UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Faculdade de Ciências e Tecnologia
Câmpus de Presidente Prudente

Elicitação para a Distribuição Gumbel

Naysha Mirella Mena Ccanto

Orientador: Prof. Dr. Fernando Antonio Moala

Programa: Matemática Aplicada e Computacional

Presidente Prudente, julho de 2026

UNIVERSIDADE ESTADUAL PAULISTA

Faculdade de Ciências e Tecnologia de Presidente Prudente

Programa de Pós-Graduação em Matemática Aplicada e Computacional

Elicitação para a Distribuição Gumbel

Naysha Mirella Mena Ccanto

Orientador: Prof. Dr. Fernando Antonio Moala

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Matemática Aplicada e Computacional da Faculdade de Ciências e Tecnologia da UNESP.

Presidente Prudente, julho de 2026

C386e

Ccanto, Naysha Mirella Mena
Elicitação para a Distribuição Gumbel / Naysha Mirella Mena Ccanto. -- , 2026
84 p. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Tecnologia, Presidente Prudente,
Orientador: Fernando Antonio Moala

1. Distribuição Gumbel. 2. Inferência Bayesiana. 3. Elicitação. 4. Aproximação de Laplace. 5. Estatística. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Dados fornecidos pelo autor(a).

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Elicitação para a Distribuição Gumbel

AUTORA: NAYSHA MIRELLA MENA CCANTO
ORIENTADOR: FERNANDO ANTONIO MOALA

Aprovada como parte das exigências para obtenção do Título de Mestra em Matemática Aplicada e Computacional, pela Comissão Examinadora:

gov.br Documento assinado digitalmente
FERNANDO ANTONIO MOALA
Data: 01/08/2026 18:11:47-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. FERNANDO ANTONIO MOALA (Participaçao Presencial)
Departamento de Estatística / UNESP / Câmpus de Presidente Prudente - FCT

Prof. Dr. MARCOS VINICIUS DE OLIVEIRA PERES (Participaçao Presencial)
Estatística / Universidade Estadual de Maringá

Professor Doutor RICK ANDERSON FREIRE MANGUEIRA (Participaçao Presencial)
Departamento de Estatística / UNESP / Câmpus de Presidente Prudente - FCT

Presidente Prudente, 30 de abril de 2026.

Dedico este trabalho à minha mãe, Madeleine Ccanto, com todo o meu amor e gratidão.

Agradecimentos

Gostaria de expressar, em primeiro lugar, minha gratidão a Deus, pois sem Ele nada seria possível. Ao longo de toda esta trajetória, concedeu-me paz, serenidade e força, estando presente em todos os momentos.

Aos meus pais, Madeleine Ccanto e Juan Mena, minha profunda gratidão pelo amor, apoio incondicional, por sempre me darem forças para seguir em frente e acreditarem em mim em todos os momentos, contribuindo de maneira essencial para que eu pudesse chegar até aqui. À minha mãe, por ser sempre a primeira a me apoiar, incentivar e encorajar, por ser uma fonte constante de inspiração, pelo exemplo de dedicação, valentia e perseverança, por enfrentar as adversidades com coragem e nunca recuar diante das dificuldades. Ao meu pai, pela confiança, pelo incentivo, pelo exemplo de esforço e dedicação que me impulsionou a seguir adiante.

Ao meu irmão, Andres Mena, por estar presente quando mais precisei, pela alegria, pela companhia e, sobretudo, por sua existência, que é para mim uma fonte constante de felicidade.

Agradeço ao meu orientador, Fernando Moala, pela orientação ao longo do desenvolvimento deste trabalho.

Aos meus avós, Melania Curo e Felix Ccanto, que sempre estiveram presentes, oferecendo amor, cuidado e acolhimento em todos os momentos, e por me lembrarem, com seu exemplo, que Deus está sempre comigo.

A todos os meus tios, em especial às minhas tias Yeny, Miriam, Mily e Yesica, pelo carinho, pelos conselhos, pelo incentivo e pelo apoio, que foram muito importantes para mim ao longo desta caminhada.

Agradeço também aos meus amigos, que foram um suporte fundamental nesta etapa. Em especial, nos primeiros anos do mestrado, à minha amiga Susan Ayala, pelo apoio nos momentos mais difíceis, pelos conselhos e pela companhia constante, assim como a Elard Hurtado, pelos conselhos e pelo apoio.

Ao meu amigo Yofer Flores, pelos ânimos e por estar presente quando precisei. A Lito Bocanegra, pelo carinho, apoio e presença constante. Às minhas amigas Mishelle, Yhajayra e Karina, pela companhia e alegria.

Agradeço também aos meus amigos Ever e Narayan, que estiveram presentes na fase final desta etapa, pela amizade e pelo apoio.

Por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) pelo apoio financeiro concedido durante o desenvolvimento deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001.

Resumo

Este trabalho propõe uma abordagem metodológica, no contexto da Inferência Bayesiana, para a estimação dos hiperparâmetros de uma distribuição a priori informativa construída a partir de quantis fornecidos por um especialista para a distribuição Gumbel. A metodologia utiliza a distribuição preditiva a priori para formalizar o conhecimento subjetivo, viabilizando a resolução de integrais complexas por meio da aproximação de Laplace proposta por Kass, Tierney e Kadane em 1986. O método foi validado por meio de simulações, um exemplo ilustrativo e uma aplicação a dados reais de precipitação máxima anual da estação de Penta-di-Casinca (França). Os resultados demonstram que a aproximação de 1986 apresenta alta fidelidade e precisão ao replicar os pontos fornecidos pelo especialista. Além disso, a aplicação prática evidenciou a eficácia da atualização bayesiana, na qual a incorporação da evidência empírica permitiu corrigir e ajustar as crenças iniciais do especialista de forma coerente. Conclui-se que a estratégia baseada na aproximação de Laplace constitui uma ferramenta resiliente e tecnicamente fundamentada para a modelagem de eventos extremos, equilibrando rigor matemático com flexibilidade prática na incorporação de informações a priori.

Palavras-chave: Inferência Bayesiana; elicitação; aproximação de Laplace; distribuição Gumbel; valores extremos.

Abstract

This study proposes a methodological approach, within the Bayesian inference framework, for estimating the hyperparameters of an informative prior distribution constructed from quantiles provided by an expert for the Gumbel distribution. The methodology relies on the prior predictive distribution to formalize subjective knowledge, enabling the evaluation of complex integrals through the Laplace approximation proposed by Kass, Tierney, and Kadane (1986).

The method is validated through simulation studies, an illustrative example, and an application to real data on annual maximum precipitation from the Penta-di-Casinca station (France). The results show that the 1986 approximation achieves high fidelity and accuracy in reproducing the expert's specified quantiles. Furthermore, the practical application highlights the effectiveness of Bayesian updating, where the incorporation of empirical evidence allows for coherent correction and refinement of the expert's initial beliefs.

It is concluded that the strategy based on the Laplace approximation constitutes a robust and theoretically grounded tool for modeling extreme events, balancing mathematical rigor with practical flexibility in incorporating prior information.

Keywords: Bayesian inference; elicitation; Laplace approximation; Gumbel distribution; extreme values.

Lista de Figuras

5.1.1 Densidade da Distribuição Gumbel para diferentes valores dos parâmetros ϕ e φ. 32

6.3.1 Distribuições a priori do parâmetro ϕ. 43
6.3.2 Distribuições a priori do parâmetro φ. 44

7.3.1 Viés das estimativas dos parâmetros sob os diferentes cenários considerados. 49
7.3.2 Erro quadrático médio (EQM) das estimativas dos parâmetros sob os diferentes cenários. 50
7.3.3 Erro absoluto médio (MAE) das estimativas dos parâmetros sob os diferentes cenários. 50
7.3.4 Probabilidade de cobertura dos intervalos de credibilidade para os parâmetros, sob os diferentes cenários. 51
7.3.5 Distribuições a posteriori do parâmetro de localização ϕ para amostra pequena ($n = 3$). 51
7.3.6 Comportamento das distribuições a posteriori de ϕ para $n = 6$. 52
7.3.7 Distribuições a posteriori de ϕ para $n = 10$. 52
7.3.8 Distribuições a posteriori do parâmetro de escala φ para $n = 3$. 53
7.3.9 Distribuições a posteriori de φ para $n = 6$. 53
7.3.10 Distribuições a posteriori de φ para $n = 10$. 54

8.2.1 Distribuições a posteriori do parâmetro de localização ϕ sob diferentes prioris ($n = 3$). 59
8.2.2 Distribuições a posteriori do parâmetro de escala φ sob diferentes prioris ($n = 3$). 59
8.2.3 Traceplots das cadeias MCMC para os parâmetros ϕ e φ. 60
8.2.4 Função de autocorrelação das cadeias MCMC para ϕ e φ. 61

9.4.1 Atualização bayesiana: comparação entre a distribuição a priori (elicitação do especialista), a distribuição a posteriori e os pontos de elicitação. 65
9.4.2 Distribuições a priori obtidas a partir da elicitação do especialista e dos cenários perturbados (Elicitada I e Elicitada II). 65
9.5.1 Distribuições a posteriori sob diferentes prioris: elicitação do especialista, cenários perturbados e prioris não informativas. 67

Lista de Tabelas

6.3.1 Quantis utilizados nos cenários de eliciação. 42
6.3.2 Hiperparâmetros obtidos via aproximação de Laplace (1986). 43

7.3.1 Resultados da Simulação MCMC (1.000 Réplicas): Viés, EAM, EQM e Cobertura para ϕ e φ 48

8.2.1 Estimativas a posteriori ($n = 3$) 58

9.2.1 Informação a priori fornecida pelo especialista. 63
9.3.1 Hiperparâmetros estimados para cada cenário de elicitação. 64
9.4.1 Comparação entre probabilidades do especialista, priori e posteriori. . . 64
9.5.1 Comparação das probabilidades fornecidas pelo especialista e das probabilidades obtidas sob diferentes prioris. 66

A.6.1 Hiperparâmetros estimados via Laplace (1989). 80

Sumário

Resumo **5**

Abstract **7**

Lista de Figuras **9**

Lista de Tabelas **11**

Sumário **13**

1 Introdução **15**

2 Análise Bayesiana **17**
2.1 Distribuições a Priori 18
2.1.1 Classificação das Prioris 18
2.1.2 Priori Informativa 19
2.1.3 Priori Não Informativa 19
2.1.4 Prioris Conjugadas 20
2.2 Distribuições Preditivas 20
2.2.1 Preditivas a Priori 20
2.2.2 Preditivas a Posteriori 21

3 Método de Laplace para Aproximação de Integrais **23**
3.1 Método proposto por Tierney e Kadane (1986) 23

4 Elicitação **27**
4.1 Introdução à Elicitação 27
4.2 Classes Não-paramétricas de Prioris 28
4.3 Modelos Gerais de Elicitação 29
4.3.1 Modelo Linear Normal 29
4.3.2 Modelo Beta-Binomial 29

5 Distribuição Gumbel **31**
5.1 Definição e aplicações 31
5.2 Propriedades da distribuição Gumbel 32
5.3 Estimação de máxima verossimilhança 33
5.4 Matriz de informação de Fisher e inferência frequentista 33
5.4.1 Intervalos de confiança 34
5.5 Inferência Bayesiana com prioris não-informativas e fracamente informativas 35
5.5.1 Estrutura de priori independente para ϕ e φ 35

5.5.2 Priori de Jeffreys (não informativa) 36
5.5.3 Priori Flat (não informativa uniforme) 37
5.5.4 Priori fracamente informativa 37
5.5.5 Posteriori a partir de prioris não-informativas 37

6 Elicitação de priori via preditiva a priori sob aproximação de Laplace 39
6.1 O método 39
6.2 Aplicação da aproximação de Laplace 41
6.3 Obtenção dos hiperparâmetros via aproximação de Laplace (1986) 42

7 Estudo de Simulação 45
7.1 Desenho do Estudo de Simulação 45
7.2 Implementação do Método de Monte Carlo via Cadeias de Markov (MCMC) 46
7.3 Resultados do Estudo de Simulação 48
7.3.1 Gráfico das Distribuições a Posteriori 51

8 Ilustração Numérica 57
8.1 Exemplo ilustrativo de uma amostra de tamanho 3 57
8.1.1 Configuração do experimento 57
8.1.2 Algoritmo de inferência 58
8.2 Análise dos resultados 58

9 Aplicação a Dados Reais: Pluviometria em Penta-di-Casinca 63
9.1 Descrição da Amostra e Contextualização 63
9.2 Informação do Especialista (Pontos de Elicitação) 63
9.3 Elicitação via Aproximação de Laplace 64
9.4 Validação da Priori 64
9.5 Comparação entre Diferentes Prioris 66

10 Conclusões 69
10.1 Eficácia da Metodologia de Elicitação 69
10.2 Desempenho em Estudos de Simulação 69
10.3 Aplicação a Dados Reais 70
10.4 Considerações Metodológicas 70
10.5 Propostas Futuras 70

Referências 71

A Elicitação via Aproximação de Laplace (1989) 75
A.1 Método proposto por Tierney e Kadane (1989) 75
A.2 Formulação da Distribuição Preditiva 78
A.3 Reescrita na Forma de Laplace 78
A.4 Ponto Estacionário 79
A.5 Aproximação de Laplace 79
A.6 Obtenção dos Hiperparâmetros 80
A.7 Considerações Finais 80

Capítulo

1

Introdução

Na análise estatística de fenômenos complexos, o método de máxima verossimilhança é frequentemente a primeira escolha devido às suas propriedades assintóticas. Contudo, em cenários caracterizados por amostras pequenas ou eventos raros, tais estimadores podem apresentar elevada variabilidade e imprecisão. Diante dessa limitação, a Inferência Bayesiana surge como uma alternativa robusta, permitindo a integração sistemática entre a evidência empírica dos dados e o conhecimento prévio sobre o fenômeno, formalizado por meio de uma distribuição a priori.

A tradução do conhecimento subjetivo de especialistas em modelos probabilísticos, processo denominado elicitação, é fundamental quando a disponibilidade de dados é restrita. Em áreas como a hidrologia e a análise de confiabilidade, o julgamento de especialistas sobre eventos extremos muitas vezes constitui a fonte de informação mais rica disponível. Entretanto, o principal desafio reside em transformar percepções subjetivas em hiperparâmetros estatisticamente utilizáveis, sem introduzir vieses matemáticos indesejados.

Uma abordagem prática para superar essa barreira é a elicitação baseada em quantis de probabilidade, que são quantidades mais intuitivas para o especialista do que os parâmetros diretos de uma distribuição. Para viabilizar essa transição, utiliza-se a distribuição preditiva a priori. No entanto, o cálculo dessa distribuição frequentemente envolve integrais intratáveis analiticamente, exigindo o emprego de métodos de aproximação.

Este trabalho propõe uma metodologia para a estimação de hiperparâmetros de uma distribuição a priori informativa para a distribuição Gumbel, utilizando a distribuição preditiva a priori resolvida por meio da Aproximação de Laplace proposta por Kass, Tierney e Kadane (1986). O enfoque está na obtenção de estimativas que reproduzam com alta fidelidade os quantis fornecidos pelo especialista, preservando consistência teórica e estabilidade numérica.

A relevância deste estudo é demonstrada por meio de simulações e de uma aplicação prática a dados reais de precipitação máxima anual da estação de Penta-di-Casinca (Córsega, França), provenientes dos registros de Ruggeri e Bousquet (2017). Por meio dessa aplicação, busca-se evidenciar como a atualização bayesiana permite que a evidência empírica corrija e refine as crenças iniciais, proporcionando uma ferramenta eficaz para o planejamento e dimensionamento de projetos de engenharia em contextos de escassez de dados.

O objetivo central desta tese é, portanto, desenvolver e aplicar uma metodologia de elicitação baseada na aproximação de Laplace, fornecendo diretrizes consistentes para a modelagem de eventos extremos sob o paradigma bayesiano.

Capítulo

2

Análise Bayesiana

A valorização do Teorema de Bayes como fundamento da inferência estatística tem oscilado historicamente entre períodos de aceitação e rejeição desde a sua publicação em 1763. Em determinados momentos, quando métodos alternativos pareciam oferecer uma base satisfatória, os resultados bayesianos foram considerados meras curiosidades teóricas ou tentativas equivocadas de resolver um problema relevante. Contudo, o surgimento de dificuldades filosóficas e práticas nas abordagens clássicas, particularmente quando não se verificavam pressupostos como a normalidade e a independência dos erros, levou a um renovado interesse pelo raciocínio bayesiano, que ressurgiu com notável vigor.

A inferência bayesiana distingue-se da clássica na forma de tratar os parâmetros: enquanto a inferência frequentista os considera quantidades fixas, próprias mas desconhecidas da população, a abordagem bayesiana os modela como variáveis aleatórias. Essa concepção reflete uma visão mais flexível da incerteza, representada por distribuições de probabilidade. O núcleo deste enfoque é o Teorema de Bayes, que permite atualizar crenças prévias com base nos dados observados, combinando a distribuição a priori — $\pi(\theta)$, que expressa o conhecimento prévio sobre os parâmetros — com a função de verossimilhança - $L(x|\theta)$, que contém a informação proveniente dos dados — para obter a distribuição a posteriori:

$$p(\theta \mid x) = \frac{L(x \mid \theta)\pi(\theta)}{\int_\Theta L(x \mid \theta)\pi(\theta)}$$

onde

$$L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta) \tag{2.0.1}$$

Na prática, trabalha-se frequentemente com a proporcionalidade:

$$p(\theta \mid x) \propto L(x \mid \theta)\pi(\theta),$$

o que simplifica a análise. A partir da distribuição a posteriori podem derivar-se estimativas pontuais, como a esperança matemática $\hat{\theta} = E[p(\theta|x)]$, bem como intervalos de credibilidade e regras de decisão baseadas em risco.

Um elemento crucial da metodologia bayesiana é a escolha da distribuição a priori. Esta pode ser informativa, quando existe conhecimento sólido prévio, ou não informativa, como no caso da distribuição plana ou da distribuição de Jeffreys, que visam minimizar a influência subjetiva. Embora exista o receio de que uma escolha

inadequada da priori possa dominar a análise, uma seleção criteriosa, associada a uma estrutura de modelo apropriada, permite que os dados "falem por si" e forneçam informação clara, mesmo sobre aspectos duvidosos do modelo.

Em comparação, os métodos clássicos têm proporcionado soluções satisfatórias em contextos que cumprem pressupostos específicos, mas tendem a tornar-se inoperantes ou complexos quando estas condições não se verificam. A flexibilidade do enfoque bayesiano torna-o uma alternativa robusta, capaz de adaptar-se a cenários mais realistas e à complexidade inerente à investigação científica.

Durante muito tempo, a principal limitação prática da inferência bayesiana foi o elevado custo computacional dos cálculos, o que a tornava inviável face a métodos mais simples de executar manualmente. No entanto, o desenvolvimento da computação eletrônica e de algoritmos eficientes — como os métodos de Monte Carlo via Cadeias de Markov (MCMC) — eliminou essa barreira, permitindo que análises complexas sejam realizadas de forma rápida e acessível. Hoje, a disponibilidade de programas estatísticos especializados torna a análise bayesiana quase tão imediata quanto o enfoque clássico mais restrito.

Na prática científica, a inferência estatística insere-se num ciclo iterativo no qual a recolha e a análise de dados se alternam de forma controlada. Nesse processo, a inferência bayesiana oferece uma estrutura natural para rever hipóteses, incorporar novas evidências e reajustar modelos sem as limitações impostas por pressupostos estritamente técnicos. Essa adaptabilidade confere-lhe um valor especial como ferramenta metodológica, capaz de integrar conhecimento prévio e evidência empírica de forma coerente e quantitativa.

Finalmente, a utilidade de uma teoria estatística deve ser avaliada não apenas pela sua formalidade matemática, mas também pela sua capacidade de gerar conhecimento útil e coerente com o senso comum. A inferência bayesiana, impulsionada por avanços teóricos e computacionais, consolidou-se como um dos pilares da inferência moderna, oferecendo soluções consistentes mesmo em situações complexas e desafiadoras.

2.1 Distribuições a Priori

Uma etapa essencial na inferência bayesiana é a definição da distribuição *a priori*. Nesta seção, discutiremos sua conceituação, classificação e exemplos de aplicação.

De maneira geral, a priori, denotada por $\pi(\theta)$, representa todo o conhecimento ou suposição sobre o parâmetro θ antes da coleta de dados. Esse conhecimento pode derivar de opiniões de especialistas, resultados de pesquisas anteriores ou informações teóricas disponíveis.

De acordo com Martz (1991), uma distribuição a priori deve:

- Respeitar o espaço paramétrico do parâmetro;
- Garantir que a posteriori seja integrável;
- Refletir corretamente o conhecimento prévio de especialistas;
- Não se sobrepor à informação contida na verossimilhança.

2.1.1 Classificação das Prioris

As distribuições a priori podem ser classificadas de várias formas. Uma classificação comum distingue prioris próprias e impróprias conforme sua integrabilidade.

Priori própria: integrável no espaço paramétrico, mesmo que não seja uma função de densidade formal. Exemplo:

$$\theta \sim \text{Gamma}(a, b), \quad \pi(\theta) = I_{(0,1)}(\theta), \quad \pi(\theta) = \frac{1}{\theta} I_{(a,b)}(\theta) \tag{2.1.1}$$

Priori imprópria: não integrável. Exemplos:

$$\pi(\theta) \propto I_{(0,\infty)}(\theta) \tag{2.1.2}$$

2.1.2 Priori Informativa

Uma priori informativa incorpora conhecimento prévio sobre o parâmetro antes da observação dos dados, podendo se basear em estudos anteriores, literatura científica ou opinião de especialistas (Bernardo & Smith, 2009; Gelman et al., 2013). Ao combinar-se com a **verossimilhança**, resulta na **distribuição posteriori**, permitindo atualizar o conhecimento sobre θ com base em evidências observadas (McElreath, 2020).

As prioris informativas podem ser classificadas quanto ao grau de certeza:

- **Fortemente informativas:** baixa variância, alta concentração ao redor de valores específicos;
- **Moderadamente informativas:** incerteza razoável, permitindo que os dados influenciem mais;
- **Fracamente informativas:** restrições amplas e plausíveis, servindo como regularização e prevenindo estimativas extremas.

Vantagens incluem integração de evidências prévias, maior precisão em amostras pequenas ou incompletas e estabilização do modelo (Spiegelhalter, Abrams & Myles, 2004). Limitações incluem subjetividade e sensibilidade à escolha da priori, exigindo análise de sensibilidade (Gelman et al., 2013).

A formulação de uma priori informativa deve ser transparente, baseada em fontes empíricas, com escolha adequada da família de distribuição e definição coerente dos hiperparâmetros.

2.1.3 Priori Não Informativa

Prioris não informativas são utilizadas quando há pouco ou nenhum conhecimento prévio sobre o parâmetro. O objetivo é minimizar a influência da priori, permitindo que os dados dominem a inferência (Gelman et al., 2013).

Entre elas, destacam-se a priori uniforme (flat) e a priori de Jeffreys, com formulações e propriedades distintas.

2.1.3.1 Priori de Jeffreys

A priori de Jeffreys (Jeffreys, 1946) é fundamentada na *informação de Fisher*, que mede a quantidade de informação que uma amostra fornece sobre θ:

$$\pi_J(\theta) \propto \sqrt{I(\theta)}, \quad I(\theta) = -E\left[\frac{\partial^2 \ln L(\theta)}{\partial \theta^2}\right]$$

As vantagens são que as priors são invariantes a reparametrizações, de modo que transformações da forma $\phi = g(\theta)$ preservam sua estrutura, e que se baseiam na própria estrutura do modelo, o que contribui para reduzir a subjetividade na sua especificação.

As limitações são que a priori pode ser *imprópria* em espaços infinitos ou dependente de θ, sendo necessário verificar a integrabilidade da posteriori (Robert, 2007).

2.1.3.2 Prioris Uniformes (Flat)

Representam neutralidade total, atribuindo igual probabilidade a todos os valores possíveis:

$$\pi(\theta) = c, \quad \forall \theta \in \Theta$$

As limitações consistem no fato de que, em espaços infinitos, a integral da priori pode tornar-se divergente (*imprópria*) e que, em geral, as prioris não são invariantes sob transformações de parâmetro, o que pode induzir dependência em relação à parametrização adotada (Jaynes, 2003). Também é necessário verificar a integrabilidade da posteriori para assegurar a validade das inferências (Robert, 2007).

2.1.4 Prioris Conjugadas

Os hiperparâmetros permitem distinguir os parâmetros da priori dos parâmetros do modelo. Por exemplo, em uma distribuição Bernoulli, ρ representa o parâmetro do modelo, enquanto α e β são os hiperparâmetros da priori Beta.

Se $F = \{p(x|\theta); \theta \in \Theta\}$ denota a família de distribuições amostrais, uma família P é dita conjugada a F se, para todo $p(x|\theta) \in F$ e $p(\theta) \in P$, a posteriori $p(\theta|x)$ também pertencer a P, ou seja:

$$\forall p(x|\theta) \in F,\ p(\theta) \in P \Rightarrow p(\theta|x) \in P.$$

As principais vantagens das prioris conjugadas residem no fato de que facilitam a obtenção da posteriori, permitem incorporar uma ampla variedade de conhecimento prévio e tornam a atualização do modelo mais simples, pois envolve apenas alterações nos hiperparâmetros.

Exemplo: Binomial com n tentativas e probabilidade de sucesso p. A posteriori de uma priori conjugada Beta (α, β) resulta em:

$$p(\theta|X) = \text{Beta}(\alpha + x, \beta + n - x)$$

2.2 Distribuições Preditivas

2.2.1 Preditivas a Priori

Antes da coleta de dados, pode-se questionar: "Qual seria a distribuição esperada da observação?" ou "O que prevemos para y sem observar dados?"

A distribuição preditiva a priori corresponde à distribuição marginal de X, obtida integrando a verossimilhança sobre o espaço de parâmetros ponderado pela priori. Assim,

$$f(x) = \int_{\Theta} f(x \mid \theta)\, \pi(\theta)\, d\theta. \tag{2.2.1}$$

Para duas variáveis:

$$f(x, y) = \int_{\Theta} f(x, y|\theta)\pi(\theta) d\theta \tag{2.2.2}$$

Exemplo: Se $y \sim N(\mu, \sigma_0^2)$ com σ_0^2 conhecida e priori $\mu \sim N(\mu_a, \sigma_a^2)$, a marginal é:

$$p(y) = N(\mu_a, \sigma_0^2 + \sigma_a^2)$$

2.2.2 Preditivas a Posteriori

Nas previsões de valores futuros, a incerteza associada provém de duas fontes principais: a estimativa dos parâmetros com base nos dados observados anteriormente e a variabilidade intrínseca da variável aleatória futura.

No contexto bayesiano, a densidade preditiva a posteriori é definida como

$$f(y_p|y) = \int f(y_p|\theta)\, p(\theta|y)\, d\theta, \tag{2.2.3}$$

onde a incerteza sobre o parâmetro θ é integrada, refletindo a combinação entre o modelo amostral e a distribuição a posteriori dos parâmetros.

O cálculo da densidade preditiva pode ser realizado por diferentes abordagens. A primeira é a *analítica*, quando a integral acima pode ser resolvida de forma exata. Outra possibilidade é a *aproximação gaussiana*, na qual se assume que $[Y_p|y] \sim N(\mu_p, \sigma_p^2)$, sendo

$$\mu_p = E[Y_p|y] \quad \text{e} \quad \sigma_p^2 = \text{Var}[Y_p|y].$$

Por fim, pode-se empregar a *simulação*, que consiste em gerar amostras $\theta_i \sim p(\theta|y)$ e, para cada uma delas, valores $y_{p(i)} \sim f(y_p|\theta_i)$, obtendo-se assim uma aproximação empírica da distribuição preditiva.

Exemplo: Observação $Y \sim \text{Bin}(n, \theta)$ com priori Beta (p, q). Posteriori:

$$p(\theta|Y) = \text{Beta}(p + y, q + n - y)$$

Para N observações futuras $Z|\theta \sim \text{Bin}(N, \theta)$:

$$\begin{aligned} f(z \mid y) &= \int_0^1 \binom{N}{z} \theta^z (1-\theta)^{N-z} \frac{\theta^{p+y-1}(1-\theta)^{q+n-y-1}}{B(p+y, q+n-y)}\, d\theta \\ &= \binom{N}{z} \frac{B(z+p+y,\ N-z+q+n-y)}{B(p+y,\ q+n-y)}. \end{aligned} \tag{2.2.4}$$

Essa é a distribuição Beta-Binomial.

Capítulo 3

Método de Laplace para Aproximação de Integrais

Frequentemente não se obtêm soluções analíticas para integrais que surgem em aplicações de interesse prático; em particular, em inferência bayesiana aparecem integrais cujo valor numérico não é suficiente — requer-se uma expressão aproximada que permita manipulação analítica (por exemplo, para obter momentos a posteriori, densidades marginais ou densidades preditivas). Uma das técnicas mais utilizadas para obter aproximações analíticas dessas integrais é o *método de Laplace* (Tierney; Kadane, 1986).

O método baseia-se na expansão em série de Taylor (até segunda ordem, tipicamente) do expoente que aparece no integrando exponencial, transformando a integral numa aproximação baseada no comportamento local do integrando em torno do seu ponto de máxima contribuição. A seguir apresenta-se uma descrição do método, a sua generalização vetorial, bem como sua aplicação no contexto de elicitação bayesiana.

3.1 Método proposto por Tierney e Kadane (1986)

Seja $L(\phi \mid \boldsymbol{x})$ a função de verossimilhança associada ao parâmetro escalar ϕ e $\pi(\phi)$ a densidade a priori. Para uma função mensurável $g(\phi)$, contínua e não-negativa, o momento a posteriori é definido por

$$E[g(\phi) \mid \boldsymbol{x}] = \frac{\displaystyle\int g(\phi)\, L(\phi \mid \boldsymbol{x})\, \pi(\phi)\, d\phi}{\displaystyle\int L(\phi \mid \boldsymbol{x})\, \pi(\phi)\, d\phi}. \tag{3.1.1}$$

Em geral, as integrais que compõem o numerador e o denominador de (3.1.1) não admitem forma fechada, o que motiva o uso de métodos de aproximação assintótica. Nesse contexto, o método de Laplace fornece uma estratégia eficiente ao explorar o comportamento local da função integrando em torno de seu ponto dominante.

Aproximação de Laplace no caso unidimensional

Considere uma função duas vezes continuamente diferenciável $h : \mathbb{R} \to \mathbb{R}$, tal que $-h(\phi)$ possua um único máximo em $\hat{\phi}$. Para n suficientemente grande, deseja-se

aproximar integrais da forma

$$I = \int_{-\infty}^{\infty} \exp\{-n\, h(\phi)\}\, d\phi. \tag{3.1.2}$$

Expandindo $h(\phi)$ em torno de $\hat{\phi}$ por meio de uma expansão de Taylor de segunda ordem e desprezando termos de ordem superior, obtém-se a aproximação clássica de Laplace

$$\int_{-\infty}^{\infty} \exp\{-n\, h(\phi)\}\, d\phi \approx \sqrt{\frac{2\pi}{n\, h''(\hat{\phi})}}\, \exp\{-n\, h(\hat{\phi})\}, \tag{3.1.3}$$

onde $h''(\hat{\phi}) > 0$. Definindo $\sigma^2 = [h''(\hat{\phi})]^{-1}$, a expressão acima pode ser escrita de forma compacta como

$$\hat{I} = \left(\frac{2\pi\sigma^2}{n}\right)^{1/2} \exp\{-n\, h(\hat{\phi})\}.$$

Sob condições regulares, o erro relativo dessa aproximação é da ordem $O(n^{-2})$, conforme demonstrado em Tierney e Kadane (1986).

Aproximação do momento a posteriori

O momento a posteriori em (3.1.1) pode ser reescrito como o quociente de duas integrais exponenciais:

$$E[g(\phi) \mid \boldsymbol{x}] = \frac{\int \exp\{-n\, h^*(\phi)\}\, d\phi}{\int \exp\{-n\, h(\phi)\}\, d\phi},$$

em que

$$-n\, h(\phi) = \log \pi(\phi) + \log L(\phi \mid \boldsymbol{x}), \qquad -n\, h^*(\phi) = \log g(\phi) + \log \pi(\phi) + \log L(\phi \mid \boldsymbol{x}).$$

Aplicando a aproximação de Laplace (3.1.3) separadamente ao numerador e ao denominador, obtém-se a estimativa assintótica proposta por Tierney e Kadane (1986):

$$\widehat{E}[g(\phi) \mid \boldsymbol{x}] \approx \frac{\sigma^*}{\sigma}\, \exp\Big\{-n(h^*(\hat{\phi}^*) - h(\hat{\phi}))\Big\}, \tag{3.1.4}$$

onde $\hat{\phi}$ e $\hat{\phi}^*$ denotam, respectivamente, os pontos de máximo de $-h(\phi)$ e $-h^*(\phi)$, e $\sigma^2 = [h''(\hat{\phi})]^{-1}$, $(\sigma^*)^2 = [h^{*\prime\prime}(\hat{\phi}^*)]^{-1}$.

Extensão para o caso multidimensional

Quando o vetor de parâmetros $\boldsymbol{\phi} \in \mathbb{R}^m$, a aproximação de Laplace generaliza-se naturalmente. Seja $h(\boldsymbol{\phi})$ uma função diferenciável cujo mínimo ocorre em $\hat{\boldsymbol{\phi}}$. Então,

$$\int_{\mathbb{R}^m} \exp\{-n\, h(\boldsymbol{\phi})\}\, d\boldsymbol{\phi} \approx \left(\frac{2\pi}{n}\right)^{m/2} \det(\Sigma)^{1/2} \exp\{-n\, h(\hat{\boldsymbol{\phi}})\}, \tag{3.1.5}$$

onde $\Sigma^{-1} = \mathrm{D}^2 h(\hat{\boldsymbol{\phi}})$ representa a matriz Hessiana de h avaliada no ponto $\hat{\boldsymbol{\phi}}$.

Na prática, a aplicação da aproximação de Laplace em dimensão maior envolve a determinação numérica do ponto $\hat{\boldsymbol{\phi}}$ e da Hessiana correspondente, reduzindo um problema de integração a tarefas de otimização e cálculo matricial, o que torna o método computacionalmente atrativo.

Aplicação à elicitação via quantis preditivos

Considere agora o problema de elicitação a priori baseado em quantis da distribuição preditiva. Seja X uma variável aleatória com densidade condicional $f_X(x \mid \theta)$ e distribuição a priori $\pi(\theta \mid \alpha)$, com $\theta = (\theta_1, \ldots, \theta_m)$. O quantil preditivo x_p satisfaz

$$P(X \leq x_p) = \int_0^\infty \cdots \int_0^\infty \left[\int_0^{x_p} f_X(x \mid \theta)\, dx \right] \pi(\theta \mid \alpha)\, d\theta.$$

Reescrevendo,

$$\begin{aligned} P(X \leq x_p) &= \int_0^\infty \cdots \int_0^\infty \left[\int_0^{x_p} f_X(x \mid \theta)\, dx \right] \pi(\theta \mid \alpha)\, d\theta, \\ &= \int_0^\infty \cdots \int_0^\infty F_X(x_p \mid \theta)\, \pi(\theta \mid \alpha)\, d\theta, \\ &= \int_0^\infty \cdots \int_0^\infty \exp\{-h(\theta \mid x_p, \alpha)\}\, d\theta, \end{aligned}$$

onde

$$h(\theta \mid x_p, \alpha) = -\log\big(F_X(x_p \mid \theta)\, \pi(\theta \mid \alpha)\big).$$

Aplicando a aproximação de Laplace multidimensional, obtém-se

$$P(X \leq x_p) \approx (2\pi)^{m/2} \left|\Sigma^{-1}\right|^{1/2} \exp\{-h(\hat{\theta} \mid x_p, \alpha)\}.$$

Dado um conjunto de probabilidades acumuladas $\{p_1, \ldots, p_n\}$ fornecidas pelo especialista, associa-se a cada uma o respectivo quantil x_{p_i}, resultando no sistema

$$\begin{cases} p_1 = (2\pi)^{m/2} \left|\Sigma^{-1}\right|^{1/2} \exp\{-h(\hat{\theta} \mid x_{p_1}, \alpha)\}, \\ p_2 = (2\pi)^{m/2} \left|\Sigma^{-1}\right|^{1/2} \exp\{-h(\hat{\theta} \mid x_{p_2}, \alpha)\}, \\ \vdots \\ p_n = (2\pi)^{m/2} \left|\Sigma^{-1}\right|^{1/2} \exp\{-h(\hat{\theta} \mid x_{p_n}, \alpha)\}. \end{cases} \tag{3.1.6}$$

A solução do sistema (3.1.6) permite identificar o vetor de hiperparâmetros $\alpha = (\alpha_1, \ldots, \alpha_n)$, garantindo que a distribuição a priori seja compatível com a informação fornecida pelo especialista.

Cabe destacar que uma formulação alternativa do método de Laplace foi proposta por Tierney e Kadane (1989), na qual se consideram integrais com fatores multiplicativos fora do termo exponencial, permitindo uma generalização da aproximação clássica.

No entanto, no contexto específico deste trabalho, particularmente na aplicação à distribuição de Gumbel, essa abordagem conduz a um problema de não identificabilidade em relação a certos hiperparâmetros da priori, em especial o parâmetro b. Em consequência, diferentes valores desse hiperparâmetro produzem a mesma distribuição preditiva a priori, o que compromete a interpretação inferencial do modelo.

Por essa razão, a formulação de 1986 foi adotada como base metodológica principal desta dissertação. Para fins de completude teórica, os detalhes da abordagem de 1989 são apresentados no Apêndice A.

Capítulo

4

Elicitação

4.1 Introdução à Elicitação

A elicitação consiste no processo de extrair e formalizar o conhecimento de um especialista acerca de uma quantidade de interesse desconhecida, de modo que essa informação possa complementar os dados numéricos observados. No contexto da estatística bayesiana, a elicitação é essencial para a definição de uma distribuição a priori, a qual, combinada com o Teorema de Bayes e novas informações, resulta na distribuição a posteriori.

Uma das principais vantagens da abordagem bayesiana é justamente a capacidade de incorporar todas as informações disponíveis — tanto as provenientes dos dados observados quanto aquelas oriundas de conhecimento prévio. Essa integração permite inferências mais realistas e consistentes em comparação com métodos clássicos ou com o uso de prioris não-informativas. Conforme apresentado por Meyer e Booker (1991), existem diversas razões para utilizar a opinião a priori de um especialista:

- Em situações que envolvem fenômenos novos, raros ou complexos, a opinião de especialistas pode representar uma das poucas fontes de informação disponíveis. Um exemplo típico seria a modelagem de danos causados por terremotos;
- Quando é necessário tomar decisões sem a coleta de grandes quantidades de dados, a opinião do especialista pode suprir a insuficiência de informação estatística;
- A elicitação permite identificar quais aspectos do planejamento experimental apresentam maior incerteza, além de oferecer uma forma sistemática de documentar crenças, incorporar dados históricos e traduzir experiência prática que, muitas vezes, não se encontra registrada formalmente.

De forma geral, o processo de elicitação envolve dois elementos principais: o *especialista*, de quem se busca extrair o conhecimento, e o *analista*, responsável por traduzir essa expertise em uma distribuição de probabilidade adequada. O especialista deve possuir um conhecimento aprofundado sobre o objeto de estudo, enquanto o analista assegura que a informação fornecida seja representada corretamente em termos probabilísticos.

A elicitação desempenha um papel fundamental na inferência bayesiana, pois permite combinar a informação a priori sobre os parâmetros com os dados experimentais, produzindo uma estimativa mais atualizada da quantidade de interesse. Conforme O'Hagan (1998), "a elicitação será a chave para acessar essa informação adicional na análise bayesiana".

4.2 Classes Não-paramétricas de Prioris

Em diversos problemas práticos, a informação disponível sobre o parâmetro de interesse não é adequadamente representada por distribuições paramétricas usuais. Nesses casos, torna-se conveniente adotar abordagens não-paramétricas para a construção da distribuição a priori, permitindo maior flexibilidade na representação das crenças do especialista.

Gosling, Oakley e O'Hagan (2007) propõem uma metodologia de elicitação não-paramétrica baseada em processos gaussianos, na qual se assume um processo gaussiano a priori sobre uma função latente desconhecida $f(\theta)$. Essa função latente é utilizada para definir a distribuição a priori por meio de uma transformação exponencial seguida de normalização, garantindo que a densidade resultante seja positiva e integre-se a um.

Formalmente, assume-se

$$f(\theta) \mid m, v, b, \sigma^2 \sim GP(m(\theta), C(\theta, \phi)), \tag{4.2.1}$$

onde $m(\theta)$ representa a função média do processo e $C(\theta, \phi)$ sua função de covariância. A densidade a priori é então definida como

$$\pi(\theta) = \frac{\exp\{f(\theta)\}}{\int \exp\{f(u)\}\, du}, \tag{4.2.2}$$

o que assegura as propriedades fundamentais de uma função densidade de probabilidade.

No contexto da elicitação, a função média do processo gaussiano pode ser especificada a partir de uma densidade paramétrica inicial que reflita crenças centrais do especialista. Uma escolha comum consiste em adotar uma densidade normal como função média,

$$m(\theta) = g(\theta) = \frac{1}{\sqrt{2\pi v^2}} \exp\left\{-\frac{1}{2}\left(\frac{\theta - m}{v}\right)^2\right\}, \tag{4.2.3}$$

a qual fornece uma estrutura inicial para a distribuição a priori, enquanto a flexibilidade adicional é introduzida pelo processo gaussiano.

A função de covariância $C(\theta, \phi)$ controla a variabilidade e a suavidade da função latente, sendo frequentemente especificada como

$$C(\theta, \phi) = \sigma^2 c(\theta, \phi), \tag{4.2.4}$$

em que $c(\theta, \phi)$ é uma função de correlação positiva definida. Uma escolha amplamente utilizada é o núcleo exponencial quadrático,

$$c(\theta, \phi) = \exp\left\{-\frac{1}{2b}(\theta - \phi)^2\right\}, \tag{4.2.5}$$

o qual induz trajetórias suaves para a função latente $f(\theta)$.

Dado um conjunto de quantis ou percentis elicitados junto ao especialista, é possível obter a distribuição a posteriori da função latente e, consequentemente, da densidade a priori $\pi(\theta)$. A incerteza associada aos hiperparâmetros do processo gaussiano pode ser tratada por meio de métodos computacionais, em particular técnicas de Monte Carlo via Cadeias de Markov (MCMC).

A utilização de processos gaussianos na elicitação de distribuições a priori não-paramétricas fornece, assim, um arcabouço flexível e coerente para incorporar conhecimento especializado, permitindo capturar estruturas complexas que não seriam adequadamente representadas por modelos paramétricos tradicionais.

4.3 Modelos Gerais de Elicitação

Nesta seção, apresentam-se dois modelos clássicos amplamente discutidos na literatura de elicitação: o modelo linear normal e o modelo Beta-Binomial. Em ambos os casos, a estrutura probabilística é explicitamente definida, permitindo uma elicitação formal e coerente dos hiperparâmetros.

4.3.1 Modelo Linear Normal

Considere o modelo linear normal dado por

$$Y \mid X, \beta, \sigma^2 \sim N(X\beta, \sigma^2 I), \tag{4.3.1}$$

em que Y é o vetor de observações, X é a matriz de planejamento, β é o vetor de coeficientes de regressão e σ^2 é a variância do erro.

No contexto Bayesiano conjugado, assume-se a priori normal-inversa-gama para (β, σ^2):

$$\beta \mid \sigma^2 \sim N(b, \sigma^2 R^{-1}), \qquad \sigma^2 \sim IG\left(\frac{\delta}{2}, \frac{w\delta}{2}\right), \tag{4.3.2}$$

onde b, R, w e δ são hiperparâmetros.

A elicitação desses hiperparâmetros pode ser conduzida por meio de julgamentos sobre a distribuição preditiva de Y para valores específicos das covariáveis. Kadane (1980) demonstra que quantis e momentos da distribuição preditiva induzida pelo modelo podem ser utilizados para identificar valores plausíveis para (b, R, w, δ), evitando a interpretação direta desses parâmetros abstratos. Uma revisão detalhada desse enfoque preditivo é apresentada em Kadane e Wolfson (1998).

4.3.2 Modelo Beta-Binomial

Considere o modelo Binomial

$$Y \mid p \sim \text{Bin}(n, p), \tag{4.3.3}$$

em que p representa a probabilidade de sucesso. Uma escolha conjugada para a distribuição a priori de p é a distribuição Beta:

$$p \sim \text{Beta}(\alpha, \beta), \tag{4.3.4}$$

com densidade proporcional a $p^{\alpha-1}(1-p)^{\beta-1}$.

Nesse contexto, Chaloner e Duncan (1983) e Gavasakar (1988) propõem métodos formais de elicitação baseados em características interpretáveis da distribuição Beta, como a moda e a dispersão. Em particular, a moda da distribuição Beta é dada por

$$\text{Moda}(p) = \frac{\alpha - 1}{\alpha + \beta - 2}, \qquad \alpha, \beta > 1, \tag{4.3.5}$$

o que permite relacionar diretamente os julgamentos do especialista sobre o valor mais provável de p aos hiperparâmetros da priori.

A elicitação é conduzida de forma iterativa, ajustando os valores de (α, β) até que a distribuição a priori represente adequadamente as crenças expressas pelo especialista, preservando a conjugação e a coerência Bayesiana.

Outros Métodos de Elicitação

Outras abordagens amplamente discutidas na literatura incluem (cf. Cooke, 1991; Kadane e Wolfson, 1998; O'Hagan et al., 2006):

1. **Método da amostra a priori equivalente:** o especialista determina valores n e s tais que seu conhecimento seja equivalente a ter observado s sucessos em uma amostra aleatória de tamanho n, resultando em uma distribuição a priori Beta com parâmetros $(s, n - s)$;

2. **Método da função de distribuição acumulada:** o especialista fornece probabilidades acumuladas, como medianas e percentis, que são ajustadas a uma distribuição suave compatível com essas avaliações;

3. **Método da densidade de probabilidade:** solicita-se diretamente a forma da densidade, por exemplo por meio da moda e de valores próximos, permitindo a construção de uma densidade aproximada, embora esse método requeira maior familiaridade estatística por parte do especialista.

Capítulo

5

Distribuição Gumbel

5.1 Definição e aplicações

A distribuição de Valor Extremo do tipo I, mais conhecida como distribuição Gumbel, foi nomeada em homenagem a Emil Julius Gumbel (1891-1966), pioneiro na aplicação da Teoria do Valor Extremo. Essa distribuição é amplamente utilizada para modelar os valores máximos ou mínimos de um fenômeno de interesse em diferentes áreas do conhecimento, incluindo climatologia, hidrologia, geologia, ciências ambientais, finanças, seguros, engenharia civil, aeronáutica e naval conforme apresentado nos trabalhos clássicos de Gumbel (1958), Coles (2001) e Kotz e Nadarajah (2000).

Um exemplo clássico de aplicação é o estudo da precipitação máxima mensal, que é crucial para o planejamento e dimensionamento de projetos de engenharia hidráulica e agrícola, como canais de irrigação, drenagem, conservação de solos, estradas e barragens. Nesses casos, é necessário conhecer a distribuição dos valores extremos para evitar falhas ou subdimensionamento de estruturas críticas, conforme discutido em Coles (2001) e Embrechts et al. (1997).

Seja X uma variável aleatória seguindo a distribuição Gumbel. A função densidade de probabilidade (f.d.p.) é dada por:

$$f(x;\phi,\varphi) = \frac{1}{\varphi}\exp\left[-\frac{x-\phi}{\varphi} - \exp\left(-\frac{x-\phi}{\varphi}\right)\right], \quad x \in (-\infty,\infty), \tag{5.1.1}$$

onde $\phi \in (-\infty,\infty)$ é o parâmetro de locação e $\varphi > 0$ é o parâmetro de escala, responsáveis, respectivamente, por deslocar a distribuição ao longo do eixo horizontal e controlar a dispersão dos dados.

A Figura 5.1.1 apresenta a forma da densidade da distribuição Gumbel para diferentes combinações de ϕ e φ, evidenciando como esses parâmetros afetam a localização e a dispersão dos valores extremos.

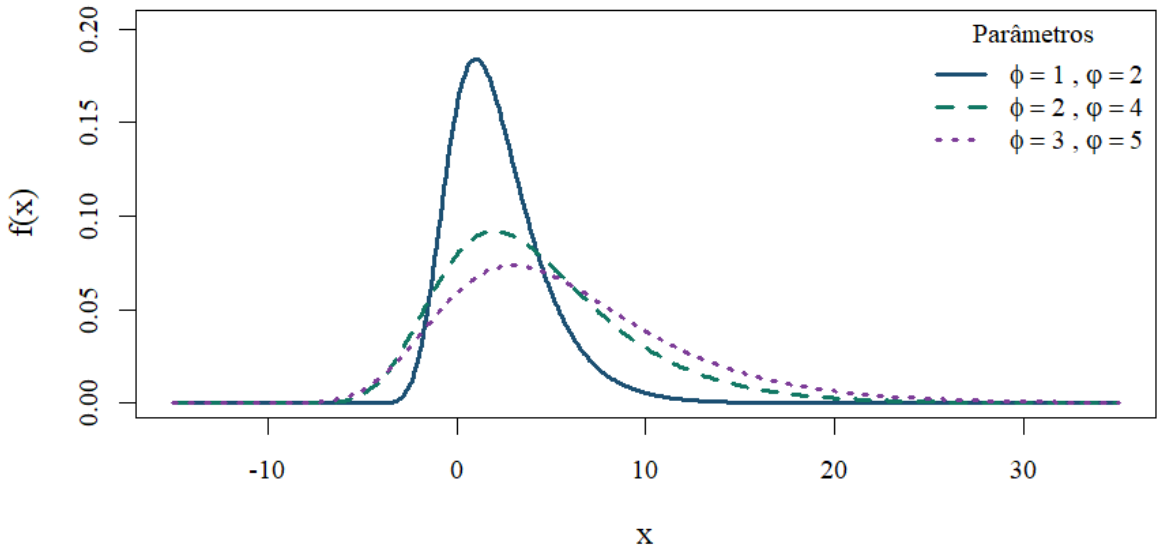


Figura 5.1.1: Densidade da Distribuição Gumbel para diferentes valores dos parâmetros ϕ e φ.

5.2 Propriedades da distribuição Gumbel

A distribuição Gumbel possui propriedades matemáticas e estatísticas importantes, que serão detalhadas a seguir.

As propriedades a seguir são bem estabelecidas na literatura clássica sobre distribuições de valores extremos, conforme Kotz e Nadarajah (2000) e Coles (2001).

P1: Função de distribuição acumulada (CDF). A função de distribuição acumulada, que fornece a probabilidade de que a variável aleatória X seja menor ou igual a x, é dada por:

$$F(x|\phi,\varphi) = \exp\left[-\exp\left(-\frac{x-\phi}{\varphi}\right)\right]. \tag{5.2.1}$$

P2: Esperança e variância. O valor esperado (média) e a variância da distribuição são expressos como:

$$E[X] = \phi + \varphi\gamma, \quad Var[X] = \frac{\pi^2}{6}\varphi^2, \tag{5.2.2}$$

onde $\gamma \approx 0.5772$ é a constante de Euler-Mascheroni. Esses parâmetros fornecem informações sobre a localização central e a dispersão da distribuição.

P3: Funções de confiabilidade e risco. A função de confiabilidade, ou função sobrevivência, expressa a probabilidade de que a variável X exceda um determinado valor x:

$$R(x) = 1 - F(x|\phi,\varphi) = 1 - \exp\left[-\exp\left(-\frac{x-\phi}{\varphi}\right)\right]. \tag{5.2.3}$$

A função de risco instantâneo, ou hazard, é dada por:

$$h(x) = \frac{1}{\varphi}\left[\exp\left(-\frac{x-\phi}{\varphi}\right) - 1 + \exp\left(\exp\left(-\frac{x-\phi}{\varphi}\right)\right)\right]. \tag{5.2.4}$$

P4: Momentos da distribuição. A função geradora de momentos (FGM) permite calcular os momentos de qualquer ordem da variável aleatória:

$$M_X(t) = \Gamma(1-\varphi t)e^{\phi t}, \tag{5.2.5}$$

onde $\Gamma(\cdot)$ é a função gama. Diferenciando $M_X(t)$ e avaliando em $t = 0$ obtém-se a média, variância e momentos centrais superiores.

P5: Transformação inversa da uniforme. Para gerar amostras da distribuição Gumbel a partir de uma variável uniforme $U \sim \mathcal{U}(0,1)$:

$$X = \phi - \varphi \log(-\log U). \tag{5.2.6}$$

5.3 Estimação de máxima verossimilhança

Seja $X_1, X_2, \ldots, X_n$ uma amostra de tamanho n. A função de verossimilhança é:

$$L(\phi, \varphi | x) = \prod_{i=1}^{n} \frac{1}{\varphi} \exp\left[-\frac{x_i-\phi}{\varphi} - \exp\left(-\frac{x_i-\phi}{\varphi}\right)\right]. \tag{5.3.1}$$

O logaritmo da verossimilhança simplifica os cálculos:

$$\log L(\phi, \varphi | x) = -n\log(\varphi) - \frac{1}{\varphi}\sum_{i=1}^{n}(x_i-\phi) - \sum_{i=1}^{n}\exp\left(-\frac{x_i-\phi}{\varphi}\right). \tag{5.3.2}$$

Derivando em relação a ϕ e φ e igualando a zero, obtemos as equações de máxima verossimilhança (EMV):

$$\frac{\partial \log L}{\partial \phi} = \frac{n}{\varphi} - \frac{1}{\varphi}\sum_{i=1}^{n}\exp\left(-\frac{x_i-\phi}{\varphi}\right) = 0, \tag{5.3.3}$$

$$\frac{\partial \log L}{\partial \varphi} = -\frac{n}{\varphi} + \frac{1}{\varphi^2}\sum_{i=1}^{n}(x_i-\phi) - \frac{1}{\varphi^2}\sum_{i=1}^{n}(x_i-\phi)\exp\left(-\frac{x_i-\phi}{\varphi}\right) = 0. \tag{5.3.4}$$

Essas equações não possuem solução fechada e exigem métodos numéricos. Após determinar $\hat{\varphi}$, substitui-se em $\partial \log L/\partial\phi = 0$ para obter $\hat{\phi}$.

O procedimento de estimação por máxima verossimilhança para a distribuição Gumbel é discutido em Coles (2001) e Davison (2003).

"'latex

5.4 Matriz de informação de Fisher e inferência frequentista

A matriz de informação de Fisher desempenha um papel fundamental na inferência frequentista, pois quantifica a quantidade de informação que a amostra fornece sobre os parâmetros desconhecidos. Para um vetor de parâmetros $\boldsymbol{\theta} = (\phi, \varphi)$, a matriz de informação de Fisher é definida por

$$\mathcal{I}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}\right], \tag{5.4.1}$$

onde $\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta}|x)$ representa a função de log-verossimilhança.

A matriz de informação de Fisher é obtida a partir das derivadas de segunda ordem da log-verossimilhança. Derivando novamente as equações de máxima verossimilhança em relação aos parâmetros ϕ e φ, obtêm-se as seguintes expressões:

$$\frac{\partial^2 \ell}{\partial \phi^2} = -\frac{1}{\varphi^2} \sum_{i=1}^{n} \exp\left(-\frac{x_i - \phi}{\varphi}\right), \tag{5.4.2}$$

$$\frac{\partial^2 \ell}{\partial \phi \partial \varphi} = -\frac{n}{\varphi^2} + \frac{1}{\varphi^2} \sum_{i=1}^{n} \exp\left(-\frac{x_i - \phi}{\varphi}\right) - \frac{1}{\varphi^3} \sum_{i=1}^{n} (x_i - \phi) \exp\left(-\frac{x_i - \phi}{\varphi}\right), \tag{5.4.3}$$

$$\frac{\partial^2 \ell}{\partial \varphi^2} = \frac{n}{\varphi^2} - \frac{2}{\varphi^3} \sum_{i=1}^{n} (x_i - \phi) + \frac{2}{\varphi^3} \sum_{i=1}^{n} (x_i - \phi) \exp\left(-\frac{x_i - \phi}{\varphi}\right) - \frac{1}{\varphi^4} \sum_{i=1}^{n} (x_i - \phi)^2 \exp\left(-\frac{x_i - \phi}{\varphi}\right). \tag{5.4.4}$$

Tomando a esperança matemática das derivadas de segunda ordem da log-verossimilhança, obtém-se a matriz de informação de Fisher da distribuição Gumbel,

$$\mathcal{I}(\phi, \varphi) = \frac{n}{\varphi^2} \begin{bmatrix} 1 & 1-\gamma \\ 1-\gamma & 1 - 2\gamma + \dfrac{\pi^2}{6} \end{bmatrix}, \tag{5.4.5}$$

onde $\gamma \approx 0.577216$ representa a constante de Euler–Mascheroni, a qual surge naturalmente durante a avaliação das esperanças das derivadas de segunda ordem da log-verossimilhança da distribuição Gumbel.

Os elementos da diagonal principal representam a quantidade de informação disponível sobre cada parâmetro individualmente. Em particular, o elemento $(1, 1)$ está associado ao parâmetro de localização ϕ, enquanto o elemento $(2, 2)$ corresponde ao parâmetro de escala φ. Os elementos fora da diagonal descrevem a dependência assintótica entre os estimadores desses parâmetros.

Além disso, sob condições regulares de regularidade, a inversa da matriz de informação de Fisher fornece uma aproximação assintótica para a matriz de variâncias e covariâncias dos estimadores de máxima verossimilhança. Consequentemente, essa matriz constitui a base para a construção de intervalos de confiança assintóticos e para a realização de testes de hipóteses sobre os parâmetros do modelo.

Para maiores detalhes sobre a obtenção da matriz de informação de Fisher para a distribuição Gumbel e suas propriedades assintóticas, recomenda-se consultar Coles (2001), Davison (2003), Casella e Berger (2002) e Cox e Hinkley (1974).

5.4.1 Intervalos de confiança

A matriz de informação de Fisher permite construir intervalos de confiança assintóticos para os parâmetros da distribuição Gumbel com base na distribuição assintótica dos estimadores de máxima verossimilhança. Sob condições regulares de regularidade e para amostras suficientemente grandes, os estimadores de máxima verossimilhança apresentam distribuição assintoticamente normal, isto é,

$$(\hat{\phi}, \hat{\varphi}) \overset{a}{\sim} N_2 \left[(\phi, \varphi), I^{-1}(\hat{\phi}, \hat{\varphi}) \right], \tag{5.4.6}$$

onde $I^{-1}(\hat{\phi}, \hat{\varphi})$ representa a inversa da matriz de informação de Fisher avaliada nos estimadores de máxima verossimilhança.

Assim, os intervalos de confiança assintóticos para os parâmetros são dados por

$$\hat{\phi} \pm z_{\alpha/2} \sqrt{\text{Var}(\hat{\phi})}, \tag{5.4.7}$$

$$\hat{\varphi} \pm z_{\alpha/2} \sqrt{\text{Var}(\hat{\varphi})}, \tag{5.4.8}$$

em que

$$\begin{bmatrix} \text{Var}(\hat{\phi}) & \text{Cov}(\hat{\phi}, \hat{\varphi}) \\ \text{Cov}(\hat{\phi}, \hat{\varphi}) & \text{Var}(\hat{\varphi}) \end{bmatrix} = I^{-1}(\hat{\phi}, \hat{\varphi}),$$

e $z_{\alpha/2}$ corresponde ao quantil da distribuição Normal padrão associado ao nível de confiança $100(1-\alpha)\%$.

Esses intervalos fornecem uma região plausível para os valores dos parâmetros, incorporando a variabilidade inerente ao processo de amostragem. Em particular, um intervalo de confiança de 95% indica que, em repetidas amostragens sob as mesmas condições experimentais, aproximadamente 95% dos intervalos construídos conterão o verdadeiro valor dos parâmetros.

5.5 Inferência Bayesiana com prioris não-informativas e fracamente informativas

Na inferência Bayesiana, a distribuição a priori representa o conhecimento prévio sobre os parâmetros antes da observação dos dados, conforme Bernardo e Smith (1994) e Gelman et al. (2014). Quando se deseja adotar uma abordagem conservadora, minimizando o impacto da priori na análise, utilizam-se prioris não-informativas, que permitem que a informação contida nos dados domine a inferência.

Para a distribuição Gumbel, consideramos dois parâmetros: o parâmetro de localização ϕ e o de escala φ. O objetivo é definir prioris que não introduzam informações fortes, mas que sejam matematicamente consistentes e que permitam derivar uma distribuição a posteriori adequada.

As informações a seguir sobre as prioris podem ser encontradas em Kass e Wasserman (1996) e Gelman et al. (2014).

5.5.1 Estrutura de priori independente para ϕ e φ

Embora a priori de Jeffreys apresente propriedades teóricas importantes, uma abordagem frequentemente utilizada em aplicações Bayesianas consiste em assumir independência a priori entre os parâmetros de localização e escala. Essa estratégia simplifica a modelagem, facilita a implementação computacional e permite especificar separadamente o conhecimento prévio sobre cada parâmetro.

Assim, considera-se

$$\phi \sim \mathcal{N}(a, b^2), \qquad \varphi \sim \text{Gamma}(c, d), \tag{5.5.1}$$

onde a, b, c e d representam hiperparâmetros previamente especificados.

Sob a hipótese de independência, a distribuição a priori conjunta é dada por

$$\pi(\phi, \varphi) = \pi_{\phi}(\phi)\pi_{\varphi}(\varphi). \tag{5.5.2}$$

Consequentemente, utilizando o Teorema de Bayes, a distribuição a posteriori pode ser escrita como

$$p(\phi, \varphi|x) \propto L(\phi, \varphi|x)\pi_{\phi}(\phi)\pi_{\varphi}(\varphi), \tag{5.5.3}$$

em que $L(\phi, \varphi|x)$ representa a função de verossimilhança da distribuição Gumbel.

Essa estrutura foi adotada ao longo deste trabalho, pois permite controlar separadamente a informação prévia atribuída aos parâmetros de localização e escala, além de facilitar a implementação computacional do algoritmo de Metropolis-Hastings utilizado na obtenção das distribuições a posteriori.

5.5.2 Priori de Jeffreys (não informativa)

A priori de Jeffreys constitui uma das prioris objetivas mais conhecidas da inferência Bayesiana. Sua principal característica é a invariância sob reparametrizações, isto é, sua forma depende exclusivamente da estrutura probabilística do modelo e permanece inalterada diante de mudanças de parametrização.

Formalmente, a priori de Jeffreys é definida por

$$\pi_J(\phi, \varphi) \propto \sqrt{\det\left[I(\phi, \varphi)\right]}, \tag{5.5.4}$$

onde $I(\phi, \varphi)$ representa a matriz de informação de Fisher apresentada na Equação (5.4.5).

Para a distribuição Gumbel, o determinante da matriz de informação de Fisher pode ser escrito na forma

$$\det\left[I(\phi, \varphi)\right] = \frac{n^2}{\varphi^4}\, C, \tag{5.5.5}$$

em que C representa uma constante positiva que não depende dos parâmetros do modelo.

Assim,

$$\sqrt{\det\left[I(\phi, \varphi)\right]} = \frac{n}{\varphi^2}\sqrt{C} \propto \frac{1}{\varphi^2},$$

obtendo-se a priori de Jeffreys

$$\pi_J(\phi, \varphi) \propto \frac{1}{\varphi^2}. \tag{5.5.6}$$

Observa-se que essa priori não incorpora informações provenientes de especialistas ou de estudos anteriores, sendo determinada exclusivamente pela geometria da função de verossimilhança. Por essa razão, é amplamente utilizada como uma priori objetiva ou não informativa em problemas de inferência Bayesiana.

Embora apresente propriedades teóricas bastante atrativas, sua utilização em modelos multiparamétricos requer alguns cuidados, especialmente quando coexistem parâmetros de localização e escala. Nesses casos, é comum adotar estruturas de prioris independentes, tanto por razões de simplicidade analítica quanto pela facilidade de implementação computacional em algoritmos MCMC.

5.5.3 Priori Flat (não informativa uniforme)

Uma priori flat ou uniforme busca minimizar a influência do conhecimento prévio. Para os parâmetros da Gumbel, podemos adotar:

$$\pi_F(\phi, \varphi) = \pi_\phi(\phi|0, 1000^2)\, \pi_\varphi(\varphi|0.001, 0.001), \tag{5.5.7}$$

onde a variância extremamente grande para ϕ e os parâmetros pequenos para φ tornam a distribuição praticamente uniforme sobre a faixa de valores plausíveis, garantindo que a posteriori seja dominada pelos dados.

5.5.4 Priori fracamente informativa

As prioris fracamente informativas são especificadas de modo a introduzir apenas um grau mínimo de regularização nos parâmetros, sem impor crenças fortes a priori. No caso da distribuição Gumbel, pode-se adotar, por exemplo,

$$\phi \sim \mathcal{N}(a_\phi,\, b_\phi^2), \qquad \varphi \sim \text{Gamma}(c, d),$$

onde a_ϕ representa uma média a priori neutra (por exemplo, centrada em zero ou em torno do valor médio dos dados) e b_ϕ^2 é escolhido suficientemente grande para produzir uma distribuição ampla.

De modo análogo, os hiperparâmetros c e d da distribuição Gamma são selecionados de forma a gerar uma variância elevada, refletindo incerteza considerável sobre o parâmetro de escala φ. Essas escolhas garantem que a priori tenha caudas mais leves que uma distribuição não informativa, mas ainda suficientemente amplas para permitir flexibilidade inferencial.

5.5.5 Posteriori a partir de prioris não-informativas

Com qualquer uma das prioris acima, a distribuição a posteriori é obtida via regra de Bayes:

$$p(\phi, \varphi|x) \propto L(\phi, \varphi|x)\, \pi(\phi, \varphi), \tag{5.5.8}$$

onde $L(\phi, \varphi|x)$ é a função de verossimilhança da Gumbel. Cada escolha de priori gera uma posteriori diferente, refletindo diferentes suposições iniciais sobre os parâmetros.

A escolha da distribuição a priori desempenha um papel central na inferência Bayesiana, influenciando diretamente a forma como a informação dos dados é combinada com o conhecimento prévio disponível. As diferentes classes de prioris apresentadas nesta seção refletem distintas filosofias de modelagem. A priori de Jeffreys, por exemplo, possui uma formulação teoricamente elegante e invariável a transformações reparametrizáveis, porém tende a induzir dependência entre parâmetros de natureza distinta, como os de locação e escala, o que pode gerar interpretações menos intuitivas.

Por outro lado, a utilização de prioris independentes representa uma alternativa prática e flexível, permitindo controlar separadamente o comportamento de cada parâmetro e evitando possíveis correlações artificiais introduzidas pelo modelo. Já a priori flat, ou não informativa, caracteriza-se por seu caráter extremamente conservador, atribuindo variância muito elevada e deixando que os dados observados exerçam a influência predominante na determinação da distribuição a posteriori.

Em contrapartida, as prioris fracamente informativas (*weakly informative*) surgem como um meio-termo entre a neutralidade extrema e a necessidade de estabilidade numérica. Elas introduzem uma leve regularização nos parâmetros, prevenindo estimativas não realistas e facilitando a convergência de métodos numéricos, sem impor crenças fortes sobre o modelo.

Em síntese, mesmo ao se adotar prioris essencialmente não informativas, é possível conduzir uma análise Bayesiana robusta, coerente e interpretável, equilibrando o papel da informação empírica dos dados com a estrutura teórica subjacente ao modelo estatístico.

Capítulo

6

Elicitação de priori via preditiva a priori sob aproximação de Laplace

Nesta seção, emprega-se a aproximação de Laplace, conforme a formulação proposta por Kass, Tierney e Kadane (1986), no processo de elicitação da distribuição a priori. Nesse procedimento, a informação fornecida pelo especialista é tratada como fonte de dados e deve ser compatível com a densidade preditiva a priori. A aproximação de Laplace é então utilizada para obter uma expressão analítica aproximada dessas probabilidades, permitindo a determinação dos hiperparâmetros.

6.1 O método

Considere uma variável aleatória T com distribuição Valor Extremo do tipo I (Gumbel), com parâmetros de locação $\phi \in \mathbb{R}$ e escala $\varphi > 0$, cuja função de distribuição acumulada é dada por

$$F(t \mid \phi, \varphi) = \exp\left\{-\exp\left(-\frac{t-\phi}{\varphi}\right)\right\}, \qquad t \in \mathbb{R}. \tag{6.1.1}$$

Assume-se independência a priori entre os parâmetros ϕ e φ, sendo

$$\phi \sim N(a, b^2), \qquad \varphi \sim \text{Gamma}(c, d), \tag{6.1.2}$$

onde a distribuição Gamma é parametrizada em termos do parâmetro de taxa $d > 0$, com densidade

$$\pi(\varphi) = \frac{d^c}{\Gamma(c)}\varphi^{c-1}\exp(-d\varphi), \qquad \varphi > 0. \tag{6.1.3}$$

Assim, a densidade a priori conjunta é expressa por

$$\pi(\phi, \varphi) = \frac{1}{\sqrt{2\pi b^2}}\exp\left\{-\frac{(\phi-a)^2}{2b^2}\right\} \cdot \frac{d^c}{\Gamma(c)}\varphi^{c-1}\exp(-d\varphi). \tag{6.1.4}$$

Seja (t_i, p_i), $i = 1, \ldots, 4$, um conjunto de percentis elicitados por um especialista, tais que

$$P(T \leq t_i) = p_i. \tag{6.1.5}$$

Essas probabilidades são interpretadas como probabilidades preditivas a priori e podem ser escritas como

$$P(T \leq t_i) = \int_{-\infty}^{\infty} \int_{0}^{\infty} F(t_i \mid \phi, \varphi)\, \pi(\phi, \varphi)\, d\varphi\, d\phi. \tag{6.1.6}$$

Substituindo as expressões de $F(t_i \mid \phi, \varphi)$ e $\pi(\phi, \varphi)$, obtemos

$$P(T \leq t_i) = \frac{d^c}{\Gamma(c)\sqrt{2\pi b^2}} \int_{-\infty}^{\infty} \int_{0}^{\infty} \exp\left\{-\exp\left(-\frac{t_i - \phi}{\varphi}\right)\right\} \varphi^{c-1} \times \exp\left\{-\frac{(\phi - a)^2}{2b^2} - d\varphi\right\} d\varphi\, d\phi. \tag{6.1.7}$$

A integral acima não admite forma analítica fechada. Para aplicar a aproximação de Laplace conforme apresentada na Seção 3.1, definimos

$$h_i(\phi, \varphi) = \frac{(\phi - a)^2}{2b^2} + d\varphi + \exp\left(-\frac{t_i - \phi}{\varphi}\right) - (c - 1)\ln(\varphi). \tag{6.1.8}$$

Assim, a probabilidade preditiva pode ser escrita como

$$P(T \leq t_i) = \frac{d^c}{\Gamma(c)\sqrt{2\pi b^2}} \int_{-\infty}^{\infty} \int_{0}^{\infty} \exp\{-h_i(\phi, \varphi)\}\, d\varphi\, d\phi. \tag{6.1.9}$$

A aproximação de Laplace baseia-se na expansão da função $h_i(\phi, \varphi)$ em torno do seu ponto crítico. Esse ponto é obtido resolvendo

$$\nabla h_i(\phi, \varphi) = \mathbf{0}. \tag{6.1.10}$$

Calculando as derivadas parciais, temos

$$\frac{\partial h_i}{\partial \phi} = \frac{1}{\varphi} \exp\left(-\frac{t_i - \phi}{\varphi}\right) + \frac{\phi - a}{b^2} = 0, \tag{6.1.11}$$

$$\frac{\partial h_i}{\partial \varphi} = \frac{t_i - \phi}{\varphi^2} \exp\left(-\frac{t_i - \phi}{\varphi}\right) + d - \frac{c - 1}{\varphi} = 0. \tag{6.1.12}$$

Esse sistema depende do quantil elicitado t_i e dos hiperparâmetros (a, b, c, d) e, em geral, não admite solução analítica fechada, devendo ser resolvido numericamente para determinar o ponto estacionário $(\hat{\phi}_i, \hat{\varphi}_i)$.

A matriz Hessiana de $h_i(\phi, \varphi)$ é definida por

$$\mathrm{H}_i(\phi, \varphi) = \begin{pmatrix} \dfrac{\partial^2 h_i}{\partial \phi^2} & \dfrac{\partial^2 h_i}{\partial \phi \partial \varphi} \\ \dfrac{\partial^2 h_i}{\partial \varphi \partial \phi} & \dfrac{\partial^2 h_i}{\partial \varphi^2} \end{pmatrix}. \tag{6.1.13}$$

Os elementos dessa matriz são

$$\frac{\partial^2 h_i}{\partial \phi^2} = \frac{1}{\varphi^2} \exp\left(-\frac{t_i - \phi}{\varphi}\right) + \frac{1}{b^2}, \tag{6.1.14}$$

$$\frac{\partial^2 h_i}{\partial \phi \partial \varphi} = \left(\frac{t_i - \phi}{\varphi^3} - \frac{1}{\varphi^2}\right) \exp\left(-\frac{t_i - \phi}{\varphi}\right), \tag{6.1.15}$$

$$\frac{\partial^2 h_i}{\partial \varphi^2} = \left[\frac{(t_i - \phi)^2}{\varphi^4} - \frac{2(t_i - \phi)}{\varphi^3}\right] \exp\left(-\frac{t_i - \phi}{\varphi}\right) + \frac{c - 1}{\varphi^2}. \tag{6.1.16}$$

Essa matriz é avaliada no ponto estacionário $(\hat{\phi}_i, \hat{\varphi}_i)$.

6.2 Aplicação da aproximação de Laplace

A aproximação de Laplace para integrais da forma

$$\int \exp\{-h(\boldsymbol{\theta})\}\, d\boldsymbol{\theta}$$

foi apresentada na Seção 3.1 e fornece a expressão geral dada em (3.1.5).

No presente problema, a função $h_i(\phi, \varphi)$ já foi definida anteriormente e o vetor de parâmetros é $\boldsymbol{\theta} = (\phi, \varphi)^\top$, de modo que a dimensão do problema é $m = 2$. Assim, aplicando a aproximação de Laplace proposta por Kass, Tierney e Kadane (1986), obtém-se uma aproximação para a probabilidade preditiva

$$P(T \leq t_i) \approx \frac{d^c}{\Gamma(c)\sqrt{2\pi b^2}} (2\pi)^{\frac{m}{2}} \left|\mathrm{H}_i(\hat{\phi}_i, \hat{\varphi}_i)\right|^{-1/2} \exp\left\{-h_i(\hat{\phi}_i, \hat{\varphi}_i)\right\}, \qquad i = 1, \ldots, 4. \tag{6.2.1}$$

Como $m = 2$, tem-se $(2\pi)^{m/2} = 2\pi$.

Por definição, a função de distribuição acumulada associada à variável aleatória T é dada por

$$F(t_i) = P(T \leq t_i).$$

Portanto, a aproximação de Laplace fornece a seguinte expressão aproximada para a função de distribuição preditiva

$$F(t_i) \approx \frac{d^c}{\Gamma(c)\sqrt{2\pi b^2}} (2\pi) \left|\mathrm{H}_i(\hat{\phi}_i, \hat{\varphi}_i)\right|^{-1/2} \exp\left\{-h_i(\hat{\phi}_i, \hat{\varphi}_i)\right\}. \tag{6.2.2}$$

Nessa expressão, $(\hat{\phi}_i, \hat{\varphi}_i)$ representa o ponto estacionário da função $h_i(\phi, \varphi)$, obtido pela solução numérica do sistema

$$\nabla h_i(\phi, \varphi) = \mathbf{0}.$$

Igualando $F(t_i)$ às probabilidades elicitadas p_i, obtém-se um sistema não linear nas incógnitas (a, b, c, d), cuja solução determina os hiperparâmetros da distribuição a priori elicitada.

A função de confiabilidade correspondente é

$$R(t_i) = 1 - F(t_i). \tag{6.2.3}$$

6.3 Obtenção dos hiperparâmetros via aproximação de Laplace (1986)

Nesta seção apresenta-se o procedimento utilizado para a obtenção dos hiperparâmetros das distribuições a priori por meio da aproximação de Laplace proposta por Tierney e Kadane (1986). Considera-se a estrutura de priori independente apresentada na Seção 5.5.1, na qual o parâmetro de localização é modelado por $\phi \sim N(a, b^2)$ e o parâmetro de escala por $\varphi \sim \text{Gamma}(c, d)$. Nessa parametrização, os hiperparâmetros a e b correspondem, respectivamente, à média e ao desvio-padrão da distribuição Normal atribuída a ϕ, enquanto c e d representam os parâmetros de forma e taxa da distribuição Gamma associada a φ.

O objetivo consiste em obter os valores dos hiperparâmetros (a, b, c, d) de forma que a distribuição preditiva a priori reproduza, o mais próximo possível, as probabilidades especificadas pelo especialista.

Inicialmente, considera-se que a variável aleatória T segue uma distribuição Gumbel com parâmetros verdadeiros $\phi = 0.5$ e $\varphi = 0.5$. A partir desses valores foram obtidos os quantis correspondentes às probabilidades

$$p = (0{,}75,\ 0{,}50,\ 0{,}25,\ 0{,}10),$$

definidas em ordem decrescente. Esses quantis foram calculados por meio da função quantil da distribuição Gumbel,

$$t_p = \phi - \varphi \log\big[-\log(p)\big].$$

Com o objetivo de avaliar a robustez do procedimento de eliciação, foram considerados três cenários distintos:

- **Elicitação nominal:** os quantis são obtidos diretamente a partir dos parâmetros verdadeiros do modelo;
- **Elicitada I:** erro de aproximadamente -10% em relação aos quantis verdadeiros;
- **Elicitada II:** erro de aproximadamente -15% em relação aos quantis verdadeiros.

Os quantis utilizados em cada cenário são apresentados na Tabela 6.3.1.

Tabela 6.3.1: Quantis utilizados nos cenários de eliciação.

Probabilidade	Verdadeiro	Elicitada I	Elicitada II
0,75	1,12294966	1,01065470	0,95450721
0,50	0,68325646	0,61493080	0,58076799
0,25	0,33668287	0,30301460	0,28618044
0,10	0,08298378	0,07468540	0,07053621

A obtenção dos hiperparâmetros foi realizada resolvendo-se numericamente um sistema de equações não lineares que relaciona as probabilidades preditivas a priori com os quantis especificados pelo especialista. Como esse sistema não admite solução analítica fechada, foi empregado um procedimento de otimização numérica implementado no software *R*. Esse procedimento permitiu determinar os valores dos hiperparâmetros (a, b, c, d) que melhor reproduzem, por meio da distribuição preditiva a priori, as probabilidades estabelecidas no processo de eliciação.

Os hiperparâmetros obtidos para cada cenário são apresentados na Tabela 6.3.2.

Tabela 6.3.2: Hiperparâmetros obtidos via aproximação de Laplace (1986).

Cenário	a	b	c	d
Verdadeiro	0,4970465	0,2031098	4,2217382	6,7369531
Elicitada I	0,5109335	0,3335553	4,2823007	12,9687762
Elicitada II	0,4540676	0,2855911	4,2629569	12,9750141

Observa-se que cada cenário de eliciação conduz a um conjunto distinto de hiperparâmetros. Em particular, os hiperparâmetros a e b caracterizam a distribuição Normal utilizada para representar o parâmetro de localização ϕ, enquanto c e d caracterizam a distribuição Gamma atribuída ao parâmetro de escala φ. Dessa forma, os valores apresentados na Tabela 6.3.2 refletem a adaptação das distribuições a priori às diferentes informações fornecidas pelo especialista em cada cenário considerado.

Para fins de simplificação na apresentação dos gráficos e tabelas dos estudos de simulação, a priori fracamente informativa será identificada pela denominação *Weak*, preservando-se a mesma interpretação ao longo deste trabalho.

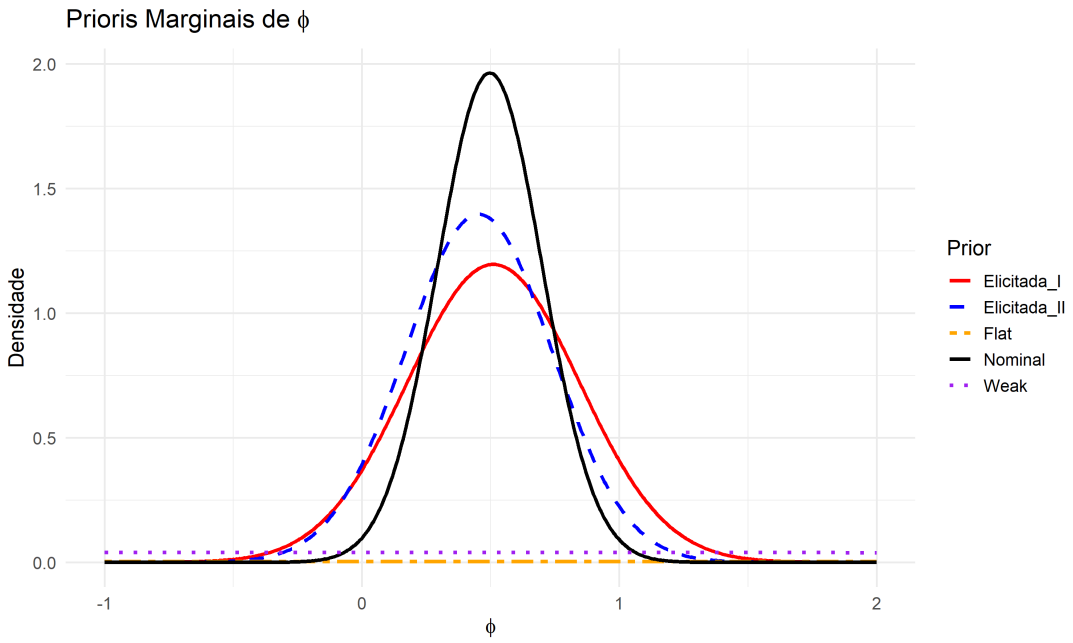


Figura 6.3.1: Distribuições a priori do parâmetro ϕ.

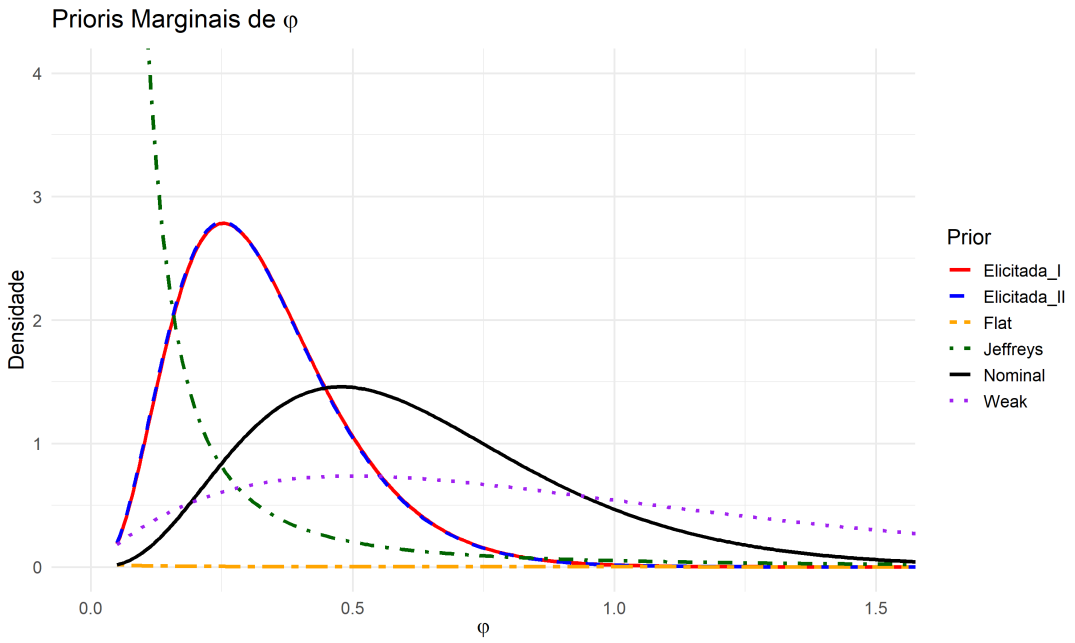


Figura 6.3.2: Distribuições a priori do parâmetro φ.

A análise dos resultados mostra que cada cenário de eliciação produz um conjunto distinto de hiperparâmetros, refletindo as diferentes informações incorporadas na construção das distribuições a priori. Conforme apresentado na Tabela 6.3.2, pequenas alterações nos quantis elicidados resultam em modificações nos hiperparâmetros estimados, evidenciando a sensibilidade esperada do processo de eliciação às informações fornecidas pelo especialista.

As Figuras 6.3.1 e 6.3.2 permitem comparar graficamente as distribuições a priori obtidas para cada cenário. Observa-se que as prioris Elicitada I e Elicitada II apresentam comportamento semelhante ao da priori nominal, sendo que a Elicitada I exibe maior sobreposição visual com a distribuição de referência. Para o parâmetro de localização ϕ, essa diferença é ligeiramente mais perceptível, enquanto, para o parâmetro de escala φ, as curvas apresentam elevada sobreposição, indicando diferenças visuais pouco pronunciadas entre os cenários de eliciação.

Em contraste, as prioris Jeffreys, Flat e Weak apresentam comportamento gráfico mais distinto em relação à priori nominal, como esperado, por não incorporarem as mesmas informações prévias utilizadas no processo de eliciação. De modo geral, os resultados mostram que perturbações moderadas nos quantis especificados pelo especialista produzem alterações correspondentes nas distribuições a priori, preservando, entretanto, a estrutura geral das distribuições obtidas nos cenários elicidados.

Capítulo

7

Estudo de Simulação

Neste capítulo é apresentado um estudo de simulação de Monte Carlo com o objetivo de avaliar o desempenho do procedimento de inferência Bayesiana proposto para a distribuição Gumbel. Em particular, busca-se comparar o comportamento das prioris elicitadas por meio da aproximação de Laplace de Tierney e Kadane (1986) com prioris não informativas e fracamente informativas, considerando cenários de pequena dimensão amostral, nos quais a informação a priori exerce maior influência sobre a inferência.

7.1 Desenho do Estudo de Simulação

O estudo de simulação foi conduzido considerando como valores verdadeiros dos parâmetros da distribuição Gumbel $\phi = 0,5$ (parâmetro de localização) e $\varphi = 0,5$ (parâmetro de escala).

Para cada tamanho amostral considerado (n = 3, 6 e 10), foram geradas $N = 500$ amostras independentes da distribuição Gumbel. Cada uma dessas amostras foi utilizada para realizar uma inferência Bayesiana independente, permitindo avaliar o comportamento dos estimadores sob repetidas amostragens.

Em cada uma das 500 amostras simuladas foi implementado um algoritmo de Monte Carlo via Cadeias de Markov (MCMC), composto por $R = 11.000$ iterações. As primeiras 1.000 iterações foram descartadas como período de *burn-in*, com o objetivo de reduzir a influência dos valores iniciais da cadeia. Em seguida, foi aplicado um procedimento de *thinning* de ordem 10, mantendo-se uma observação a cada dez iterações, com o propósito de reduzir a autocorrelação entre as amostras geradas.

As amostras foram simuladas a partir da distribuição Gumbel utilizando o método da transformação inversa, garantindo consistência com o modelo probabilístico adotado ao longo deste trabalho.

Foram consideradas seis especificações de distribuições a priori: a priori elicitada correspondente ao cenário verdadeiro (Elicitada Nominal), duas versões obtidas a partir de perturbações nos quantis elicitados (Elicitada I e Elicitada II), além das prioris Jeffreys, Flat e Weak.

Para cada combinação entre tamanho amostral e distribuição a priori foram calculadas as medidas de desempenho viés (*Bias*), erro absoluto médio (EAM), erro quadrático médio (EQM) e probabilidade de cobertura dos intervalos de credibilidade. Essas medidas permitem avaliar, respectivamente, a presença de erro sistemático nas estimativas, a magnitude média dos erros de estimação, a precisão global dos estimadores

e a capacidade dos intervalos de credibilidade em conter os verdadeiros valores dos parâmetros.

7.2 Implementação do Método de Monte Carlo via Cadeias de Markov (MCMC)

Para a atualização bayesiana e obtenção das estimativas das distribuições a posteriori, utilizou-se o método de Monte Carlo via Cadeias de Markov (MCMC) (Metropolis et al., 1953; Hastings, 1970). Este método é fundamental quando a distribuição resultante da combinação entre a priori elicitada e a função de verossimilhança não possui uma forma analítica fechada, o que ocorre frequentemente na modelagem de valores extremos.

O MCMC opera gerando uma sequência de valores (uma cadeia) de modo que, após um número suficiente de iterações, a distribuição desses valores converge para a distribuição a posteriori desejada. No contexto desta simulação, o procedimento permitiu extrair amostras dos parâmetros ϕ e φ da distribuição Gumbel, possibilitando o cálculo de estimativas pontuais e intervalos de credibilidade.

"'latex A implementação do algoritmo considerou um período de aquecimento (*burn-in*) de 1.000 iterações, utilizado para reduzir a influência dos valores iniciais da cadeia, seguido da aplicação de um procedimento de *thinning* de ordem 10, com o objetivo de diminuir a autocorrelação entre as amostras sucessivas. A avaliação da convergência das cadeias foi realizada por meio da inspeção gráfica dos traceplots, das funções de autocorrelação (ACF) e das distribuições marginais a posteriori, verificando-se a estabilização das cadeias após o período de aquecimento e a adequada exploração do espaço paramétrico pelos parâmetros ϕ e φ. Esses diagnósticos forneceram evidências da convergência das cadeias e da adequação das amostras utilizadas na inferência Bayesiana.

A atualização dos parâmetros foi realizada por meio do algoritmo de Metropolis-Hastings, utilizado para gerar amostras da distribuição a posteriori conjunta de ϕ e φ. O procedimento consistiu em atualizações sucessivas dos parâmetros por meio de distribuições propostas e decisões de aceitação ou rejeição baseadas nas respectivas razões de aceitação. No caso do parâmetro de escala φ, a formulação incorporou uma transformação logarítmica, com o correspondente ajuste jacobiano. O esquema geral de atualização utilizado é apresentado no Algoritmo 1.

Algorithm 1 Algoritmo Metropolis-Hastings para ϕ e φ

1: Inicializar $\phi^{(0)}$ e $\varphi^{(0)}$
2: **for** $i = 1, \ldots, R-1$ **do**
3: Gerar proposta

$$\phi^* \sim N(\phi^{(i-1)}, \sigma^2_\phi)$$

4: Calcular

$$\alpha_\phi = \min\left(1, \frac{\pi(\phi^*, \varphi^{(i-1)}|x)}{\pi(\phi^{(i-1)}, \varphi^{(i-1)}|x)}\right)$$

5: Aceitar ou rejeitar ϕ^*
6: Gerar

$$w^* \sim N(\log \varphi^{(i-1)}, \sigma^2_\varphi)$$

$$\varphi^* = \exp(w^*)$$

7: Calcular

$$\alpha_\varphi = \min\left(1, \frac{\pi(\phi^{(i)}, \varphi^*|x)\varphi^*}{\pi(\phi^{(i)}, \varphi^{(i-1)}|x)\varphi^{(i-1)}}\right)$$

8: Aceitar ou rejeitar φ^*
9: **end for**
10: Descartar burn-in
11: Aplicar thinning

7.3 Resultados do Estudo de Simulação

Nesta seção são apresentados os resultados do estudo de simulação considerando o procedimento de elicitação baseado na aproximação de Laplace proposta por Kass, Tierney e Kadane (1986).

Tabela 7.3.1: Resultados da Simulação MCMC (1.000 Réplicas): Viés, EAM, EQM e Cobertura para ϕ e φ

		Parâmetro ϕ (Localização)				Parâmetro φ (Escala)			
n	**Priori**	**Viés**	**EAM**	**EQM**	**PC (95%)**	**Viés**	**EAM**	**EQM**	**PC (95%)**
3	Nominal	0,0018	0,0731	0,0079	1,000	0,0555	0,1079	0,0187	0,998
	Elicitada I	0,0020	0,1603	0,0401	0,970	-0,1143	0,1272	0,0224	0,936
	Elicitada II	-0,0224	0,1434	0,0310	0,978	-0,1150	0,1277	0,0225	0,952
	Jeffreys	-0,0030	0,2465	0,0987	0,854	-0,0065	0,2280	0,0852	0,876
	Flat	-0,0218	0,2533	0,1091	0,946	0,3046	0,4159	0,3360	0,950
	Frac. Inf.	-0,0009	0,2449	0,0973	0,988	0,2262	0,2632	0,1044	0,970
6	Nominal	0,0034	0,0836	0,0111	0,994	0,0386	0,0999	0,0155	0,984
	Elicitada I	-0,0046	0,1389	0,0304	0,920	-0,0723	0,0965	0,0141	0,942
	Elicitada II	-0,0218	0,1301	0,0261	0,922	-0,0752	0,0976	0,0143	0,932
	Jeffreys	0,0033	0,1857	0,0533	0,900	-0,0095	0,1430	0,0316	0,908
	Flat	0,0173	0,1856	0,0542	0,916	0,0690	0,1702	0,0479	0,956
	Frac. Inf.	0,0260	0,1841	0,0541	0,950	0,1117	0,1674	0,0445	0,966
10	Nominal	0,0085	0,0742	0,0087	0,988	0,0331	0,0885	0,0126	0,974
	Elicitada I	-0,0019	0,1050	0,0172	0,956	-0,0431	0,0791	0,0095	0,930
	Elicitada II	-0,0137	0,0992	0,0154	0,960	-0,0458	0,0797	0,0096	0,926
	Jeffreys	0,0072	0,1293	0,0262	0,952	-0,0005	0,1053	0,0179	0,922
	Flat	0,0190	0,1301	0,0267	0,956	0,0398	0,1165	0,0226	0,944
	Frac. Inf.	0,0265	0,1303	0,0270	0,962	0,0707	0,1214	0,0241	0,960

Análise dos Resultados da Simulação

A Tabela 7.3.1 apresenta os resultados do estudo de simulação em termos de viés, erro quadrático médio (EQM) e probabilidade de cobertura (PC) dos intervalos de credibilidade de 95%, para diferentes tamanhos amostrais e especificações de priori. A seguir, discute-se o desempenho dos estimadores em cada cenário.

No caso de amostra extremamente reduzida ($n = 3$), a priori elicitada nominal apresenta desempenho superior para o parâmetro de localização ϕ, com viés praticamente nulo, menor EQM e cobertura perfeita (PC = 1.000), evidenciando a consistência da elicitação quando a informação do especialista é corretamente especificada. As prioris elicitadas (I e II) também apresentam bom desempenho, embora com maior EQM e leve deterioração na cobertura, especialmente para o parâmetro de escala φ, onde se observa viés negativo mais acentuado.

Em contraste, as prioris não informativas (Jeffreys e Flat) exibem maior variabilidade, com EQMs elevados e coberturas menos estáveis, particularmente para φ. A priori fracamente informativa apresenta desempenho intermediário, com boa cobertura, porém ainda com variabilidade relativamente alta. Esses resultados destacam a importância da incorporação de informação prévia em cenários com poucos dados.

Para $n = 6$, observa-se melhora geral no desempenho dos estimadores. A priori nominal mantém sua superioridade, com baixos níveis de viés e EQM e cobertura

próxima do nominal. As prioris elicitadas (I e II) tornam-se mais precisas, com redução significativa do EQM, embora ainda apresentem leve viés em φ.

As prioris não informativas também melhoram, com redução da variabilidade e aumento da cobertura, mas ainda permanecem inferiores às prioris informativas. A priori fracamente informativa destaca-se por apresentar um bom equilíbrio entre viés, EQM e cobertura, aproximando-se do desempenho das prioris elicitadas.

Para $n = 10$, observa-se convergência entre os resultados das diferentes prioris, refletindo a predominância da informação amostral. A priori nominal continua apresentando excelente desempenho, enquanto as prioris elicitadas (I e II) tornam-se ainda mais precisas, com redução adicional do EQM e melhora na cobertura. O viés em φ diminui consideravelmente.

As prioris não informativas passam a apresentar desempenho semelhante em termos de cobertura, embora ainda exibam maior variabilidade em alguns casos. A priori fracamente informativa mantém comportamento estável, constituindo uma alternativa robusta quando há incerteza na especificação da priori.

De forma geral, os resultados indicam que a metodologia de elicitação baseada na aproximação de Laplace (1986) é eficaz na construção de prioris informativas que melhoram significativamente a qualidade das inferências, sobretudo em amostras pequenas. À medida que o tamanho amostral aumenta, a influência da priori diminui, mas as prioris elicitadas permanecem competitivas, sem comprometer a consistência dos estimadores.

Esses achados reforçam a relevância do uso de informação prévia estruturada em problemas de valores extremos, caracterizados pela escassez de dados.

A seguir, apresentam-se as figuras 7.3.1, 7.3.2, 7.3.3 e 7.3.4 correspondentes aos resultados do estudo de simulação.

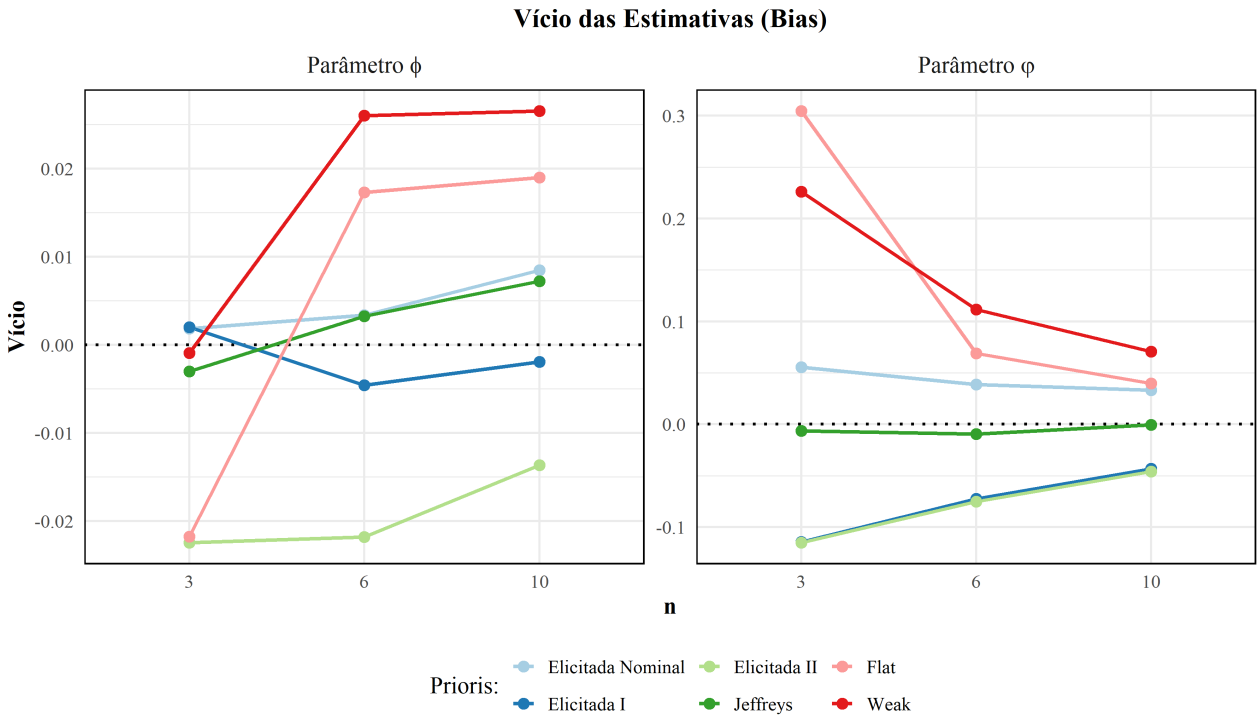


Figura 7.3.1: Viés das estimativas dos parâmetros sob os diferentes cenários considerados.

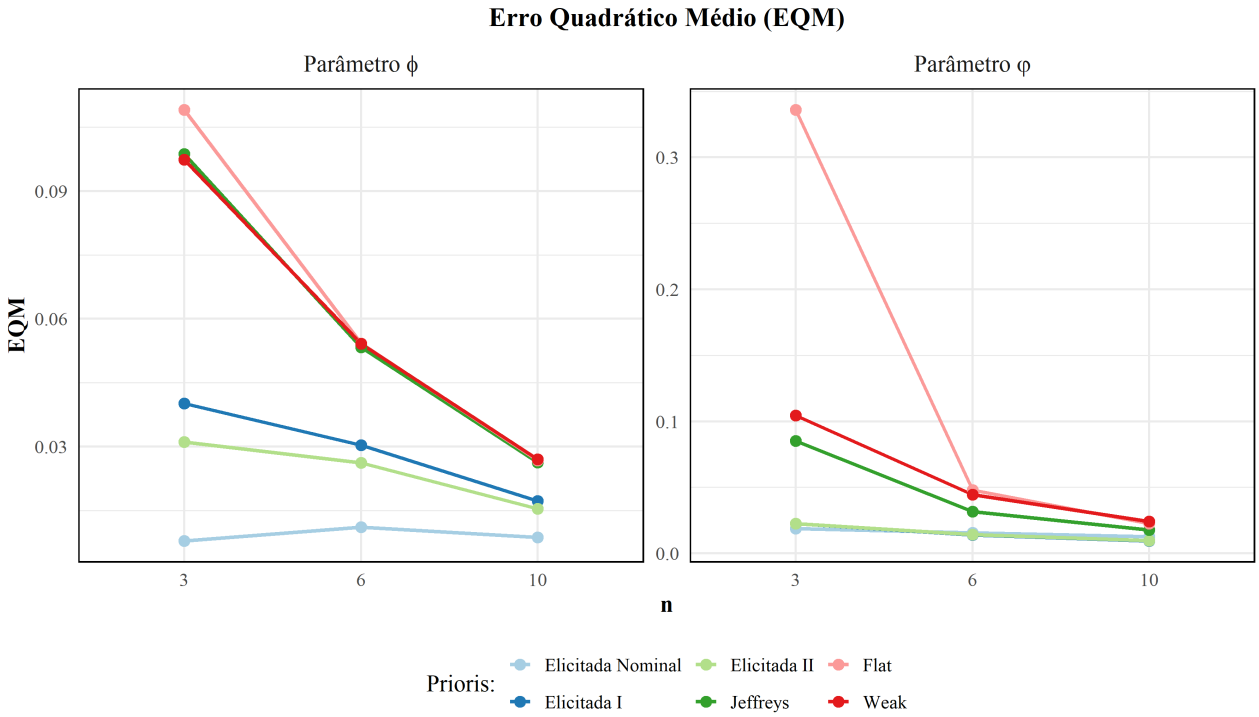


Figura 7.3.2: Erro quadrático médio (EQM) das estimativas dos parâmetros sob os diferentes cenários.

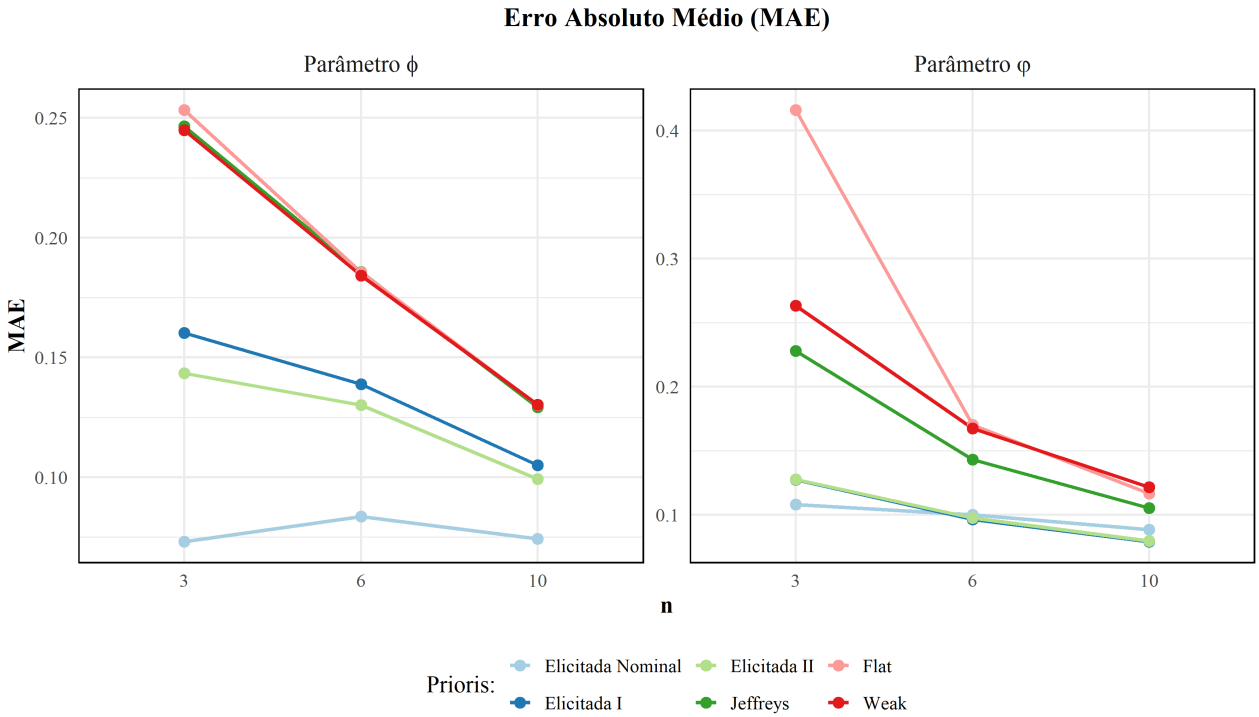


Figura 7.3.3: Erro absoluto médio (MAE) das estimativas dos parâmetros sob os diferentes cenários.

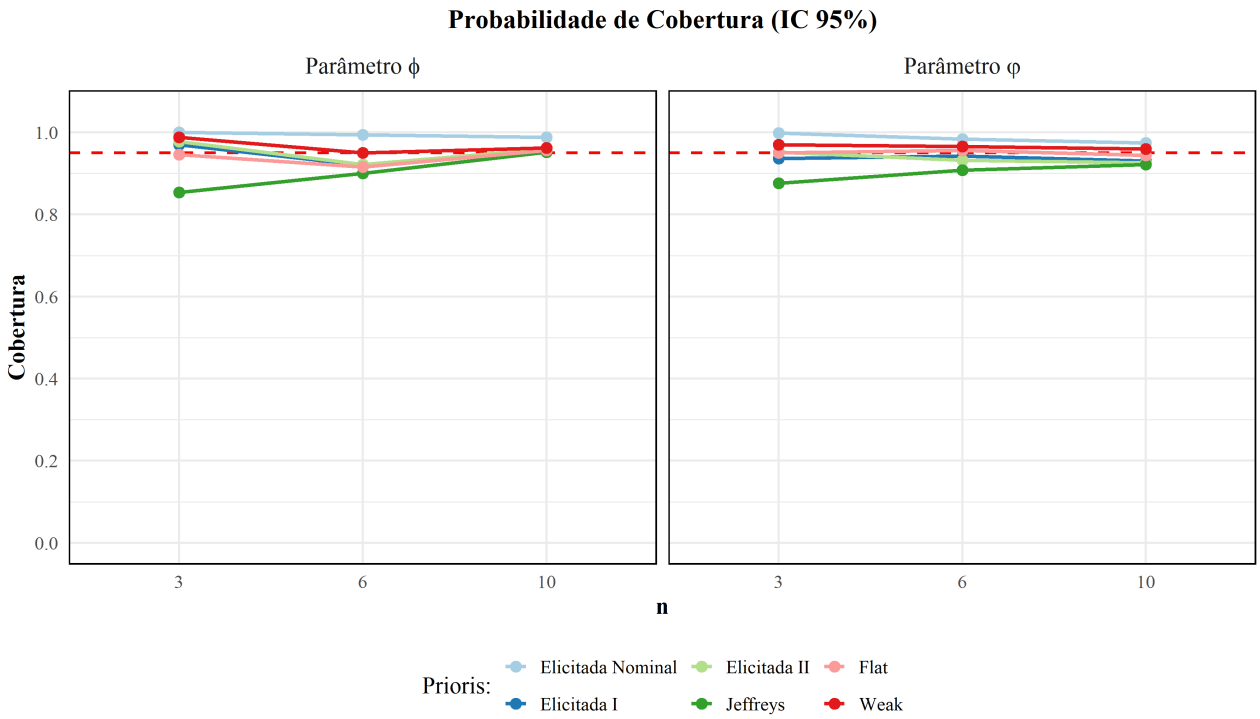


Figura 7.3.4: Probabilidade de cobertura dos intervalos de credibilidade para os parâmetros, sob os diferentes cenários.

7.3.1 Gráfico das Distribuições a Posteriori

As Figuras correspondentes às distribuições a posteriori dos parâmetros ϕ (localização) e φ (escala), para diferentes tamanhos amostrais, permitem uma avaliação visual do impacto das diferentes escolhas de priori sobre a inferência bayesiana.

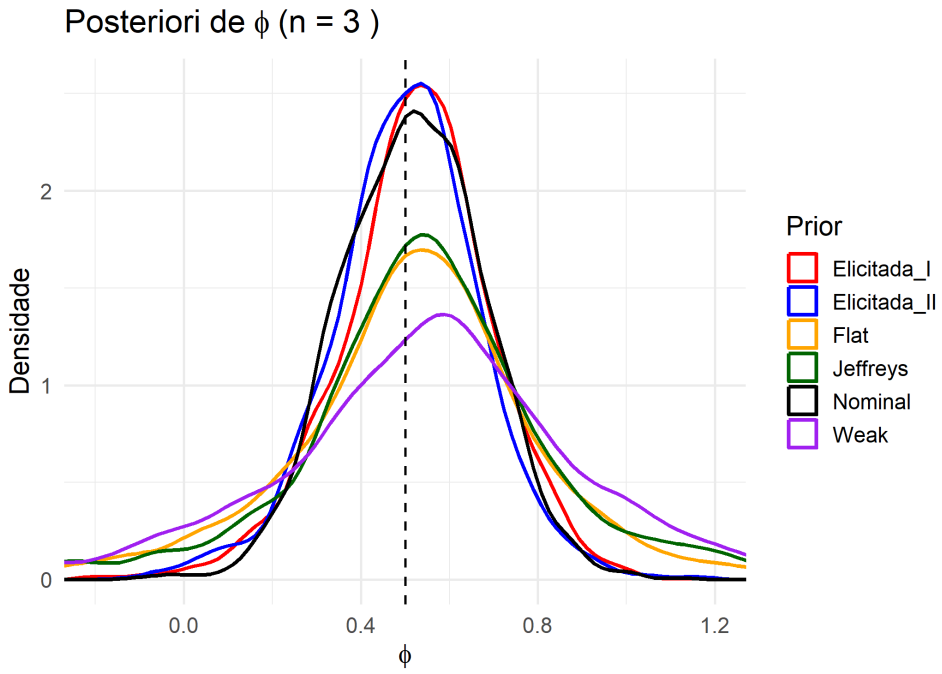


Figura 7.3.5: Distribuições a posteriori do parâmetro de localização ϕ para amostra pequena ($n = 3$).

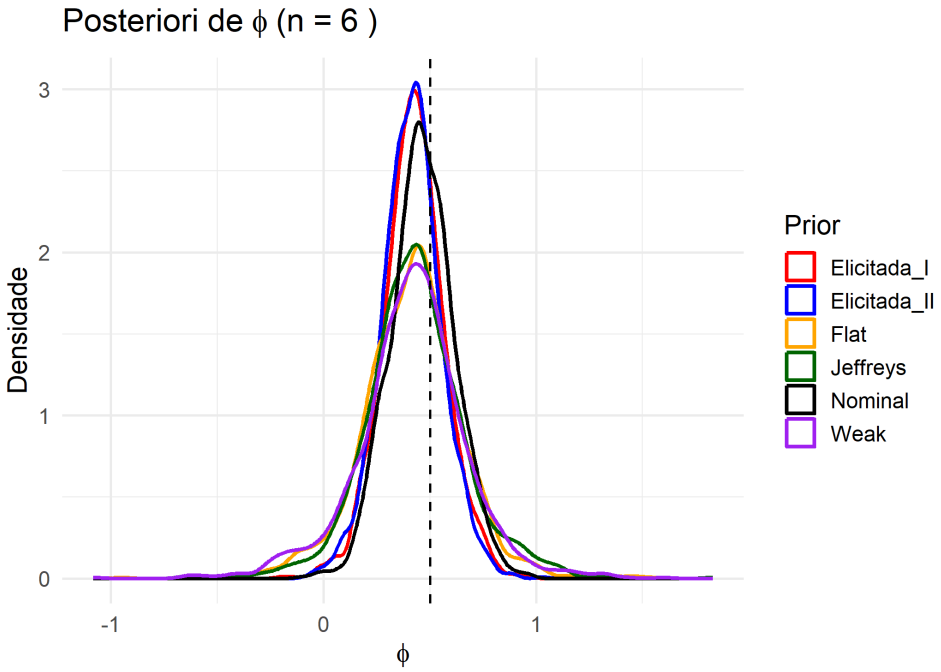


Figura 7.3.6: Comportamento das distribuições a posteriori de ϕ para $n = 6$.

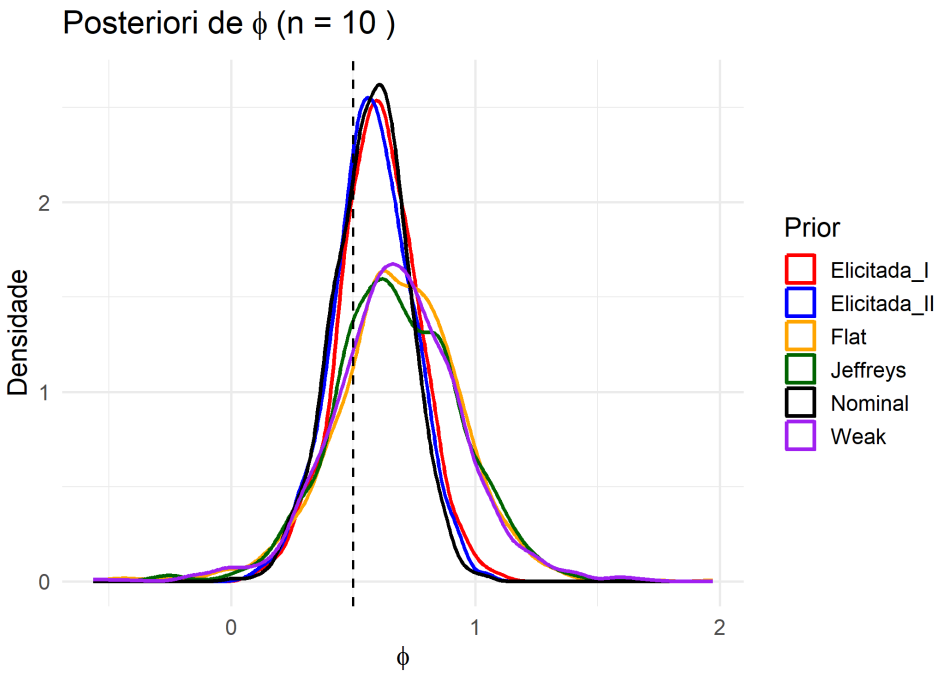


Figura 7.3.7: Distribuições a posteriori de ϕ para $n = 10$.

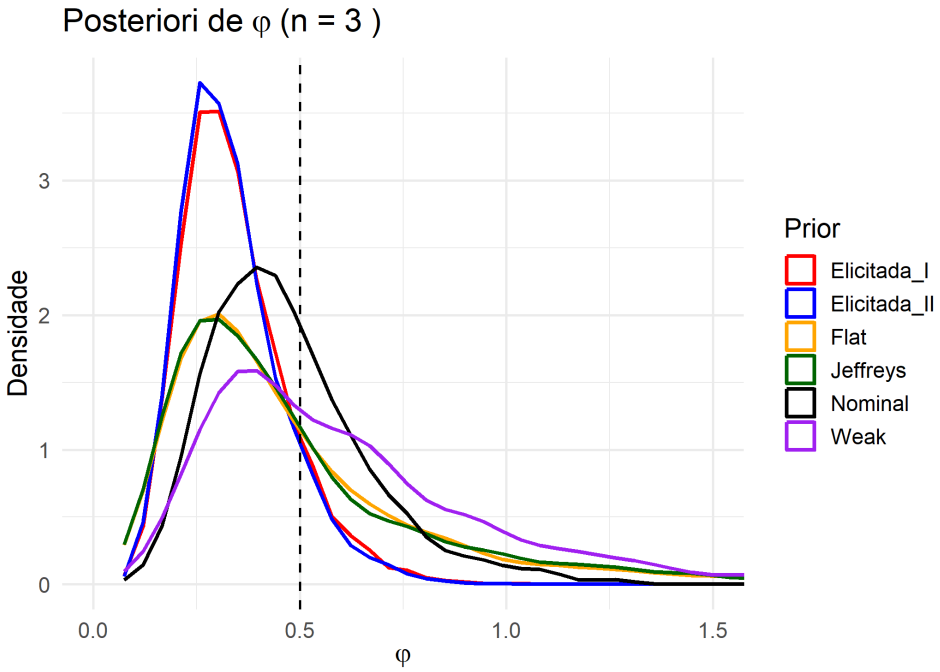


Figura 7.3.8: Distribuições a posteriori do parâmetro de escala φ para $n = 3$.

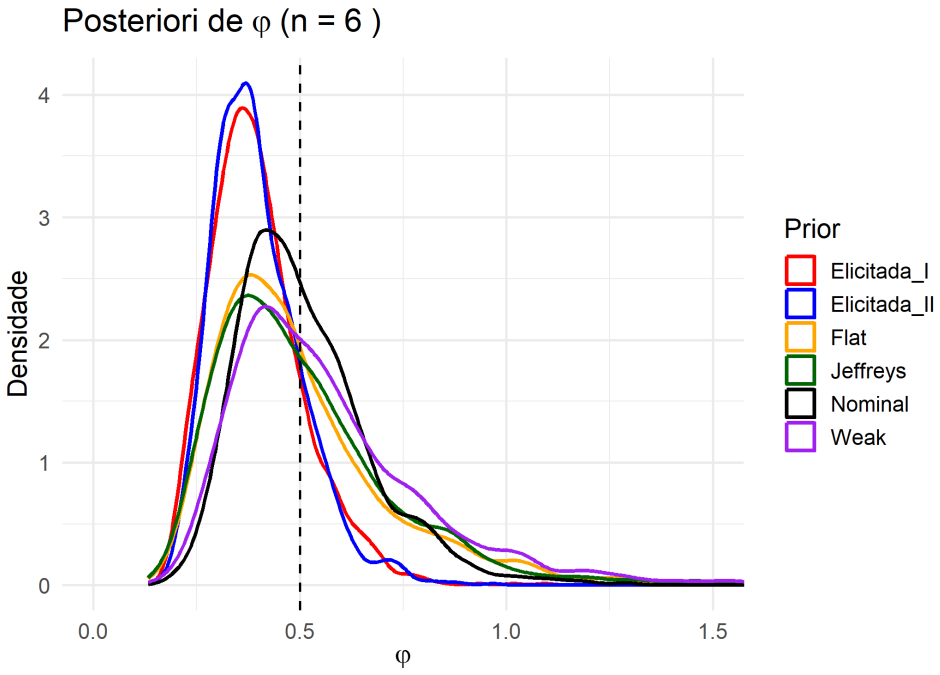


Figura 7.3.9: Distribuições a posteriori de φ para $n = 6$.

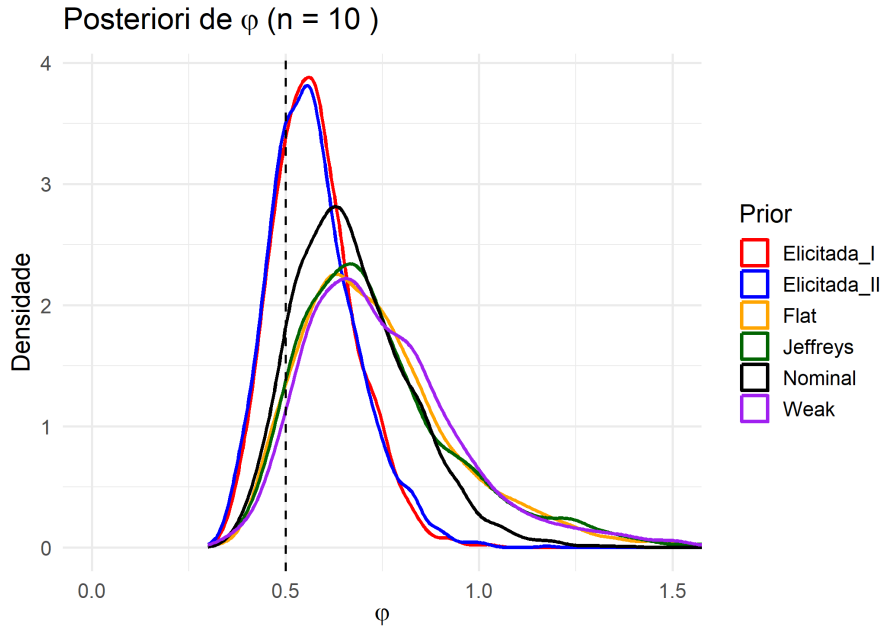


Figura 7.3.10: Distribuições a posteriori de φ para $n = 10$.

Parâmetro de localização ϕ

Para $n = 3$ na figura 7.3.5, observa-se que todas as distribuições a posteriori estão relativamente concentradas em torno do valor verdadeiro (indicado pela linha pontilhada). No entanto, há diferenças claras na dispersão: as prioris elicitadas (I e II) produzem distribuições mais concentradas e com maior pico, indicando menor variabilidade. A priori nominal apresenta comportamento semelhante, embora ligeiramente deslocada em relação ao centro das elicitadas. Em contraste, as prioris não informativas (Jeffreys e Flat) e a fracamente informativa exibem maior dispersão, refletindo maior incerteza na estimação quando a informação prévia é limitada.

Para $n = 6$ na figura 7.3.6, nota-se uma maior concentração das distribuições a posteriori, com redução geral da variabilidade. As prioris elicitadas continuam apresentando distribuições mais estreitas e centradas próximas ao valor verdadeiro, enquanto as prioris não informativas ainda demonstram maior dispersão, embora menos acentuada que no caso anterior. A priori nominal apresenta boa centralização, mas ainda com leve assimetria.

Para $n = 10$ na figura 7.3.7, as diferenças entre as prioris tornam-se menos pronunciadas. Todas as distribuições a posteriori convergem para uma forma semelhante, evidenciando a dominância da informação amostral. Ainda assim, as prioris elicitadas mantêm leve vantagem em termos de concentração, enquanto as prioris não informativas apresentam caudas um pouco mais longas.

Parâmetro de escala φ

Para $n = 3$ na figura 7.3.8, observa-se maior assimetria nas distribuições a posteriori, característica típica do parâmetro de escala. As prioris elicitadas (I e II) produzem distribuições mais concentradas, porém ligeiramente deslocadas à esquerda do valor verdadeiro, indicando um possível viés negativo. A priori nominal apresenta melhor alinhamento com o valor real, embora com menor pico. As prioris não informativas e a fracamente informativa exibem maior dispersão e caudas mais longas, refletindo elevada incerteza.

Para $n = 6$ na figura 7.3.9, há uma redução significativa da variabilidade. As distribuições tornam-se mais concentradas, e o deslocamento observado nas prioris elicitadas diminui. A priori nominal apresenta bom equilíbrio entre centralização e dispersão. As prioris não informativas ainda apresentam caudas mais pronunciadas, mas já se aproximam das demais em termos de forma.

Para $n = 10$ na figura 7.3.10, verifica-se uma clara convergência entre todas as distribuições a posteriori. As diferenças entre prioris tornam-se sutis, com todas as curvas concentradas próximas ao valor verdadeiro. Ainda assim, observa-se que as prioris não informativas mantêm leve assimetria e caudas mais longas, enquanto as prioris elicitadas apresentam distribuições mais simétricas e concentradas.

De forma geral, os resultados gráficos corroboram os achados quantitativos da simulação. Em amostras pequenas, a escolha da priori exerce influência significativa sobre a forma da distribuição a posteriori, sendo as prioris elicitadas capazes de reduzir a variabilidade e produzir estimativas mais estáveis. À medida que o tamanho amostral aumenta, a influência da priori diminui e as distribuições tendem a convergir, evidenciando a consistência do procedimento bayesiano.

Esses resultados reforçam que a elicitação de prioris, quando bem especificada, constitui uma ferramenta eficaz para melhorar a inferência em contextos de escassez de dados, sem comprometer a robustez do modelo à medida que mais informação é incorporada.

Capítulo

8

Ilustração Numérica

Neste capítulo, apresenta-se uma ilustração numérica com o objetivo de demonstrar a aplicação prática da metodologia de elicitação de prioris baseada na aproximação de Laplace (1986) no contexto da distribuição de Gumbel.

A análise é conduzida em um cenário de extrema escassez de dados, com tamanho amostral reduzido, a fim de evidenciar o impacto da incorporação de informação prévia na inferência Bayesiana. Em particular, busca-se avaliar a capacidade das prioris elicitadas em produzir inferências estáveis e coerentes quando comparadas com prioris não informativas.

8.1 Exemplo ilustrativo de uma amostra de tamanho 3

Com o intuito de ilustrar o comportamento do método proposto em condições críticas, considera-se uma amostra de tamanho $n = 3$. O objetivo principal deste experimento é analisar o desempenho inferencial das prioris elicitadas via aproximação de Laplace (1986) em comparação com prioris de referência, tais como a priori de Jeffreys, a priori Flat e a priori fracamente informativa (Weak).

A escolha de uma amostra extremamente pequena permite destacar a relevância da informação prévia na estabilização das estimativas, especialmente para o parâmetro de escala.

8.1.1 Configuração do experimento

A amostra observada utilizada no estudo foi:

$$\boldsymbol{x} = \{0.3898,\ 1.2183,\ 0.5558\}. \tag{8.1.1}$$

Os hiperparâmetros das prioris informativas foram obtidos previamente por meio da aproximação de Laplace (1986), a partir de quantis fornecidos por um especialista.

A inferência a posteriori foi realizada utilizando o algoritmo de Metropolis-Hastings implementado no software R. Foram adotadas as seguintes configurações:

- Número total de iterações: 20.000;
- Período de burn-in: 5.000 iterações;
- Thinning: 10;

- Valores iniciais: média e desvio padrão amostral;
- Propostas: passeio aleatório para ϕ e transformação logarítmica para φ.

8.1.2 Algoritmo de inferência

Devido à ausência de forma analítica fechada para a distribuição a posteriori conjunta, utilizou-se o algoritmo de Metropolis-Hastings para gerar amostras da posteriori.

Algorithm 2 Algoritmo MCMC (Metropolis-Hastings)

Dados: $\boldsymbol{x} = (x_1, x_2, x_3)$
Inicializar $\phi^{(0)}$ e $\varphi^{(0)}$
for $i = 1, \ldots, N$ **do**
    Gerar $\phi^* \sim N(\phi^{(i-1)}, \sigma^2_\phi)$
    Calcular a razão de aceitação:

$$\alpha = \min\left(1, \frac{p(\boldsymbol{x}|\phi^*, \varphi^{(i-1)})\pi(\phi^*, \varphi^{(i-1)})}{p(\boldsymbol{x}|\phi^{(i-1)}, \varphi^{(i-1)})\pi(\phi^{(i-1)}, \varphi^{(i-1)})}\right)$$

    Aceitar ou rejeitar ϕ^*
    Gerar $w^* \sim N(\log \varphi^{(i-1)}, \sigma^2)$
    Definir $\varphi^* = \exp(w^*)$
    Calcular nova razão de aceitação (com Jacobiano)
    Aceitar ou rejeitar φ^*
end for
Descartar burn-in e aplicar thinning

8.2 Análise dos resultados

A Tabela apresenta as estimativas a posteriori para os parâmetros ϕ e φ, incluindo média, desvio-padrão e intervalos de credibilidade de 95%.

Tabela 8.2.1: Estimativas a posteriori ($n = 3$)

Priori	ϕ Média (DP)	IC 95%	φ Média (DP)	IC 95%
Elicitada I	0,53 (0,18)	[0,16; 0,87]	0,33 (0,12)	[0,16; 0,62]
Elicitada II	0,51 (0,17)	[0,14; 0,84]	0,34 (0,12)	[0,16; 0,62]
Nominal	0,53 (0,16)	[0,19; 0,83]	0,49 (0,21)	[0,20; 0,99]
Jeffreys	0,51 (0,27)	[0,01; 1,01]	0,34 (0,24)	[0,12; 0,96]
Flat	0,51 (0,48)	[-0,51; 1,29]	0,58 (0,65)	[0,14; 2,08]
Weak	0,47 (0,58)	[-0,76; 1,33]	0,65 (0,79)	[0,16; 2,52]

Os resultados indicam que, para o parâmetro de localização ϕ, todas as prioris produzem estimativas semelhantes, evidenciando baixa sensibilidade da inferência à escolha da priori. Tomando a priori nominal como referência para comparação, observa-se que as prioris elicitadas apresentam comportamento bastante semelhante, sendo a Elicitada I ligeiramente mais próxima da especificação nominal.

Para o parâmetro de escala φ, por outro lado, observa-se maior influência da escolha da priori. As prioris elicitadas produzem distribuições a posteriori mais concentradas, refletindo menor variabilidade, enquanto as prioris não informativas apresentam maior dispersão e intervalos de credibilidade mais amplos.

As densidades a posteriori obtidas evidenciam diferenças relevantes entre as especificações de priori, especialmente em cenários de pequena amostra, nos quais a informação a priori exerce maior influência sobre a inferência.

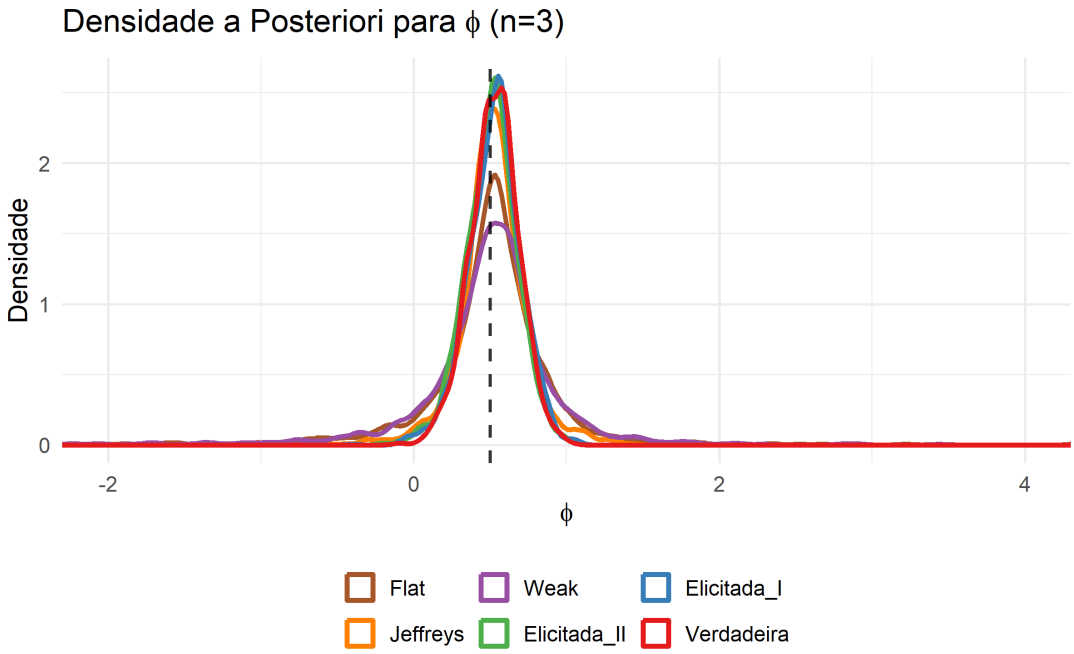


Figura 8.2.1: Distribuições a posteriori do parâmetro de localização ϕ sob diferentes prioris ($n = 3$).

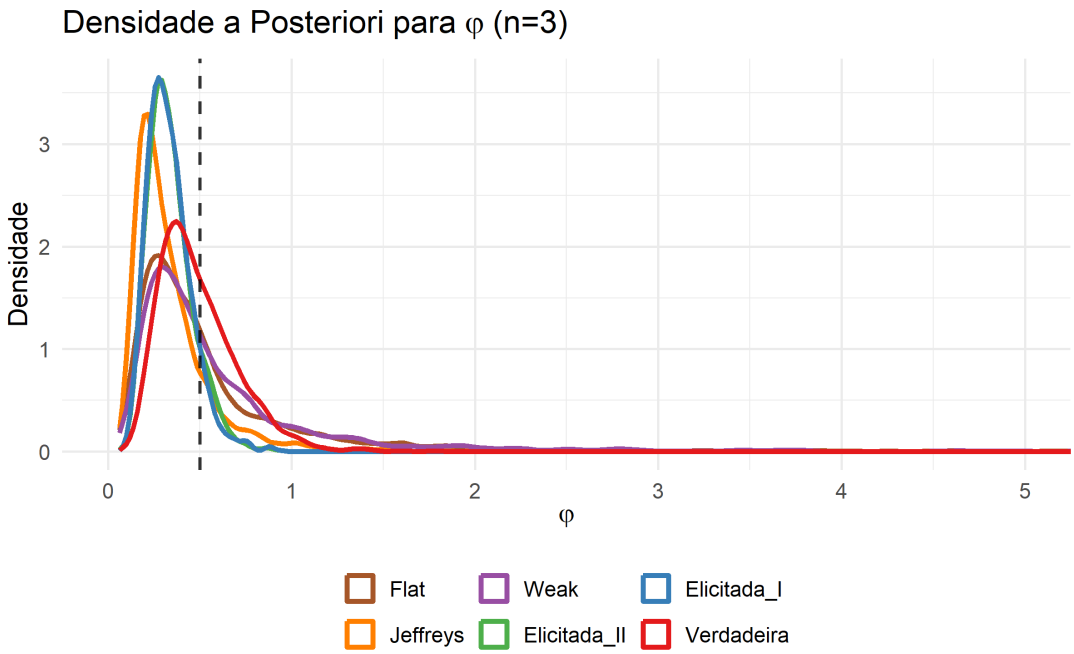


Figura 8.2.2: Distribuições a posteriori do parâmetro de escala φ sob diferentes prioris ($n = 3$).

As Figuras 8.2.1 e 8.2.2 mostram que o impacto da escolha da priori depende do parâmetro considerado. Para o parâmetro de localização ϕ, as distribuições apresentam formas bastante semelhantes, com diferenças pouco pronunciadas entre as prioris, indicando baixa sensibilidade da inferência.

Por outro lado, para o parâmetro de escala φ, observam-se diferenças mais evidentes. As prioris elicitadas produzem distribuições mais concentradas, refletindo menor variabilidade, enquanto as prioris não informativas apresentam maior dispersão e caudas mais longas. Nota-se ainda um leve deslocamento das distribuições elicitadas em relação à priori nominal, evidenciando o efeito da informação prévia incorporada.

A qualidade das cadeias geradas foi avaliada por meio da inspeção dos traceplots e das funções de autocorrelação (ACF), após a remoção do período de *burn-in* e a aplicação do *thinning*. As Figuras 8.2.3 e 8.2.4 indicam comportamento adequado das cadeias de Markov. Os traceplots apresentam oscilações em torno de um nível aproximadamente constante, sem tendências sistemáticas ou mudanças estruturais ao longo das iterações, sugerindo estacionariedade. Complementarmente, as funções de autocorrelação evidenciam rápido decaimento das correlações, especialmente para ϕ, enquanto para φ observa-se uma dependência ligeiramente maior nos primeiros defasamentos, que se reduz rapidamente, indicando boa mistura das cadeias e fornecendo evidências de convergência sob os diagnósticos gráficos considerados.

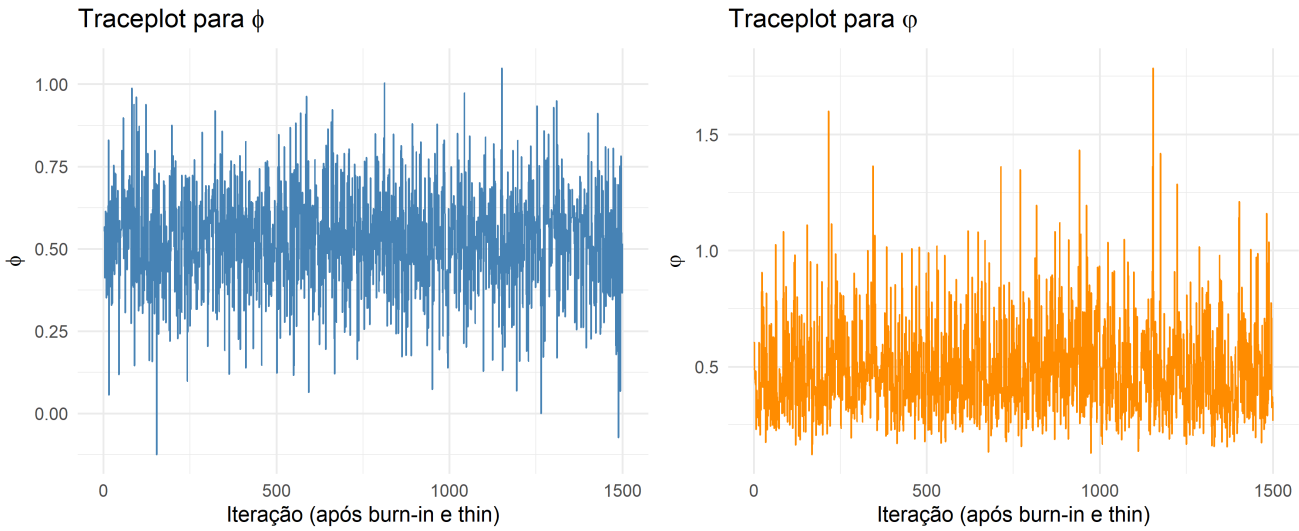


Figura 8.2.3: Traceplots das cadeias MCMC para os parâmetros ϕ e φ.

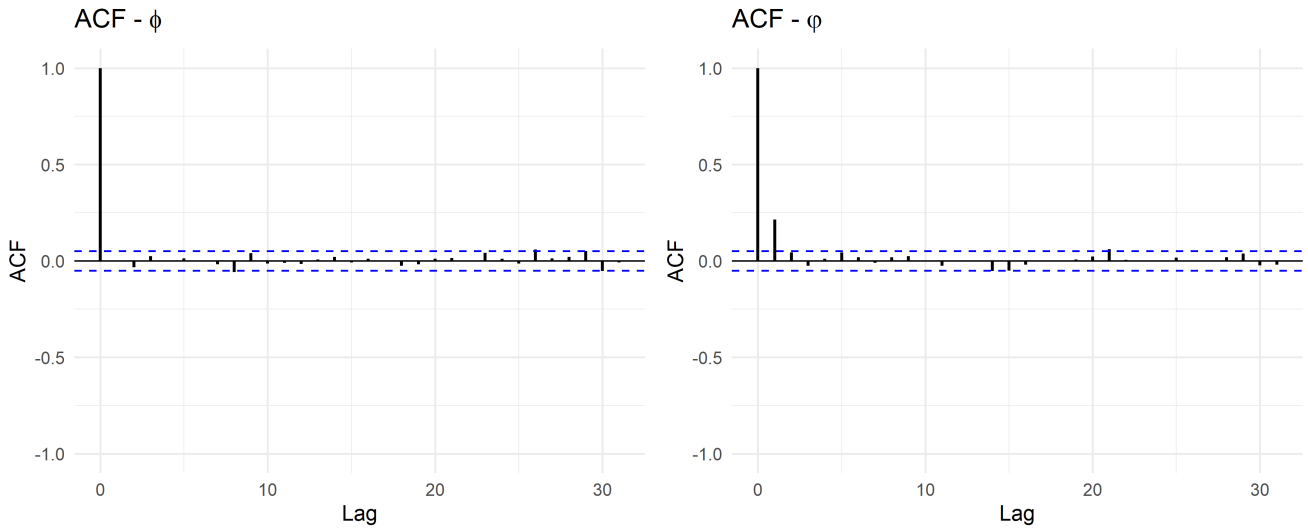


Figura 8.2.4: Função de autocorrelação das cadeias MCMC para ϕ e φ.

Em conjunto, esses resultados indicam que o algoritmo MCMC apresentou bom desempenho, produzindo amostras confiáveis para a inferência a posteriori.

Os resultados evidenciam que a elicitação de prioris via aproximação de Laplace (1986) proporciona ganhos substanciais em termos de estabilidade inferencial.

Para o parâmetro ϕ, a informação contida nos dados é suficiente para garantir estimativas estáveis, independentemente da priori utilizada.

Entretanto, para o parâmetro φ, a utilização de prioris informativas reduz significativamente a incerteza, embora possa introduzir leve viés. Esse comportamento reflete o compromisso entre viés e variância característico da inferência Bayesiana.

Assim, as prioris elicitadas mostram-se particularmente adequadas para aplicações práticas, nas quais a informação prévia desempenha papel fundamental.

Capítulo 9

Aplicação a Dados Reais: Pluviometria em Penta-di-Casinca

9.1 Descrição da Amostra e Contextualização

Nesta etapa, validamos o processo de elicitação de priori via preditiva a priori utilizando dados reais de precipitação. Os dados referem-se aos registros de pluviometria máxima anual na estação meteorológica de Penta-di-Casinca, na Córsega, França.

Esta base de dados foi extraída do trabalho de Ruggeri e Bousquet (2017). A amostra compreende $n = 29$ observações coletadas entre o período de 1987 a 2015, configurando-se como um estudo de caso central para a metodologia discutida por Bousquet e Keller (2018).

> **Amostra (X_{raw}):** 107.8, 72.4, 204.5, 83.8, 142.0, 95.5, 316.1, 177.9, 87.3, 81.9, 109.5, 89.5, 150.7, 122.1, 98.2, 113.2, 104.4, 66.9, 136.4, 275.4, 125.0, 199.8, 51.2, 75.0, 166.2, 106.0, 72.8, 190.4, 105.0.

Os dados e os índices foram escalonados por um fator de 10 ($X = X_{raw}/10$) para estabilidade numérica durante os processos de otimização.

9.2 Informação do Especialista (Pontos de Elicitação)

A elicitação foi baseada em três quantis de probabilidade acumulada (S_i) reportados por Ruggeri e Bousquet (2017), os quais representam a informação fornecida por um especialista hidrológico no estudo original. Neste trabalho, esses pontos de elicitação foram adotados para ilustrar a aplicação da metodologia de obtenção de distribuições a priori por meio da aproximação de Laplace descrita nos capítulos anteriores.

Tabela 9.2.1: Informação a priori fornecida pelo especialista.

Índice (i)	Pluviometria (t_i)	Probabilidade Estimada (S_i)
1	75 mm	25%
2	100 mm	50%
3	150 mm	75%

9.3 Elicitação via Aproximação de Laplace

A estimação dos hiperparâmetros foi realizada conforme a metodologia descrita nos capítulos anteriores, utilizando a aproximação de Laplace proposta por Tierney e Kadane (1986). Os valores obtidos para cada cenário de elicitação são apresentados na Tabela 9.3.1.

Tabela 9.3.1: Hiperparâmetros estimados para cada cenário de elicitação.

Cenário	a	b	c	d
Elicitação do especialista	8,5343	0,6922	2,9492	0,6454
Elicitada I	7,7872	2,0469	17,6050	3,9594
Elicitada II	7,2860	0,0210	3,5646	0,8807

A Tabela 9.3.1 mostra que pequenas variações nos quantis de elicitação produzem alterações nos hiperparâmetros estimados. Esse comportamento é esperado, uma vez que os hiperparâmetros são obtidos de forma a reproduzir as informações especificadas na etapa de elicitação. Assim, a sensibilidade observada reflete a dependência natural do procedimento em relação às informações fornecidas pelo especialista, não caracterizando, por si só, uma limitação da metodologia proposta.

9.4 Validação da Priori

A adequação da priori construída foi avaliada comparando-se as probabilidades preditivas com os valores fornecidos pelo especialista. Os resultados são apresentados na Tabela 9.4.1.

Tabela 9.4.1: Comparação entre probabilidades do especialista, priori e posteriori.

Índice	Especialista	A priori	A posteriori
$i = 1$	25%	25%	20,6%
$i = 2$	50%	50%	43,4%
$i = 3$	75%	75%	79,3%

Observa-se, na Tabela 9.4.1, que a distribuição a priori reproduz exatamente os quantis de elicitação, conforme esperado pelo procedimento adotado. Após a incorporação da informação proveniente dos dados observados, as probabilidades a posteriori apresentam pequenas alterações em relação às probabilidades inicialmente especificadas pelo especialista. Em particular, para os níveis de 75 mm e 100 mm verifica-se uma redução das probabilidades, enquanto para o nível de 150 mm observa-se um ligeiro aumento. Essas diferenças refletem o processo de atualização Bayesiana, no qual a informação contida na amostra complementa o conhecimento prévio, produzindo uma distribuição a posteriori que concilia a informação fornecida pelo especialista com a evidência observada.

A Figura 9.4.1 ilustra o processo de atualização bayesiana.

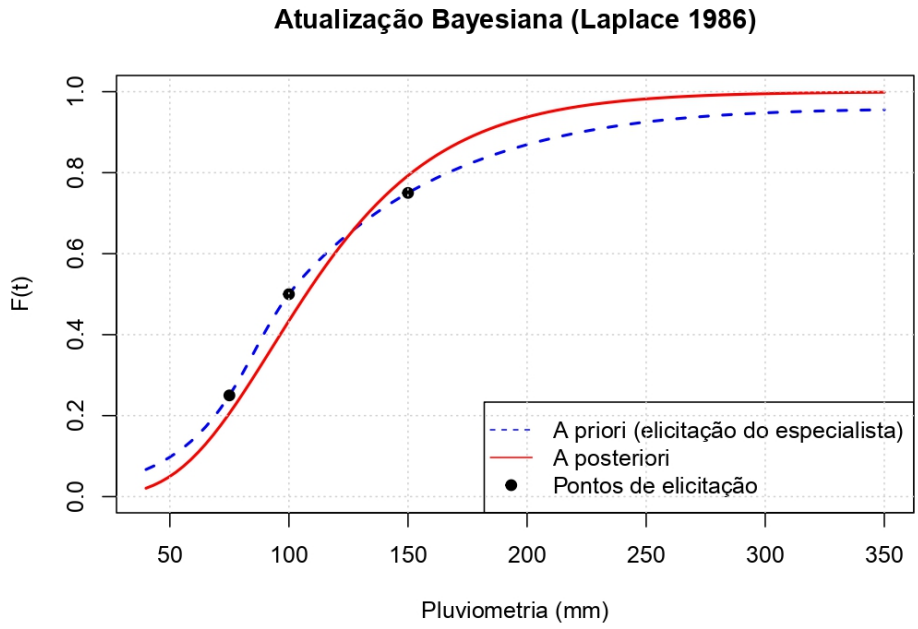


Figura 9.4.1: Atualização bayesiana: comparação entre a distribuição a priori (elicitação do especialista), a distribuição a posteriori e os pontos de elicitação.

Na Figura 9.4.1, observa-se que a distribuição a priori passa exatamente pelos pontos de elicitação, enquanto a distribuição a posteriori se ajusta em função da informação contida nos dados observados. Dessa forma, a distribuição a posteriori representa a atualização da informação inicial do especialista após a incorporação da evidência amostral.

A Figura 9.4.2 apresenta as distribuições a priori obtidas sob diferentes cenários de elicitação.

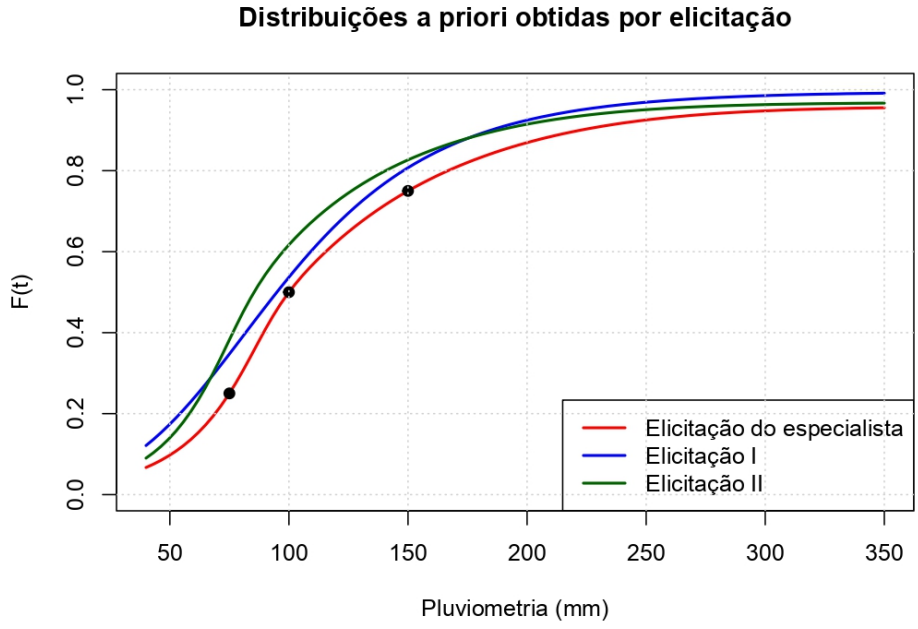


Figura 9.4.2: Distribuições a priori obtidas a partir da elicitação do especialista e dos cenários perturbados (Elicitada I e Elicitada II).

Observa-se que a priori baseada na elicitação do especialista reproduz fielmente os pontos de elicitação, enquanto as versões perturbadas apresentam deslocamentos decorrentes das alterações introduzidas nos quantis especificados. Esse comportamento evidencia o impacto de pequenas perturbações nas informações elicitadas sobre a forma das distribuições a priori.

9.5 Comparação entre Diferentes Prioris

A influência da escolha da priori foi analisada considerando diferentes especificações, cujos resultados são apresentados na Tabela 9.5.1.

Tabela 9.5.1: Comparação das probabilidades fornecidas pelo especialista e das probabilidades obtidas sob diferentes prioris.

t	Especialista	Elicitação do especialista	Elicitada I	Elicitada II	Jeffreys	Flat	Weak
75	25%	20,6%	15,8%	38,5%	13,3%	13,5%	13,8%
100	50%	43,4%	37,0%	57,8%	34,8%	34,5%	35,1%
150	75%	79,3%	74,9%	83,4%	74,9%	74,1%	74,5%

A partir da Tabela 9.5.1, verifica-se que a priori obtida a partir da elicitação do especialista conduz a resultados mais equilibrados e coerentes com os dados observados. Em contraste, a Elicitada I tende a subestimar as probabilidades para os níveis de precipitação considerados, enquanto a Elicitada II apresenta superestimação para alguns desses níveis, evidenciando o impacto de pequenas perturbações nos quantis elicitados sobre as probabilidades preditivas obtidas.

Observa-se ainda que as prioris não informativas (Jeffreys, Flat e Weak) apresentam comportamento bastante semelhante entre si, indicando que, na ausência de informação prévia específica, a inferência é predominantemente influenciada pela função de verossimilhança.

A Figura 9.5.1 apresenta a comparação das distribuições a posteriori. Observa-se que a posteriori baseada na elicitação do especialista apresenta melhor alinhamento com os dados, enquanto as prioris não informativas produzem resultados semelhantes entre si.

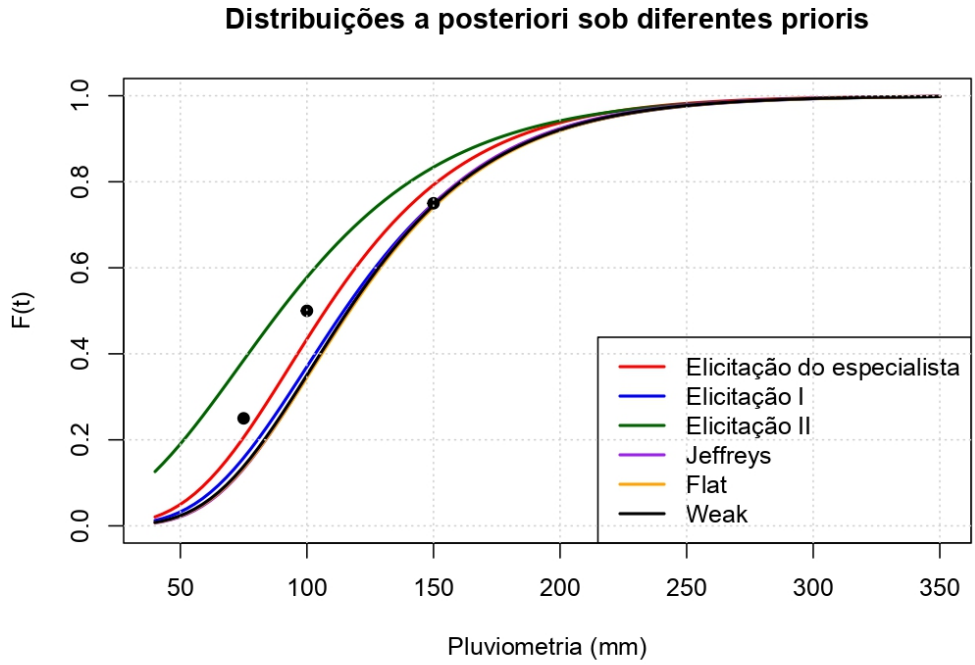


Figura 9.5.1: Distribuições a posteriori sob diferentes prioris: elicitação do especialista, cenários perturbados e prioris não informativas.

Os resultados obtidos nesta aplicação evidenciam que a escolha da priori exerce influência relevante na inferência bayesiana. A incorporação adequada da informação do especialista contribui para estimativas mais realistas, enquanto erros na elicitação podem comprometer a qualidade dos resultados.

Além disso, observa-se que prioris não informativas tendem a ignorar informações relevantes disponíveis, podendo levar a inferências menos consistentes em contextos onde há conhecimento prévio.

Dessa forma, conclui-se que a elicitação bem conduzida representa um elemento fundamental na modelagem bayesiana, especialmente em aplicações envolvendo eventos extremos.

Capítulo

10

Conclusões

Este trabalho apresentou o desenvolvimento e a aplicação de uma metodologia de elicitação de priori para a distribuição de Gumbel, baseada na distribuição preditiva a priori como mecanismo para incorporar o conhecimento de especialistas. A resolução das integrais envolvidas nesse processo foi realizada por meio da aproximação de Laplace proposta por Tierney e Kadane (1986), a qual se mostrou adequada tanto do ponto de vista analítico quanto computacional. As principais conclusões são apresentadas a seguir.

10.1 Eficácia da Metodologia de Elicitação

A metodologia baseada na minimização da discrepância entre as probabilidades fornecidas pelo especialista e aquelas geradas pelo modelo mostrou-se eficaz na construção de prioris informativas. Tanto nos estudos de simulação quanto na aplicação a dados reais, o método permitiu traduzir quantis subjetivos em hiperparâmetros consistentes para os parâmetros de localização (ϕ) e escala (φ).

Esse resultado é particularmente relevante no contexto da inferência bayesiana, onde a especificação da priori constitui um dos principais desafios. A abordagem adotada garante uma incorporação estruturada da informação prévia, mantendo coerência probabilística e interpretabilidade.

10.2 Desempenho em Estudos de Simulação

Os resultados da simulação evidenciaram que as prioris elicitadas apresentam desempenho superior em cenários de pequena amostra, com menores níveis de viés e erro quadrático médio, além de melhores probabilidades de cobertura, quando comparadas às prioris não informativas.

A priori elicitada nominal destacou-se como a melhor alternativa quando a informação do especialista é consistente, enquanto as versões perturbadas (Elicitada I e II) mostraram que o método mantém robustez mesmo na presença de imprecisões moderadas. Por outro lado, prioris não informativas, como Jeffreys e Flat, apresentaram maior variabilidade, especialmente para o parâmetro de escala.

À medida que o tamanho amostral aumenta, observou-se a redução da influência da priori, com convergência dos resultados entre diferentes especificações, evidenciando a predominância da informação amostral.

10.3 Aplicação a Dados Reais

A aplicação aos dados de pluviometria da estação de **Penta-di-Casinca** demonstrou a utilidade prática da metodologia proposta. Verificou-se que a distribuição a priori construída a partir da elicitação do especialista foi consistente com as probabilidades fornecidas, enquanto a atualização bayesiana permitiu ajustar essas crenças à evidência observada.

O deslocamento entre as curvas a priori e a posteriori indicou que os dados sugerem uma maior frequência de eventos extremos em relação à percepção inicial do especialista, evidenciando o papel da inferência bayesiana como mecanismo de atualização do conhecimento.

10.4 Considerações Metodológicas

Embora a formulação alternativa baseada na aproximação de Laplace proposta por Tierney e Kadane (1989) tenha sido analisada, verificou-se que, no contexto da distribuição de Gumbel, essa abordagem conduz a um problema de não identificabilidade em relação ao hiperparâmetro b. Em particular, a expressão da preditiva a priori torna-se independente desse parâmetro, implicando que diferentes valores de b resultam na mesma priori.

Esse comportamento é inadequado do ponto de vista inferencial, uma vez que impede a correta representação da incerteza associada ao parâmetro de localização. Como consequência, optou-se por não utilizar essa formulação como método principal, sendo sua apresentação mantida apenas no apêndice para fins de completude teórica.

10.5 Propostas Futuras

Os resultados obtidos indicam que a elicitação via preditiva a priori, combinada com a aproximação de Laplace (1986), constitui uma ferramenta robusta para a modelagem bayesiana de valores extremos.

Como continuidade deste trabalho, sugerem-se as seguintes direções:

1. Extensão da metodologia para outras distribuições da família de valores extremos, como a Distribuição Generalizada de Valores Extremos (GEV);

2. Desenvolvimento de ferramentas computacionais interativas que auxiliem o processo de elicitação junto ao especialista;

3. Investigação de métodos alternativos de aproximação de integrais, com o objetivo de comparar desempenho numérico e estabilidade em diferentes cenários.

Referências

ACHCAR, J. A.; SMITH, A. F. M. *Aspects of Reparametrization in Approximate Bayesian Inference*. In: HODGES, J. L. (Ed.). Bayesian and Likelihood Methods in Statistics and Econometrics. Amsterdam: North-Holland, 1990. p. 439–452.

AZEVEDO-FILHO, A.; SCHACHTER, R. D. Laplace's method approximations for probabilistic inference in belief networks with continuous variables. In: *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*. p. 28–36, 1994.

BEDFORD, T. J.; COOKE, R. M. *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge: Cambridge University Press, 2001.

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

BERNARDO, J. M.; SMITH, A. F. M. *Bayesian Theory*. Hoboken: John Wiley & Sons, 2009.

BOUSQUET, N.; KELLER, M. *Bayesian Prior Elicitation and Selection for Extreme Values*. arXiv preprint, 2018.

BREZIS, H. *Functional Analysis Sobolev Spaces and Partial Differential Equations*. Springer, New York, 2010.

DE BRUIJN, N. G. *The Laplace Method for Integrals*. v. 1. North-Holland Publishing Co., 1958.

CASELLA, G.; BERGER, R. L. *Statistical inference*. 2. ed. Pacific Grove: Duxbury, 2002.

CAVALCANTI, M. M.; CAVALCANTI, V. N. D. *Introdução à Teoria das Distribuições e aos Espaços de Sobolev*. Ed. Universidade Estadual de Maringá, 2009.

CAZENAVE, T. *Semilinear Schrödinger Equations*. Courant Lecture Notes in Mathematics, vol. 10, New York University, Courant Institute of Mathematical Sciences/American Mathematical Society, New York/Providence, RI, 2003.

CHALONER, K. M.; DUNCAN, G. T. Assessment of a beta prior distribution: modal elicitation and posterior inference. *The Statistician*, v. 32, n. 3, p. 174–180, 1983.

COLES, S. *An introduction to statistical modeling of extreme values*. London: Springer, 2001.

COOKE, R. M. *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Oxford: Oxford University Press, 1991.

COX, D. R.; HINKLEY, D. V. *Theoretical statistics*. London: Chapman and Hall, 1974.

DAVISON, A. C. *Statistical models*. Cambridge: Cambridge University Press, 2003.

DE GROOT, M. H. *Optimal Statistical Decisions*. New York: McGraw–Hill, 1970.

EMBRECHTS, P.; KLÜPPELBERG, C.; MIKOSCH, T. *Modelling extremal events for insurance and finance*. Berlin: Springer, 1997.

GAVASAKAR, P. Modal elicitation for the beta-binomial model. *Communications in Statistics – Theory and Methods*, v. 17, n. 9, p. 2989–3002, 1988.

GAVASAKAR, U. A comparison of two elicitation methods for a prior distribution for a binomial parameter. *Management Science*, v. 34, p. 784–790, 1988.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; DUNSON, D. B.; VEHTARI, A.; RUBIN, D. B. *Bayesian Data Analysis*. 3. ed. Boca Raton: Chapman and Hall/CRC, 2013.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; DUNSON, D. B.; VEHTARI, A.; RUBIN, D. B. *Bayesian data analysis*. 3. ed. Boca Raton: CRC Press, 2014.

GOSLING, J. P.; OAKLEY, J. E.; O'HAGAN, A. Nonparametric elicitation for heavy-tailed prior distributions. *Bayesian Analysis*, v. 2, n. 4, p. 841–866, 2007.

GUMBEL, E. J. *Statistics of extremes*. New York: Columbia University Press, 1958.

HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, v. 57, n. 1, p. 97–109, 1970. doi: `https://doi.org/10.1093/biomet/57.1.97`.

HOSKING, J. R. M.; WALLIS, J. R.; WOOD, E. F. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, v. 27, n. 3, p. 251–261, 1985.

JAYNES, E. T. *Probability Theory: The Logic of Science*. Cambridge: Cambridge University Press, 2003.

JEFFREYS, H. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, v. 186, n. 1007, p. 453–461, 1946.

JEFFREYS, H. *Theory of Probability*. 3. ed. Oxford: Oxford University Press, 1961.

KADANE, J. B. Predictive and structural methods for eliciting prior distributions. In: ZELLNER, A. (Ed.). *Bayesian Analysis in Econometrics and Statistics: Essays in Honor of Harold Jeffreys*. Amsterdam: North-Holland, 1980. p. 89–109.

KADANE, J. B.; WOLFSON, L. J. Experiences in elicitation. *The Statistician*, v. 47, n. 1, p. 3–19, 1998.

KASS, R. E.; TIERNEY, L.; KADANE, J. B. Approximate marginal densities of nonlinear functions. *Biometrika*, v. 75, n. 3, p. 425–433, 1988.

KASS, R. E.; TIERNEY, L.; KADANE, J. B. The validity of posterior expansions based on Laplace's method. In: GEISSER, S.; HODGES, J. S.; PRESS, S. J.; ZELLNER, A. (Eds.). *Bayesian and Likelihood Methods in Statistics and Econometrics*. Amsterdam: North-Holland, 1990. p. 473–488.

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.

KOTZ, S.; NADARAJAH, S. *Extreme value distributions: theory and applications*. London: Imperial College Press, 2000.

MARTZ, H. F.; WALLER, R. A. *Bayesian Reliability Analysis*. Krieger Publishing Company, 1991.

MCELREATH, R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. 2. ed. Boca Raton: CRC Press, 2020.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, v. 21, n. 6, p. 1087–1092, 1953. doi: https://doi.org/10.1063/1.1699114.

MEYER, M. A.; BOOKER, J. M. *Eliciting and Analyzing Expert Judgment: A Practical Guide*. London: Academic Press, 1991.

MIYATA, Y. Laplace approximations to means and variances with asymptotic modes. *Journal of Statistical Planning and Inference*, v. 140, n. 2, p. 382–392, 2010.

MOALA, F. A.; RODRIGUES, J. *Bayesian inference of the Weibull reliability function via Laplace approximation*. n. 36. São Carlos: Instituto de Ciências Matemáticas de São Carlos, Universidade de São Paulo, 1997. (Notas do ICMSC – Série Estatística).

MOALA, F. A.; DEY, S. Objective and subjective prior distributions for the Gompertz distribution. *Anais da Academia Brasileira de Ciências*, v. 90, p. 2643–2661, 2018. Disponível em: https://doi.org/10.1590/0001-3765201820171040. Acesso em: 8 out. 2025.

O'HAGAN, A. Eliciting expert beliefs in substantial practical applications. *Journal of the Royal Statistical Society: Series D (The Statistician)*, v. 47, n. 1, p. 21–35, 1998.

O'HAGAN, A.; BUCK, C.; DANESHKHAH, A.; EISER, J.; GARTHWAITE, P.; JENKINSON, D.; OAKLEY, J.; RAKOW, T. *Uncertain Judgements: Eliciting Expert Probabilities*. John Wiley & Sons, 2006. ISBN: 978-0-470-02999-2. doi: https://doi.org/10.1002/0470033312.

PAGOLA, P.; PÉREZ, E.; LÓPEZ, J. L. A simplification of Laplace's method for double integrals: application to the second Appell function. *Journal of Computational and Applied Mathematics*, v. 206, n. 1, p. 159–172, 2007.

PENHA, D. L. *Inferência bayesiana não-paramétrica para elicitação da função de confiabilidade*. 2014. Tese (Doutorado em Estatística) – Universidade de São Paulo, São Paulo, 2014.

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*, v. 7, n. 4, p. 493–497, 1992.

RAIFFA, H.; SCHLAIFER, R. *Applied Statistical Decision Theory*. Cambridge, MA: Harvard University Press, 1961.

ROBERT, C. P. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. New York: Springer, 2007.

ROBERT, C. P.; CASELLA, G. *Monte Carlo Statistical Methods*. 2. ed. New York: Springer, 2004.

RUGGERI, F.; BOUSQUET, N. *Robust Bayesian Analysis for Generalized Extreme Value Models*. submitted, 2017.

SPIEGELHALTER, D. J.; ABRAMS, K. R.; MYLES, J. P. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Chichester: John Wiley & Sons, 2004.

TIERNEY, L.; KADANE, J. B. Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, v. 81, n. 393, p. 82–86, 1986. Disponível em: http://www.jstor.org/stable/2287970. Acesso em: 8 out. 2025.

TIERNEY, L.; KASS, R. E.; KADANE, J. B. Fully Exponential Laplace Approximations to Expectations and Variances of Nonpositive Functions. *Journal of the American Statistical Association*, 84(407), 710–716, 1989. Disponível em: https://doi.org/10.2307/2289652.

VAN DE SCHOOT, R.; DEPAOLI, S.; KING, R.; KRAMER, B.; MÄRTENS, K.; TADESSE, M. G.; et al. Bayesian statistics and modelling. *Nature Reviews Methods Primers*, v. 1, n. 1, p. 1–26, 2021. Disponível em: https://doi.org/10.1038/s43586-020-00001-2. Acesso em: 8 out. 2025.

Apêndice

A

Elicitação via Aproximação de Laplace (1989)

Este apêndice apresenta a formulação alternativa do método de Laplace proposta por Tierney e Kadane (1989), a qual estende a abordagem clássica ao permitir a presença de fatores multiplicativos fora do termo exponencial na integral a ser aproximada.

Essa formulação constitui uma generalização natural do método apresentado no Capítulo 3, oferecendo maior flexibilidade analítica em problemas onde a função integranda não pode ser escrita exclusivamente na forma exponencial.

No entanto, quando aplicada ao contexto específico deste trabalho — em particular, à distribuição de Gumbel no processo de elicitação via preditiva a priori — essa abordagem conduz a uma limitação importante: a expressão final da aproximação não depende de certos hiperparâmetros da priori, especialmente do parâmetro b, resultando em um problema de não identificabilidade.

Em termos práticos, isso implica que diferentes valores de b produzem a mesma distribuição preditiva a priori, o que não é desejável no contexto da inferência bayesiana, pois impede a interpretação adequada da influência da priori sobre o modelo.

Por essa razão, a formulação de 1989 não foi adotada como método principal nesta dissertação. Ainda assim, sua apresentação é relevante do ponto de vista teórico e metodológico, sendo incluída neste apêndice para fins de completude e comparação conceitual.

Dessa forma, optou-se por utilizar a formulação de Laplace (1986) como abordagem principal ao longo deste trabalho. Ainda assim, a formulação de 1989 é apresentada neste apêndice por sua relevância teórica e por evidenciar diferenças importantes entre as abordagens.

A.1 Método proposto por Tierney e Kadane (1989)

Seja $L(\phi \mid \boldsymbol{x})$ a função de verossimilhança associada ao parâmetro escalar ϕ e $\pi(\phi)$ a densidade a priori. Para uma função mensurável $g(\phi)$, contínua e não-negativa, o momento a posteriori é definido por

$$E[g(\phi) \mid \boldsymbol{x}] = \frac{\displaystyle\int g(\phi)\, L(\phi \mid \boldsymbol{x})\, \pi(\phi)\, d\phi}{\displaystyle\int L(\phi \mid \boldsymbol{x})\, \pi(\phi)\, d\phi}. \tag{A.1.1}$$

Diferentemente da formulação apresentada em Tierney e Kadane (1986), o método proposto em Tierney e Kadane (1989) aborda diretamente a aproximação de razões de integrais por meio de expansões assintóticas mais gerais, permitindo a inclusão explícita de termos multiplicativos de ordem inferior e garantindo maior estabilidade numérica.

Aproximação de Laplace no caso unidimensional

Considere integrais da forma

$$I(n) = \int b(\phi) \exp\{-n\, h(\phi)\}\, d\phi, \tag{A.1.2}$$

onde $h : \mathbb{R} \to \mathbb{R}$ é duas vezes continuamente diferenciável, $-h(\phi)$ possui um único máximo em $\hat{\phi}$, $h''(\hat{\phi}) > 0$, e $b(\phi)$ é uma função contínua, estritamente positiva e de ordem constante em relação a n.

Expandindo $h(\phi)$ em série de Taylor de segunda ordem em torno de $\hat{\phi}$, a aproximação de Laplace conduz a

$$\int b(\phi) \exp\{-n\, h(\phi)\}\, d\phi \approx \sqrt{\frac{2\pi}{n\, h''(\hat{\phi})}}\, b(\hat{\phi})\, \exp\{-n\, h(\hat{\phi})\}, \tag{A.1.3}$$

com erro relativo da ordem $O(n^{-2})$, sob condições regulares (Tierney; Kadane, 1989).

Aproximação do momento a posteriori

No contexto do momento a posteriori em (A.1.1), o método de 1989 consiste em escrever separadamente o numerador e o denominador como integrais do tipo (A.1.2). Definimos

$$g(\phi)L(\phi \mid \boldsymbol{x})\pi(\phi) = b_N(\phi) \exp\{-n\, h_N(\phi)\}, \qquad L(\phi \mid \boldsymbol{x})\pi(\phi) = b_D(\phi) \exp\{-n\, h_D(\phi)\},$$

onde

$$h_D(\phi) = -\frac{1}{n} \log\big(L(\phi \mid \boldsymbol{x})\pi(\phi)\big), \qquad h_N(\phi) = h_D(\phi) - \frac{1}{n} \log g(\phi),$$

e $b_N(\phi)$ e $b_D(\phi)$ são funções suaves de ordem constante.

Aplicando a aproximação de Laplace (A.1.3) a cada integral, obtém-se

$$\int b_K(\phi) \exp\{-n\, h_K(\phi)\}\, d\phi \approx \sqrt{\frac{2\pi}{n\, h_K''(\hat{\phi}_K)}}\, b_K(\hat{\phi}_K)\, \exp\{-n\, h_K(\hat{\phi}_K)\}, \quad K \in \{N, D\}.$$

Consequentemente, o momento a posteriori é aproximado por

$$\widehat{E}[g(\phi) \mid \boldsymbol{x}] \approx \frac{b_N(\hat{\phi}_N)}{b_D(\hat{\phi}_D)} \left(\frac{h_D''(\hat{\phi}_D)}{h_N''(\hat{\phi}_N)}\right)^{1/2} \exp\Big\{-n(h_N(\hat{\phi}_N) - h_D(\hat{\phi}_D))\Big\}, \tag{A.1.4}$$

com erro da ordem $O(n^{-2})$.

Extensão para o caso multidimensional

Quando o vetor de parâmetros $\boldsymbol{\phi} \in \mathbb{R}^m$, a aproximação generaliza-se de forma natural. Considere integrais do tipo

$$\int b(\boldsymbol{\phi}) \exp\{-n\, h(\boldsymbol{\phi})\}\, d\boldsymbol{\phi},$$

onde $h(\boldsymbol{\phi})$ é diferenciável e possui um mínimo em $\hat{\boldsymbol{\phi}}$. A aproximação de Laplace resulta em

$$\int b(\boldsymbol{\phi}) \exp\{-n\, h(\boldsymbol{\phi})\}\, d\boldsymbol{\phi} \approx \left(\frac{2\pi}{n}\right)^{m/2} \det(\Sigma)^{1/2} b(\hat{\boldsymbol{\phi}}) \exp\{-n\, h(\hat{\boldsymbol{\phi}})\}, \tag{A.1.5}$$

onde $\Sigma^{-1} = \mathrm{D}^2 h(\hat{\boldsymbol{\phi}})$ é a matriz Hessiana de h avaliada em $\hat{\boldsymbol{\phi}}$.

Aplicação à elicitação via quantis preditivos

Considere o problema de elicitação a priori baseado em quantis da distribuição preditiva. Seja X uma variável aleatória com densidade condicional $f_X(x \mid \theta)$ e distribuição a priori $\pi(\theta \mid \alpha)$, em que $\theta = (\theta_1, \ldots, \theta_m)$ denota o vetor de parâmetros desconhecidos. O quantil preditivo x_p é definido implicitamente pela relação

$$P(X \le x_p) = \int_0^\infty \cdots \int_0^\infty \left[\int_0^{x_p} f_X(x \mid \theta)\, dx \right] \pi(\theta \mid \alpha)\, d\theta.$$

Definindo a função de distribuição condicional

$$F_X(x_p \mid \theta) = \int_0^{x_p} f_X(x \mid \theta)\, dx,$$

a probabilidade acumulada pode ser reescrita como

$$P(X \le x_p) = \int_0^\infty \cdots \int_0^\infty F_X(x_p \mid \theta)\, \pi(\theta \mid \alpha)\, d\theta.$$

A integral acima, em geral, não admite forma fechada, especialmente quando a dimensão do vetor de parâmetros é elevada. Para contornar essa dificuldade, recorre-se à aproximação de Laplace na forma proposta por Tierney e Kadane (1989), a qual permite a presença de um fator multiplicativo fora do termo exponencial.

Para esse fim, escrevemos a probabilidade acumulada na forma geral

$$P(X \le x_p) = \int b(\theta) \exp\{-n\, h(\theta)\}\, d\theta,$$

onde identificamos

$$b(\theta) = F_X(x_p \mid \theta), \qquad h(\theta) = -\frac{1}{n} \log \pi(\theta \mid \alpha).$$

Observe que a função $b(\theta)$ não depende de n e que $h(\theta)$ é suficientemente regular, satisfazendo as condições usuais para a aplicação da aproximação de Laplace. Seja $\hat{\theta}$ o ponto que maximiza $-h(\theta)$, ou equivalentemente, o ponto que maximiza a densidade a priori $\pi(\theta \mid \alpha)$. Denote ainda por

$$\Sigma^{-1} = \mathrm{D}^2 h(\hat{\theta})$$

a matriz Hessiana de $h(\theta)$ avaliada em $\hat{\theta}$.

De acordo com o resultado assintótico de Tierney e Kadane (1989), obtém-se a seguinte aproximação:

$$\int b(\theta)\exp\{-n\,h(\theta)\}\,d\theta \approx \left(\frac{2\pi}{n}\right)^{m/2} \det(\Sigma)^{1/2} b(\hat{\theta})\exp\{-n\,h(\hat{\theta})\}.$$

Aplicando esse resultado à probabilidade acumulada $P(X \leq x_p)$, resulta

$$P(X \leq x_p) \approx \left(\frac{2\pi}{n}\right)^{m/2} \det(\Sigma)^{1/2} F_X(x_p \mid \hat{\theta})\exp\{-n\,h(\hat{\theta})\}.$$

Finalmente, substituindo a definição de $h(\theta)$, obtemos

$$\exp\{-n\,h(\hat{\theta})\} = \pi(\hat{\theta} \mid \alpha),$$

de modo que a aproximação final assume a forma

$$P(X \leq x_p) \approx \left(\frac{2\pi}{n}\right)^{m/2} \det(\Sigma)^{1/2} F_X(x_p \mid \hat{\theta})\pi(\hat{\theta} \mid \alpha).$$

Essa expressão fornece uma aproximação assintótica direta para a probabilidade acumulada associada ao quantil preditivo x_p, constituindo a base para o procedimento de elicitação a priori sob a abordagem de Laplace com fator multiplicativo, conforme proposto por Tierney e Kadane (1989).

A.2 Formulação da Distribuição Preditiva

Considera-se o mesmo modelo apresentado no corpo principal do trabalho (ver Capítulo 5), no qual T segue uma distribuição de Gumbel com parâmetros ϕ e φ, e assume-se independência a priori entre esses parâmetros.

A probabilidade preditiva a priori é dada por

$$P(T \leq t_i) = \int_{-\infty}^{\infty}\int_{0}^{\infty} F(t_i \mid \phi, \varphi)\,\pi(\phi, \varphi)\,d\varphi\,d\phi, \tag{A.2.1}$$

a qual não admite forma fechada.

Substituindo a forma funcional correspondente, obtém-se

$$P(T \leq t_i) = \int_{-\infty}^{\infty}\int_{0}^{\infty} \exp\left\{-\exp\left(-\frac{t_i - \phi}{\varphi}\right)\right\}\pi(\phi, \varphi)\,d\varphi\,d\phi. \tag{A.2.2}$$

A.3 Reescrita na Forma de Laplace

Para aplicar a aproximação de Laplace, reescreve-se a integral na forma

$$\int g(\phi, \varphi)\exp\{-f(\phi, \varphi)\}\,d\phi d\varphi, \tag{A.3.1}$$

onde

$$f(\phi, \varphi) = d\varphi + \frac{(\phi - a)^2}{2b^2} - (c-1)\ln(\varphi), \tag{A.3.2}$$

e

$$g(\phi, \varphi) = \exp\left\{-\exp\left(-\frac{t_i - \phi}{\varphi}\right)\right\}. \tag{A.3.3}$$

A.4 Ponto Estacionário

O ponto estacionário é obtido resolvendo

$$\frac{\partial f}{\partial \phi} = \frac{1}{b^2}(\phi - a) = 0, \tag{A.4.1}$$

$$\frac{\partial f}{\partial \varphi} = d - \frac{c-1}{\varphi} = 0. \tag{A.4.2}$$

Daí,

$$\hat{\phi} = a, \qquad \hat{\varphi} = \frac{c-1}{d}. \tag{A.4.3}$$

A matriz Hessiana é

$$H(\phi, \varphi) = \begin{pmatrix} \frac{1}{b^2} & 0 \\ 0 & \frac{c-1}{\varphi^2} \end{pmatrix}, \tag{A.4.4}$$

com determinante

$$|H| = \frac{c-1}{b^2\varphi^2}. \tag{A.4.5}$$

A.5 Aproximação de Laplace

Aplicando a aproximação de Laplace para $m = 2$,

$$P(T \leq t_i) \approx \frac{d^c}{\Gamma(c)\sqrt{2\pi b^2}}(2\pi)|H(\hat{\phi}, \hat{\varphi})|^{-1/2} g(\hat{\phi}, \hat{\varphi}) \exp\{-f(\hat{\phi}, \hat{\varphi})\}. \tag{A.5.1}$$

Substituindo,

$$|H|^{-1/2} = \frac{b\hat{\varphi}}{\sqrt{c-1}}, \tag{A.5.2}$$

obtém-se

$$P(T \leq t_i) \approx \frac{d^c}{\Gamma(c)}\sqrt{2\pi}\frac{\hat{\varphi}}{\sqrt{c-1}} g(\hat{\phi}, \hat{\varphi}) \exp\{-f(\hat{\phi}, \hat{\varphi})\}. \tag{A.5.3}$$

Observa-se que o termo b desaparece da expressão final, evidenciando a ausência de dependência da preditiva a priori em relação a esse hiperparâmetro.

A.6 Obtenção dos Hiperparâmetros

Dada a não identificabilidade em relação a b, optou-se por fixar esse hiperparâmetro e estimar (a, c, d) por meio da minimização da discrepância entre as probabilidades preditivas aproximadas e os quantis elicitados.

Tabela A.6.1: Hiperparâmetros estimados via Laplace (1989).

Cenário	a	b	c	d
Verdadeiro	0.5000	0.3	1.3217	0.6436
Elicitada I	0.4500	0.3	1.2976	0.6613
Elicitada II	0.4249	0.3	1.2862	0.6736

A.7 Considerações Finais

A formulação de 1989 apresenta a vantagem de permitir a obtenção analítica do ponto estacionário, simplificando a aplicação da aproximação de Laplace. No entanto, essa simplificação conduz, no caso da distribuição de Gumbel, a um problema de não identificabilidade do hiperparâmetro b, comprometendo a flexibilidade do modelo.

Os resultados apresentados neste apêndice evidenciam que diferentes formulações da aproximação de Laplace podem produzir comportamentos distintos no processo de elicitação, reforçando a importância da escolha metodológica na inferência bayesiana.

Esse resultado reforça a escolha da formulação de 1986 como abordagem principal neste trabalho.