j, \sigma_j^2)) + \log(f_Z(z|\alpha, \beta))) dj dz \\ &= \int_0^\infty \int_0^\infty \exp(-h(j, z|\mu_j, \sigma_j^2, \alpha, \beta)) dj dz, \end{aligned}$$

onde $h(j, z) = j - \log(f_J) - \log(f_Z)$, $\mu_j = \frac{\mu - x}{z}$ e $\sigma_j = \frac{\sigma}{z}$. No método de Laplace, as derivadas de primeira e segunda ordem da função são calculadas. A derivada de primeira ordem é utilizada para encontrar o ponto de máximo da função h , enquanto a derivada de segunda ordem é empregada para obter a matriz hessiana.

O ponto de máximo $\hat{j}$ e $\hat{z}$ é dado pela raiz do seguinte sistema

$$\begin{cases} \frac{\partial h}{\partial j} = 0 \\ \frac{\partial h}{\partial z} = 0 \end{cases}$$

$$\begin{cases} 1 + \frac{1}{\hat{j}\sigma^2} (\hat{z}^2 \log(\hat{j}) + \sigma^2 + x\hat{z} - \mu\hat{z}) = 0 \\ \beta - \frac{\alpha}{\hat{z}} + \frac{\log(\hat{j})}{\sigma^2} (\hat{z} \log(\hat{j} - \mu + x)) = 0 \end{cases}$$

O sistema em questão é não linear e apresenta elevada complexidade para solução analítica. Uma alternativa viável para resolver esse tipo de sistema é a aplicação de métodos numéricos, como o método de Newton-Raphson. Diversas linguagens de programação oferecem ferramentas para implementação desses métodos. Neste trabalho, optou-se pela utilização da linguagem de programação **R**, será utilizada a função **optim**, disponível no pacote **stats**, para a obtenção dos pontos de máximo e mínimo.

A matriz hessiana é dada por

$$\Sigma = \begin{bmatrix} \frac{\partial^2 h}{\partial j^2} & \frac{\partial^2 h}{\partial j \partial z} \\ \frac{\partial^2 h}{\partial z \partial j} & \frac{\partial^2 h}{\partial z^2} \end{bmatrix},$$

onde:

$$\frac{\partial^2 h}{\partial j^2} = -\frac{z^2 \log(j) + \sigma^2 + xz - z^2 - z\mu}{j^2 \sigma^2}$$

$$\frac{\partial^2 h}{\partial z^2} = \frac{\log^2(j)}{\sigma^2} + \frac{\alpha}{z^2}$$

$$\frac{\partial^2 h}{\partial j \partial z} = \frac{\partial^2 h}{\partial z \partial j} = -\frac{2z \log(j) - \mu + x}{j \sigma^2}$$

Finalmente, apresentamos o seguinte sistema de quantis. No lado esquerdo, encontram-se as probabilidades (0.25, 0.50, 0.75, 0.90), solicitadas pelo estatístico ao especialista. No lado direito, estão os quantis correspondentes, fornecidos pelo especialista e associados a essas probabilidades:

$$\left\{ \begin{array}{l} 0.25 = 2\pi \frac{\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\hat{z}^\alpha}{j\sqrt{2\pi}\sigma} \exp\left(-\frac{(\log(\hat{j}) - \frac{x_{0.25} - \mu}{\hat{z}})^2}{2\frac{\sigma^2}{\hat{z}^2}} - \beta\hat{z}\right)}{\sqrt{\left| -\frac{\hat{z}^2 \log(\hat{j}) + \sigma^2 + x_{0.25}\hat{z} - \hat{z}^2 - \hat{z}\mu}{j^2 \sigma^2} \left(\frac{\log^2(\hat{j})}{\sigma^2} + \frac{\alpha}{\hat{z}^2}\right) - \left(\frac{2\hat{z} \log(\hat{j}) - \mu + x_{0.25}}{j \sigma^2}\right)^2 \right|}} \\ 0.50 = 2\pi \frac{\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\hat{z}^\alpha}{j\sqrt{2\pi}\sigma} \exp\left(-\frac{(\log(\hat{j}) - \frac{x_{0.50} - \mu}{\hat{z}})^2}{2\frac{\sigma^2}{\hat{z}^2}} - \beta\hat{z}\right)}{\sqrt{\left| -\frac{\hat{z}^2 \log(\hat{j}) + \sigma^2 + x_{0.50}\hat{z} - \hat{z}^2 - \hat{z}\mu}{j^2 \sigma^2} \left(\frac{\log^2(\hat{j})}{\sigma^2} + \frac{\alpha}{\hat{z}^2}\right) - \left(\frac{2\hat{z} \log(\hat{j}) - \mu + x_{0.50}}{j \sigma^2}\right)^2 \right|}} \\ 0.75 = 2\pi \frac{\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\hat{z}^\alpha}{j\sqrt{2\pi}\sigma} \exp\left(-\frac{(\log(\hat{j}) - \frac{x_{0.75} - \mu}{\hat{z}})^2}{2\frac{\sigma^2}{\hat{z}^2}} - \beta\hat{z}\right)}{\sqrt{\left| -\frac{\hat{z}^2 \log(\hat{j}) + \sigma^2 + x_{0.75}\hat{z} - \hat{z}^2 - \hat{z}\mu}{j^2 \sigma^2} \left(\frac{\log^2(\hat{j})}{\sigma^2} + \frac{\alpha}{\hat{z}^2}\right) - \left(\frac{2\hat{z} \log(\hat{j}) - \mu + x_{0.75}}{j \sigma^2}\right)^2 \right|}} \\ 0.90 = 2\pi \frac{\frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\hat{z}^\alpha}{j\sqrt{2\pi}\sigma} \exp\left(-\frac{(\log(\hat{j}) - \frac{x_{0.90} - \mu}{\hat{z}})^2}{2\frac{\sigma^2}{\hat{z}^2}} - \beta\hat{z}\right)}{\sqrt{\left| -\frac{\hat{z}^2 \log(\hat{j}) + \sigma^2 + x_{0.90}\hat{z} - \hat{z}^2 - \hat{z}\mu}{j^2 \sigma^2} \left(\frac{\log^2(\hat{j})}{\sigma^2} + \frac{\alpha}{\hat{z}^2}\right) - \left(\frac{2\hat{z} \log(\hat{j}) - \mu + x_{0.90}}{j \sigma^2}\right)^2 \right|}} \end{array} \right.$$

Convem notar que $\hat{j}$ e $\hat{z}$ são os pontos de máximo de h e estão em função dos hiperparâmetros. O sistema apresentado é não linear e, para resolver esse tipo de problema, será utilizado o método numérico de Newton-Raphson. A implementação será realizada com a função **nleqslv**, disponível no pacote **nleqslv** da linguagem de programação **R**. Prosseguimos com a análise.

Suponha que, em um processo de elicitação, solicitamos a um especialista o quantil correspondente a probabilidade 0.25. O especialista, sujeito a erros, fornece o valor 0.3121, enquanto o valor verdadeiro seria 0.3467. Isso representa um erro relativo de aproximadamente menos 10%, calculado como $\frac{\text{valor}_{\text{especialista}} - \text{valor}_{\text{verdadeiro}}}{\text{valor}_{\text{verdadeiro}}} = \frac{0.3121 - 0.3467}{0.3467} = -0.1$. Este processo será denominado *Elicitação I*.

Em um cenário alternativo, em vez de errar em relação ao quantil, o especialista pode cometer um erro em relação à probabilidade. Isso significa que ele fornecerá o quantil correspondente a uma probabilidade que está 10% abaixo da probabilidade solicitada. Por exemplo, ao ser solicitado o quantil para uma probabilidade de 0.25, o especialista fornece

o quantil correspondente à probabilidade de 0.15, cometendo um erro de -10% em relação à probabilidade. Este processo será denominado *Elicitação II*.

A seguir, simularemos os quantis obtidos por um especialista para uma distribuição de probabilidade Gumbel com parâmetros $\phi = 1$ e $\varphi = 2$, aplicando o erro do processo de elicitación I e II temos:

Tabela 4.1: Quantis simulados obtidos por um especialista

Probabilidade (p)	Quantil Verdadeiro (x_p)	Quantil Esp. I (x_p)	Quantil Esp. II (x_p)
0.25	0.3467	0.3121	-0.2806
0.50	1.7330	1.5597	1.1748
0.75	3.4918	3.1426	2.6843
0.90	5.5007	4.9507	3.9998

Substituindo os valores ($x_{0.25}, x_{0.50}, x_{0.75}, x_{0.90}$) no sistema pelos quantis fornecidos pelo especialista no processo de *Elicitação I* (0.3121, 1.5597, 3.1426, 4.9507) e resolvendo o sistema por métodos numéricos utilizando a linguagem de programação R (o código pode ser encontrado no apêndice B.2), obtemos os parâmetros para a distribuição de probabilidade a priori dada em 3.6. da mesma forma obtivemos os parâmetros para o processo *Elicitação II*. Os resultados obtidos são apresentados na tabela a seguir.

Tabela 4.2: Hiperparâmetros obtidos resolvendo o sistema de quantis

Especialista	α	β	μ	σ
I	8.3527	4.2140	0.9180	0.4511
II	6.5620	2.5114	0.7650	0.9821

Como o sistema foi resolvido utilizando métodos numéricos, o erro absoluto total foi de 0.9111 para *Elicitação I* e 0.2294 para *Elicitação II*. Esse erro foi calculado somando os resultados de cada equação do sistema de quantis, substituindo os hiperparâmetros pelos valores apresentados na tabela 4.2.

Capítulo 5

Aplicação dos modelos

Neste capítulo, analisaremos os modelos associados a cada uma das priors descritas no capítulo 3. Primeiramente, apresentaremos o algoritmo MCMC para calcular as distribuições posteriores. Em seguida, exploraremos os modelos gerados e, por fim, investigaremos o comportamento dos estimadores, dados pelo modelos, em função do tamanho das amostras.

5.1 Método de Monte Carlo via Cadeia de Markov

Como, muitas vezes, a posteriori não apresenta fórmula fechada conhecida, isso dificulta muito a análise, já que a inferência bayesiana gira entorno dela. Essa dificuldade foi contornada com o surgimento do método de monte carlo via cadeia de markov (MCMC), que gera uma amostra da distribuição posteriori. Isso permite realizar inferências e análises a partir dela e da distribuição posteriori dos parâmetros.

5.1.1 Algoritmo para gerar ϕ e φ

A seguir, será detalhada a aplicação do MCMC para a distribuição Gumbel, utilizando priors independentes Normal e Gamma, conforme descrito no capítulo 3.2. Nesse contexto, $p(\phi, \varphi|x)$ representa o produto da verossimilhança e da distribuição de probabilidade a priori, enquanto q_ϕ e q_φ são as geradoras da proposta para os parâmetros ϕ e φ , respectivamente.

1. Determine uma quantidade n de amostras.
2. Inicialize os valores de $\phi[0]$ e $\varphi[0]$.
3. Gere um candidato para $\phi[i+1]$ a partir de uma distribuição normal com média $\phi[i]$ e variância σ^2 : $\phi[i+1] \sim \text{Normal}(\phi[i], \sigma^2)$.
4. Calcule o valor de α :

$$\alpha = \min \left\{ 1, \frac{p(\phi[i+1], \varphi[i]|x)}{p(\phi[i], \varphi[i]|x)} \cdot \frac{q_\phi(\phi[i]|\mu = \phi[i+1], \sigma^2)}{q_\phi(\phi[i+1]|\mu = \phi[i], \sigma^2)} \right\}$$

5. Gere uma variável uniforme U no intervalo $[0, 1]$.
6. Se $\alpha > U$, aceite $\phi[i + 1]$, caso contrário, rejeite o valor gerado e faça $\phi[i + 1] = \phi[i]$.
7. Gere um candidato para $\varphi[i + 1]$ a partir de uma distribuição gama com forma $\varphi[i]$ e parâmetro β : $\varphi[i + 1] \sim \text{Gamma}(\varphi[i], \beta)$.
8. Calcule o valor de α :

$$\alpha = \min \left\{ 1, \frac{p(\phi[i + 1], \varphi[i + 1]|x)}{p(\phi[i + 1], \varphi[i]|x)} \cdot \frac{q_\varphi(\varphi[i]|\alpha = \varphi[i + 1], \beta)}{q_\varphi(\phi[i + 1]|\alpha = \varphi[i], \beta)} \right\}$$

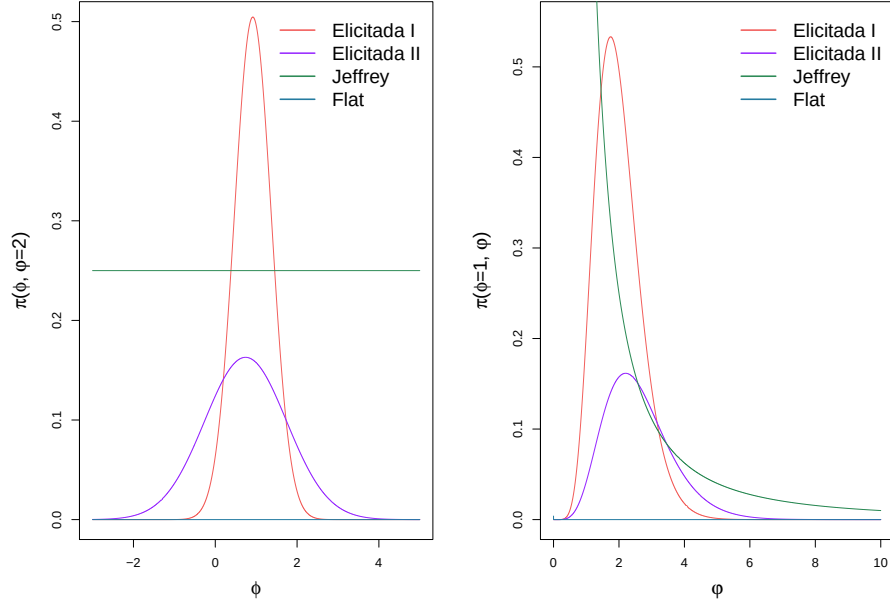
9. Gere uma variável uniforme U no intervalo $[0, 1]$.
10. Se $\alpha > U$, aceite $\varphi[i + 1]$, caso contrário, rejeite o valor gerado e faça $\varphi[i + 1] = \varphi[i]$.
11. Repita o passo 2 até gerar n quantidades de amostras.

Algumas técnicas e estratégias de monitoramento são discutidos em (RAFTERY; LEWIS, 1992). Recomenda-se que a taxa de aceitação dos valores gerado esteja em torno de 30%. Após gerar as amostras, é prática comum excluir uma fração inicial dos valores gerados, conhecida como burn-in. Isso ocorre porque a cadeia começa a partir de um ponto inicial arbitrário (semente) no espaço de parâmetros, que pode estar distante da região de alta densidade da distribuição a posteriori. Durante as primeiras iterações, a cadeia passa por um período de ajuste e convergência para a distribuição alvo. Nesse intervalo, os valores iniciais não são representativos da distribuição a posteriori e, caso incluídos na análise, podem enviesar estimativas estatísticas, como médias ou variâncias. O algoritmo descrito foi implementado na linguagem de programação R e está disponível no apêndice B.1.

5.2 Distribuições de Probabilidade a Posteriori

Será analisada uma cadeia MCMC de tamanho 1.000.000 para cada uma das priors discutidas no Capítulo 3. O objetivo desta seção é verificar se o algoritmo está funcionando corretamente e não apresenta nenhuma anormalidade.

Figura 5.1: Distribuições de Probabilidade a priori



Uma amostra da distribuição de Gumbel, com parâmetros $\phi = 1$ e $\varphi = 2$, de tamanho 3 foi gerada a partir da semente 20192024. Os valores gerados foram 1.5278848, 1.8842654 e 5.5493339

No caso da Priori Elicitada I, observa-se que a autocorrelação estabiliza a partir de um lag de tamanho 20. Além disso, as distribuições marginais indicam que essa priori apresenta uma variância menor, com o pico de maior densidade centrado próximo ao valor verdadeiro. Já a Priori Elicitada II exibe características semelhantes à Priori Elicitada I, diferenciando-se apenas por apresentar um maior grau de variância.

Figura 5.2: Cadeia MCMC da distribuição posterior com distribuição prior Elicitada I para uma amostra de tamanho $n = 3$.

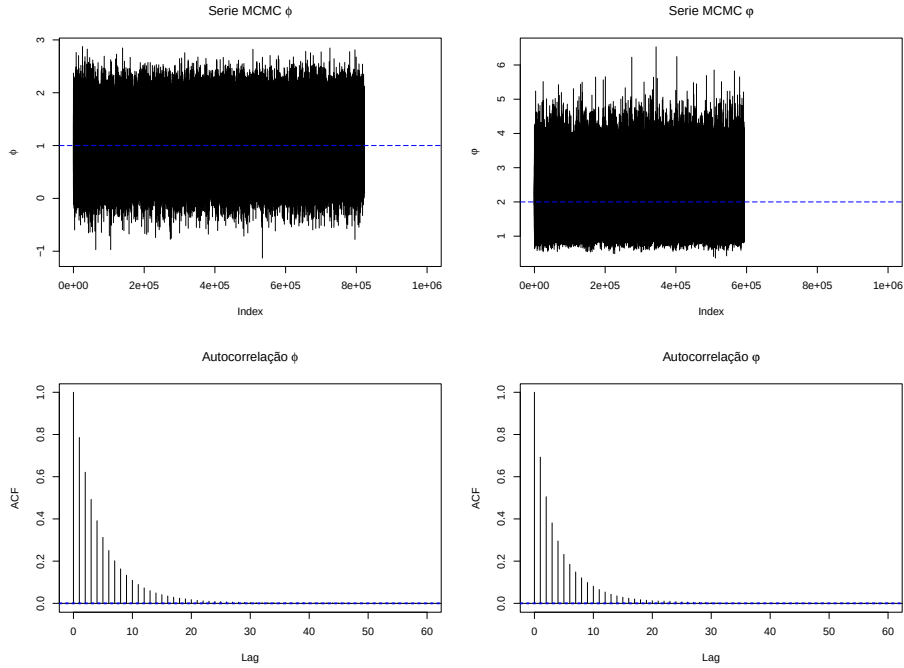
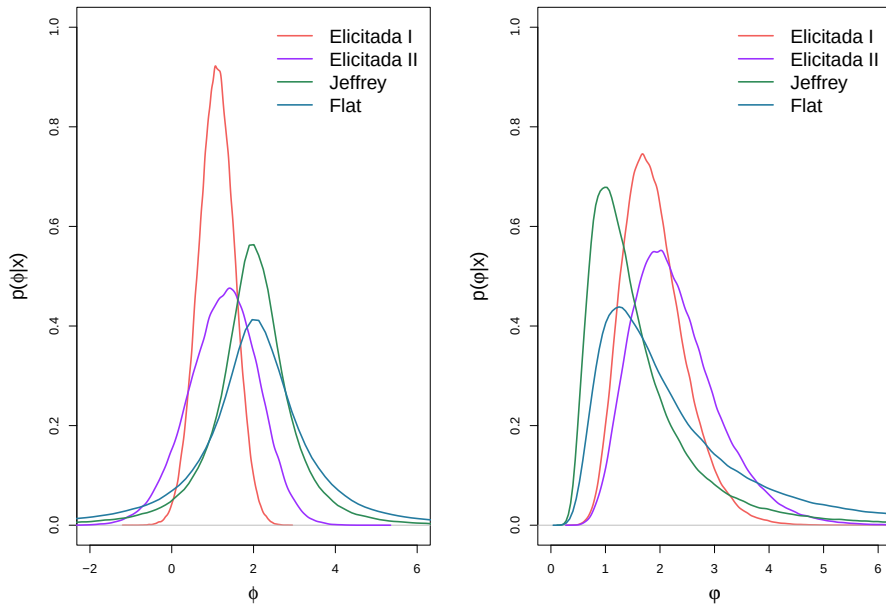


Figura 5.3: Densidade posteriori marginal de ϕ e φ para as priori Elicitada I, Elicitada II Jeffrey e Flat com tamanho de amostra $n = 3$



5.3 Dados simulados

A simulação foi realizada utilizando a linguagem de programação R. Foram geradas 1.000 amostras de tamanhos 3, 5, 10, 15 e 20, provenientes da distribuição de Gumbel com os parâmetros $\phi = 1$ e $\varphi = 2$. A partir de cada uma dessas amostras, foi gerada uma distribuição posteriori, para cada uma das prioris descritas no capítulo 3, de tama-

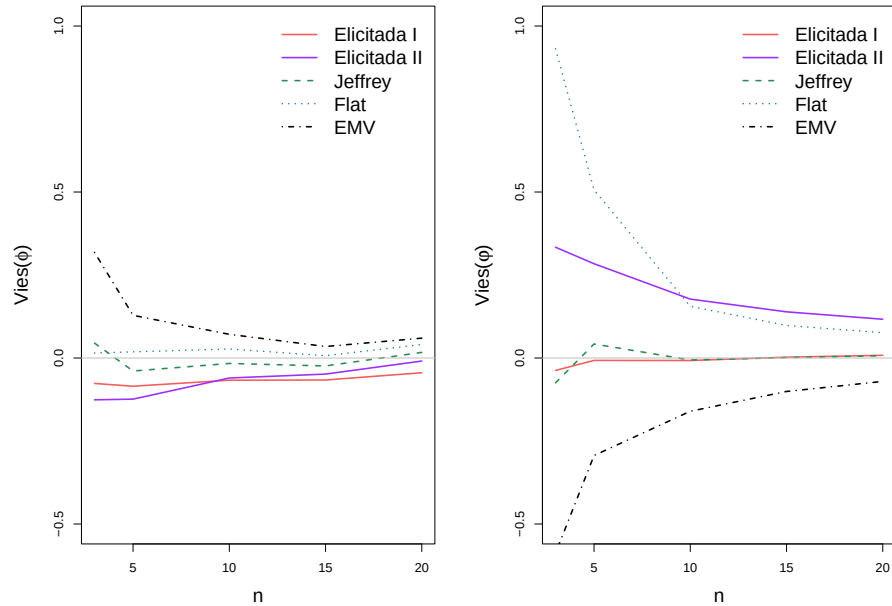
nho 20.000 com periodo de burn-in de 5000, utilizando o algoritmo MCMC descrito no Capítulo 5.1. O calculo das estimativas bayesianos foi dado pela média da posteriori como mencionado em 3.3. Os EMV foram obtidos resolvendo o sistema 2.7, para isso foi usado a função `nleqslv` da biblioteca de mesmo nome `nleqslv` da linguagem de programação R.

Para avaliar a performance dos estimadores bayesianos e EMV para cada parâmetro, foram analisados o erro quadrático médio (EQM) e o viés, definidos a seguir:

$$Vis(\theta_i) = \frac{1}{N} \sum_{j=1}^N (\hat{\theta}_{i,j} - \theta_i), \quad EQM(\theta_i) = \frac{1}{N} \sum_{j=1}^N (\hat{\theta}_{i,j} - \theta_i)^2, \quad \text{para } i = 1, 2$$

Os estimadores bayesianos serão identificados com base na distribuição de probabilidade a priori utilizada. Por exemplo, a posteriori derivada da priori de Jeffreys será referida como posteriori de Jeffreys, e o correspondente estimador será chamado de estimador de Jeffreys.

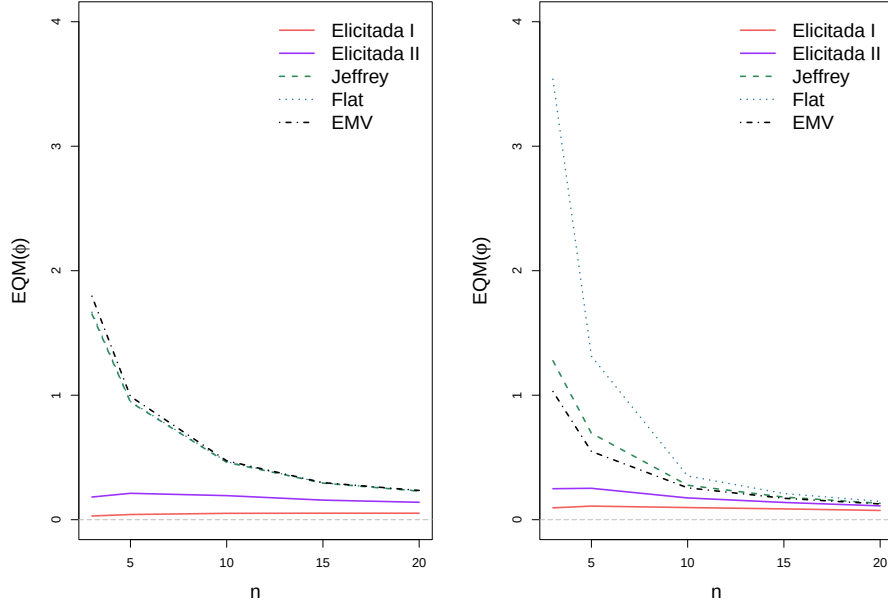
Figura 5.4: Viés para amostras de tamanho n com $\phi = 1$ e $\varphi = 2$ para diferentes tipos de prioris



Na Figura 5.4, observa-se que o estimador elicitado I para o parâmetro φ apresenta o menor viés. Para o parâmetro ϕ , por outro lado, há uma convergência mais lenta em direção a um viés igual a zero, além de apresentar um desempenho superior apenas ao estimador de máxima verossimilhança (EMV). O estimador elicitado II apresenta uma convergencia rapida para ϕ e lenta para φ .

Por outro lado, a análise do erro quadrático médio (EQM), apresentada na Figura 5.5, revela que o estimador elicitado, I ou II, possui um EQM significativamente menor para ambos os parâmetros, quando comparado aos demais estimadores. Esse resultado sugere que a curva mais acentuada observada na Figura 5.3, em relação às outras distribuições posteriores marginais, não é um caso isolado derivado de uma amostra que

Figura 5.5: EQM para amostras de tamanho n com $\phi = 1$ e $\varphi = 2$ para diferentes tipos de prioris



favoreceu a posteriori associada as prioris elicitadas. Os dados sobre o viés e o EQM estão apresentados com mais detalhes no apêndice ??.

Em seguida, utilizamos todos os estimadores gerados — ou seja, 1000 estimadores para cada tamanho de amostra — para analisar os intervalos definidos pelos quantis de 0.025 e 0.975 de cada conjunto de dados. Os intervalos obtidos foram os seguintes:

Tabela 5.1: Intervalo que contem 95% para os estimadores de $\phi = 1$ e $\varphi = 2$.

ϕ	Elicitada I	Elicitada II	Jeffrey	Flat	EMV
$n = 3$	(0.6088, 1.1465)	(0.0901, 1.5911)	(-1.1207, 3.8893)	(-1.1704, 3.7702)	(-0.9127, 4.2784)
$n = 5$	(0.5521, 1.2448)	(0.0575, 1.7576)	(-0.6928, 3.0771)	(-0.6420, 3.1850)	(-0.5430, 3.2845)
$n = 10$	(0.5219, 1.3492)	(0.1129, 1.8255)	(-0.2577, 2.4029)	(-0.2146, 2.4440)	(-0.1746, 2.5013)
$n = 15$	(0.5097, 1.3560)	(0.1919, 1.7106)	(-0.0265, 2.0674)	(0.0022, 2.0902)	(0.0223, 2.1372)
$n = 20$	(0.5373, 1.3871)	(0.3208, 1.7409)	(0.1739, 2.0040)	(0.1977, 2.0487)	(0.2206, 2.0672)
φ	Elicitada I	Elicitada II	Jeffrey	Flat	EMV
$n = 3$	(1.4503, 2.6071)	(1.7553, 3.1579)	(0.3228, 4.6267)	(0.5427, 6.7430)	(0.2332, 3.4018)
$n = 5$	(1.3563, 2.6491)	(1.5066, 3.1364)	(0.6942, 3.9669)	(0.8289, 4.8611)	(0.5928, 3.2299)
$n = 10$	(1.4251, 2.6162)	(1.4841, 2.9318)	(1.0957, 3.1255)	(1.1982, 3.3863)	(1.0093, 2.8553)
$n = 15$	(1.4527, 2.6148)	(1.4966, 2.8553)	(1.2492, 2.9292)	(1.3111, 3.0711)	(1.1860, 2.7735)
$n = 20$	(1.4937, 2.5559)	(1.5262, 2.7539)	(1.3530, 2.7534)	(1.4003, 2.8451)	(1.2970, 2.6439)

Os dados indicam que o menor intervalo é obtido pelos estimadores bayesianos elicitados, que diminui lentamente à medida que o tamanho da amostra aumenta. Além disso, observa-se que o estimador de máxima verossimilhança (EMV) apresenta desempenho semelhante ao estimador de Jeffrey, enquanto o estimador flat resulta no maior intervalo.

Capítulo 6

Conclusão

O estimador bayesiano com a priori elicitada, derivada da aproximação de Laplace, apresenta um desempenho significativamente superior aos demais estimadores, mostrando-se uma excelente opção, se não a melhor, para estimar os parâmetros da distribuição de Gumbel. Mesmo os quantis fornecidos pelo especialista, com erro de 10%, mostraram-se úteis para o modelo.

Por outro lado, temos uma maior complexidade de obter os hiperparâmetros da prior de elicitada, especialmente na resolução do sistema de quantis, que na maioria das vezes é não linear. Isso exige o uso de métodos numéricos, os quais, por sua vez, exigem uma escolha adequada da semente inicial, que nem sempre é fácil de determinar. Além disso, o algoritmo utilizado para resolver o sistema muitas vezes enfrentou dificuldades de convergência, dado que o sistema é altamente sensível a pequenas variações nos hiperparâmetros, tornando a solução do problema trabalhosa, mesmo com o auxílio computacional.

Apesar dessas dificuldades, o sistema de quantis, baseado em integrais aproximadas pelo método de Laplace, revelou-se uma ferramenta eficiente para a elicitação de hiperparâmetros, ampliando ainda mais as aplicações desse método.

6.1 Propostas futuras

Com base nas análises realizadas e nas dificuldades encontradas na aplicação da prior derivada do método de Laplace, as seguintes propostas têm como objetivo explorar os limites do método e tornar sua aplicação mais simples:

1. Analisar o comportamento do estimador elicitado para diferentes níveis de erro nos quantis, com o objetivo de identificar um limite de erro que não prejudique a performance do estimador.
2. Testar diferentes distribuições para as priors independentes, a fim de obter um sistema de quantis cuja solução seja mais fácil de convergir com o uso de métodos numéricos.
3. Investigar novas abordagens para a resolução do sistema de quantis, como o método dos mínimos quadrados ou formas de linearizar o sistema.

Bibliografia

JEFFREYS, H. Theory of Probability. [S.l.]: Oxford University Press, 1961.

MOALA, F. A.; DEY, S. Objective and subjective prior distributions for the gompertz distribution. Anais da Academia Brasileira de Ciências, SciELO Brasil, v. 90, p. 2643–2661, 2018.

O'HAGAN, A. Eliciting expert beliefs in substantial practical applications [read before the royal statistical society at a meeting on 'elicitation' on wednesday, april 16th, 1997, the president, professor a. f. m. smith in the chair]. J Royal Statistical Soc D, Wiley, v. 47, n. 1, p. 21–35, mar 1998.

RAFTERY, A. E.; LEWIS, S. M. [practical markov chain monte carlo]: comment: one long run with diagnostics: implementation strategies for markov chain monte carlo. Statistical science, JSTOR, v. 7, n. 4, p. 493–497, 1992.

TIERNEY, L.; KADANE, J. B. Accurate approximations for posterior moments and marginal densities. Journal of the American Statistical Association, [American Statistical Association, Taylor and Francis, Ltd.], v. 81, n. 393, p. 82–86, 1986. ISSN 01621459. Disponível em: <<http://www.jstor.org/stable/2287970>>.

Apêndice A

Tabelas de Resultados

As seguintes tabelas são referentes aos gráficos do capítulo 5.3

Tabela A.1: Viés para amostras de tamanho n com $\phi = 1$ e $\varphi = 2$ para diferentes tipos de prioris.

ϕ	Elicitada I	Elicitada II	Jeffrey	Flat	EMV
n=3	-0.0766	-0.1259	0.0151	0.0448	0.3180
n=5	-0.0850	-0.1238	0.0186	-0.0392	0.1286
n=10	-0.0669	-0.0602	0.0272	-0.0164	0.0714
n=15	-0.0662	-0.0484	0.0070	-0.0239	0.0347
n=20	-0.0444	-0.0091	0.0408	0.0170	0.0598
φ	Elicitada I	Elicitada II	Jeffrey	Flat	EMV
n=3	-0.0374	0.3337	0.9313	-0.0749	-0.5879
n=5	-0.0072	0.2840	0.5063	0.0423	-0.2937
n=10	-0.0075	0.1774	0.1559	-0.0060	-0.1604
n=15	0.0025	0.1391	0.0981	0.0013	-0.1007
n=20	0.0082	0.1168	0.0762	0.0070	-0.0705

Tabela A.2: EQM para amostras de tamanho n com $\phi = 1$ e $\varphi = 2$ para diferentes tipos de prioris.

ϕ	Elicitada I	Elicitada II	Jeffrey	Flat	EMV
n=3	0.0299	0.1822	1.6662	1.6517	1.7955
n=5	0.0416	0.2120	0.9489	0.9443	0.9903
n=10	0.0511	0.1934	0.4657	0.4610	0.4730
n=15	0.0522	0.1573	0.2953	0.2937	0.2977
n=20	0.0520	0.1399	0.2321	0.2294	0.2350
ϕ	Elicitada I	Elicitada II	Jeffrey	Flat	EMV
n=3	0.0952	0.2489	3.5380	1.2770	1.0300
n=5	0.1092	0.2523	1.3130	0.6947	0.5484
n=10	0.0976	0.1748	0.3512	0.2767	0.2559
n=15	0.0867	0.1387	0.2105	0.1815	0.1719
n=20	0.0743	0.1111	0.1472	0.1317	0.1267

Apêndice B

Códigos na linguagem R

Os códigos apresentados a seguir foram utilizados para obter os resultados discutidos neste trabalho. Cabe destacar que, devido a limitações do programa utilizado para elaborar este documento, optou-se por evitar o uso de caracteres especiais nos comentários e nas funções dos códigos.

```
#bibliotecas

library('evd') #gerar amostras da distribuicao gumbel
library('nleqslv') #resolve sistema de quantis
```

Listing B.1: Bibliotecas essenciais

B.1 Geração de Amostras Marginais da Distribuição de Probabilidade Posterior

```
#Posteriori proporcional (Ver 3.2)
#Distribuicao de probabilidade a priori Flat
f_flat <- function(phi, varphi, x) {
  prod(dgumbel(x, loc = phi, scale = varphi)) *
  dnorm(phi, mean = 0, sd = 1000) *
  dgamma(varphi, shape = 0.001, rate = 0.001)
}

#Distribuicao de probabilidade a priori Jeffrey
f_jeff <- function(phi, varphi, x) {
  prod(dgumbel(x, loc = phi, scale = varphi))/(varphi^2)
}

#Distribuicao de probabilidade a priori Elicitada
f_lapl <- function(phi, varphi, x) {
  prod(dgumbel(x, loc = phi, scale = varphi)) *
  dgamma(varphi, shape = 8.3526854, rate = 4.2140159) *
  dnorm(phi, mean = 0.91802, sd = 0.4511)
}
```

Listing B.2: Distribuição de probabilidade a Posteriori Proporcional

```
MH <- function(f, amostra, init, v, n) {
  n <- n + 5000 #Aumentando o numero de iteracoes
  samples <- matrix(NA, nrow = n, ncol = 2) # Matriz de
    amostras
  colnames(samples) <- c("phi", "varphi")
  samples[1, ] <- init # Iniciando a primeira amostra

  atual_phi <- init[1]
  atual_varphi <- init[2]

  for (i in 2:n) {
    # Proposta para phi
    proposta_phi <- rnorm(1, mean = atual_phi, sd = v[1])
    L_phi <- f(proposta_phi, atual_varphi, amostra) / f(atual_
      phi, atual_varphi, amostra)
    q_phi <- dnorm(atual_phi, mean = proposta_phi, sd = v[1]) /
      dnorm(proposta_phi, mean = atual_phi, sd = v[1])

    # Aceitar ou rejeitar a proposta para phi
    if (runif(1) < min(1, L_phi * q_phi)) {
      atual_phi <- proposta_phi
    }

    # Proposta para varphi
    proposta_varphi <- rgamma(1, shape = atual_varphi, rate = v
      [2])
    #verificando possivel erro
    while(proposta_varphi < 1e-100){
      proposta_varphi <- rgamma(1, shape = atual_varphi, rate =
        v[2])
    }

    L_varphi <- f(atual_phi, proposta_varphi, amostra) / f(
      atual_phi, atual_varphi, amostra)
    q_varphi <- dgamma(atual_varphi, shape = proposta_varphi,
      rate = v[2]) / dgamma(proposta_varphi, shape = atual_
        varphi, rate = v[2])
    #ratio_varphi <- L_varphi * q_varphi

    # Aceitar ou rejeitar a proposta para varphi
    if (runif(1) < min(1, L_varphi * q_varphi)) {
      atual_varphi <- proposta_varphi
    }

    # Armazenando a amostra atual
    samples[i, ] <- c(atual_phi, atual_varphi)
  }
}
```

```

#Retorna as amostras apos o burn-in
return(samples[5001:n, ])
}

#Gerando amostras
amostra = rgumbel(n=3, loc=1, scale=2)
#f      = funcao alvo
#init   = semente
#v      = variancia das distribuicoes propostas normal e gamma
#n      = Quantidade de amostras desconsiderando o periodo de
        burn-in de 5000

MH(f=f_flat, amostra, init=c(1,1), v=c(3, 1.5), n=20e3)
MH(f=f_jeff, amostra, init=c(1,1), v=c(3, 1.5), n=20e3)
MH(f=f_lapl, amostra, init=c(1,1), v=c(3, 1.5), n=20e3)

```

Listing B.3: Algoritmo de MCMC para distribuição Gumbel

B.2 Resolução de Sistema de Quantis Derivadas da Aproximação de Laplace

```

h <- function(x, v, t) {
  v = abs(v);
  v[1] - dlnorm(v[1], (x[3]-t)/v[2], x[4]/v[2], log=T) - dgamma
    (v[2], shape = x[1], rate = x[2], log=T)
}

#1a derivada
dz1 <- function(x, v, t) {
  x[2] + (log(v[1])*(v[2]*log(v[1]) - x[3] + t))/(x[4]^2) - x
    [1]/v[2]
}

dj1 <- function(x, v, t) {
  (v[1]*x[4]^2 + (v[2]^2)*log(v[1]) + x[4]^2 + t*v[2] - x[3]*v
    [2])/(v[1]*x[4]^2)
}

#2a derivada
dz2 <- function(x, v, t) {
  (log(v[1])/x[4])^2 + x[1]/(v[2]^2)
}

dj2 <- function(x, v, t) {
  - ((v[2]^2)*log(v[1]) + x[4]^2 + t*v[2] - v[2]^2 - x[3]*v[2])
    /((v[1]*x[4])^2)
}

#DTZ

```

```

djz <- function(x, v, t) {
  (2*v[2]*log(v[1]) - x[3] + t)/(v[1]*x[4]^2)
}

hessiana <- function(x, v, t) {
  matrix(c(dj2(x, v, t), djz(x, v, t),
    djz(x, v, t), dz2(x, v, t)),
    ncol = 2, byrow = T)
}

```

Listing B.4: Derivadas de Primeira, segunda Ordem e Matriz Hessiana

```

# x = vetor de hiperparametros (alpha, beta, mu, sigma) e t =
  quantil
laplace <- function(p, t) {

  v = optim(c(1, 1),
    fn = function(x) h(p, x, t)
  )$par
  v = abs(v)
  2*pi*exp(-h(p, v, t))*sqrt(abs(1/det(hessiana(p, v, t))))
}

```

Listing B.5: Aproximação de Laplace

```

#0 quantil dentro do parametro foi dado pelo especialista,
  elicitada II, contem um erro de menos 10% em relacao ao
  quantil. Por exemplo -0.2806739 corresponde ao quantil 0.15
#Obs. infelizmente as sementes iniciais para obter os
  hiperparametros 8.3526854, 4.2140159, 0.91802, 0.4511 da
  priori elicitada I nao foi encontrada.
sistema_laplace <- function(x) {
  c(
    laplace(x, -0.2806739) - 0.25,
    laplace(x, 1.1748431) - 0.50,
    laplace(x, 2.6843020) - 0.75,
    laplace(x, 3.9998800) - 0.90
  )
}

#Resolvendo o Sistema
nleqslv(c(6.56204976, 2.5113997, 0.7649900, 0.9821097),
  sistema_laplace,
  method = 'Newton'
)

```

Listing B.6: Sistema de Laplace