



Universidade Federal do Rio Grande do Norte
Centro de Ciências Exatas e da Terra
Programa de Pós-Graduação em Matemática
Aplicada e Estatística
Mestrado em Matemática Aplicada e Estatística



Inferência Bayesiana para modelos Poisson com priori conjugada baseada em misturas da distribuição Gama

Karine dos Santos Araújo

Natal - RN

Abril de 2025

Karine dos Santos Araújo

Inferência Bayesiana para modelos Poisson com priori conjugada baseada em misturas da distribuição Gama

Trabalho apresentado ao Programa de Pós-Graduação em Matemática Aplicada e Estatística da Universidade Federal do Rio Grande do Norte, em cumprimento com as exigências legais para obtenção do título de Mestre.

Área de Concentração: Probabilidade e Estatística.

Linha de Pesquisa: Modelos de Regressão.

Universidade Federal do Rio Grande do Norte

Centro de Ciências Exatas e da Terra

Programa de Pós-Graduação em Matemática Aplicada e Estatística

Mestrado em Matemática Aplicada e Estatística

Orientador

Prof. Dr. Marcelo Bourguignon Pereira

Natal - RN

Abril de 2025

Universidade Federal do Rio Grande do Norte - UFRN
Sistema de Bibliotecas - SISBI
Catalogação de Publicação na Fonte. UFRN - Biblioteca Setorial Prof. Ronaldo Xavier de Arruda - CCET

Araújo, Karine dos Santos.

Inferência bayesiana para modelos Poisson com priori conjugada baseada em misturas da distribuição Gama / Karine dos Santos Araújo. - 2025.

51 f.: il.

Dissertação (mestrado) - Universidade Federal do Rio Grande do Norte, Centro de Ciências Exatas e da Terra, Programa de Pós-Graduação em Matemática Aplicada e Estatística. Natal, RN, 2025.

Orientação: Prof. Dr. Marcelo Bourguignon Pereira.

1. Inferência bayesiana - Dissertação. 2. Misturas de Gama - Dissertação. 3. Priori conjugada - Dissertação. 4. Distribuição Poisson - Dissertação. I. Pereira, Marcelo Bourguignon. II. Título.

RN/UF/CCET

CDU 519.226(043.3)

Dissertação de Mestrado sob o título *Inferência Bayesiana para modelos Poisson com priori conjugada baseada em misturas da distribuição Gama* apresentada por Karine dos Santos Araújo e aceita pelo Programa de Pós-Graduação em Matemática Aplicada e Estatística da Universidade Federal do Rio Grande do Norte, sendo aprovada por todos os membros da banca examinadora abaixo especificada:

Prof. Dr. Marcelo Bourguignon Pereira

Orientador

Departamento de Estatística
Universidade Federal do Rio Grande do
Norte

**Prof. Dr. Fernando Ferraz do
Nascimento**

Departamento de Estatística
Universidade Federal do Piauí

Prof. Dr. Nelson Lima de Souza Filho

Departamento de Estatística
Universidade Federal do Amazonas

Natal - RN
Abril de 2025

Aos meus amados pais, Luzia e Genilson, pelo amor, apoio incansável e constante ao longo desta jornada acadêmica. Este trabalho é dedicado a vocês com profunda gratidão e admiração.

Agradecimentos

Inicio meus agradecimentos com o coração repleto de gratidão a Deus e à minha amada Nossa Senhora Aparecida, de quem sou devota, pela proteção constante e pelos cuidados silenciosos que nunca me faltaram. Sei que muitas das minhas orações, especialmente nos momentos de angústia e desânimo, foram ouvidas e acolhidas com amor.

Agradeço profundamente aos meus pais, Luzia e Genilson, por todo o amor, apoio e pelos ensinamentos que levarei para toda a vida. Foram vocês que me mostraram, desde cedo, que com esforço, dedicação e fé, eu poderia alcançar qualquer lugar. As lições que recebi de vocês são, sem dúvida, os conhecimentos mais valiosos que carrego.

Aos meus irmãos, pelo companheirismo e incentivo ao longo da jornada, e, de forma especial, à minha irmã Verônica, que sonha comigo, vibra com cada conquista e me ajuda a transformar sonhos em realidade.

Ao meu parceiro de vida e de caminhada, Lucas, obrigada por estar ao meu lado com tanto carinho, por acreditar em mim mesmo quando eu duvidei, por me apoiar nos dias difíceis e comemorar comigo cada pequena vitória.

Aos meus amigos, que tornaram o caminho mais leve e cheio de significado, meu sincero agradecimento por cada palavra de apoio, por cada momento de partilha e por todas as risadas que suavizaram os dias difíceis.

Ao meu orientador, agradeço pela paciência, pela escuta atenta e pela dedicação em cada orientação. Sua orientação foi essencial para a realização deste trabalho.

A todos que, de alguma forma, caminharam comigo até aqui, minha eterna gratidão.

*"É justo que muito custe,
o que muito vale".
(Santa Teresa D'Ávila)*

Resumo

A inferência Bayesiana é uma metodologia estatística que combina informações prévias sobre os parâmetros do modelo com dados observacionais para estimar a distribuição a posteriori dos parâmetros desconhecidos. Uma das vantagens de utilizar prioris conjugadas é que a distribuição a posteriori resultante permanece dentro da mesma família de distribuições da priori, o que facilita tanto os cálculos quanto a interpretação intuitiva dos parâmetros da posteriori. Neste estudo, adotamos misturas de distribuições Gama como prioris conjugadas para o parâmetro λ da distribuição Poisson, o que proporciona uma abordagem mais flexível e capaz de se ajustar melhor às diferentes características dos dados. As misturas de distribuições Gama exploradas neste trabalho incluem a distribuição Lindley generalizada 1 (ABOUAMMOH; ALSHANGITI; RAGAB, 2015), a distribuição Lindley generalizada 2 (RAMOS; LOUZADA; MOALA, 2021) e a distribuição Lindley generalizada 3 (ZAKERZADEH; DOLATI, 2009). Essas distribuições são extensões da clássica distribuição Lindley. Para ilustrar a aplicação prática da metodologia proposta, realizamos um estudo com dados reais.

Palavras-chave: Inferência Bayesiana. Misturas de Gama. Priori conjugada. Distribuição Poisson.

Abstract

Bayesian inference is a statistical methodology that combines prior information about model parameters with observational data to estimate the posterior distribution of unknown parameters. One of the advantages of using conjugate priors is that the resulting posterior distribution remains within the same family of distributions as the prior, which facilitates both the calculations and the intuitive interpretation of the posterior parameters. In this study, we adopted mixtures of Gamma distributions as conjugate priors for the λ parameter of the Poisson distribution, which provides a more flexible approach capable of better adjusting to different data characteristics. The mixtures of Gamma distributions explored in this work include the generalized Lindley distribution 1 ([ABOUAMMOH; ALSHANGITI; RAGAB, 2015](#)), the generalized Lindley distribution 2 ([RAMOS; LOUZADA; MOALA, 2021](#)) and the generalized Lindley distribution 3 ([ZAKERZADEH; DOLATI, 2009](#)). These distributions are extensions of the classic Lindley distribution. To illustrate the practical application of the proposed methodology, we carried out a study with real data.

Keywords: Bayesian inference. Gamma mixtures. Conjugate prior. Poisson Distribution.

Lista de Figuras

Figura 1 – Gráfico de densidade da GLD1.	26
Figura 2 – Gráfico de densidade da GLD2.	29
Figura 3 – Gráfico de densidade da GLD3.	32

Lista de Tabelas

Tabela 1 – Ataques de bombas voadoras em Londres durante a Segunda Guerra Mundial.	43
Tabela 2 – Comparação das Estimativas dos Modelos GLD1, GLD2, GLD3, Gama e WL: Parâmetros Teóricos, Amostrais, EMV, Erro Padrão, AIC, BIC e Intervalos de Credibilidade	45

Lista de abreviaturas e siglas

AIC	Informação de Akaike
BIC	Critério de Informação Bayesiano
EMV	Estimador de Máxima Verossimilhança
GLD1	Distribuição generalizada Lindley 1
GLD2	Distribuição generalizada Lindley 2
GLD3	Distribuição generalizada Lindley 3
MCMC	Monte Carlo via Cadeias de Markov
WL	Distribuição Lindley Ponderada

Lista de símbolos

α	Alpha
β	Beta
e	Constante de Euler
η	Eta
γ	gama
Γ	Gama
λ	Lambda
\in	Pertence
ϕ	Phi
τ	Tau
θ	Theta

Sumário

1	INTRODUÇÃO	14
1.1	Objetivo Geral	16
1.2	Objetivos Específicos	16
1.3	Organização do trabalho	16
2	INFERÊNCIA BAYESIANA	18
2.1	Teorema de Bayes	18
2.2	Função de verossimilhança	19
2.3	Misturas de Distribuições	20
2.4	Priori Conjugada	22
3	DISTRIBUIÇÕES LINDLEY GENERALIZADAS	25
3.1	Distribuição Lindley generalizada 1	25
3.1.1	Gerando amostras de uma distribuição GLD1	27
3.2	Distribuição Lindley generalizada 2	28
3.2.1	Gerando amostras de uma distribuição GLD2	30
3.3	Distribuição Lindley generalizada 3	31
3.3.1	Gerando amostras de uma distribuição GLD3	33
4	ABORDAGEM BAYESIANA	35
4.1	Distribuição Lindley generalizada 1	35
4.2	Distribuição Lindley generalizada 2	38
4.3	Distribuição Lindley generalizada 3	40
5	APLICAÇÃO	43
5.1	O número de ataques de bombas em Londres durante a Segunda Guerra Mundial	43
6	CONCLUSÃO	47
	REFERÊNCIAS BIBLIOGRÁFICAS	48

1 Introdução

Na estatística, destacam-se duas principais áreas de estudo: a estatística descritiva e a estatística inferencial. Ambos são extremamente importantes para a análise e interpretação de dados. A estatística descritiva auxilia os pesquisadores a entenderem as informações dos dados coletados por meio da organização e sumarização (MCHUGH, 2003). Ela permite obter e apresentar resultados como média amostral, mediana amostral, moda amostral, variância amostral e diversos gráficos, fornecendo uma visão geral dos dados por meio de organização, resumo e representação gráfica. Por outro lado, a estatística inferencial é utilizada para formular conclusões e fazer inferências com base na análise de dados coletados em pesquisas (TWYXCROSS; SHIELDS, 2004).

Dentro da estatística inferencial, há duas abordagens principais: a inferência clássica e a inferência bayesiana. A inferência clássica extrai todas as informações possíveis sobre a população a partir de amostra retirada da população. Em contraste, a inferência bayesiana permite tirar conclusões e fazer inferências com base em uma amostra retirada da população, incorporando informações prévias na análise.

A principal diferença entre a estatística bayesiana e a frequentista é que a bayesiana considera os parâmetros como variáveis. Todo parâmetro do modelo é tratado como uma variável aleatória, à qual é atribuída uma distribuição de probabilidade (YOKOO et al., 2013).

Nos últimos anos, a abordagem frequentista tem sido predominante na estimação de parâmetros, sendo caracterizada pelo uso exclusivo dos dados observados e evitando a incorporação de informações externas. No entanto, ambos os paradigmas, frequentista e bayesiano, apresentam vantagens e limitações que os tornam mais ou menos adequados dependendo de fatores como os objetivos do estudo, as características dos dados disponíveis e as condições específicas da análise. Por exemplo, a abordagem frequentista frequentemente enfrenta dificuldades em situações com amostras pequenas ou dados escassos, o que pode comprometer a precisão das estimativas. Em contrapartida, a abordagem bayesiana, apesar de oferecer flexibilidade em contextos de amostras reduzidas, é altamente sensível à escolha das distribuições a priori, que, se não forem cuidadosamente definidas, podem resultar em estimativas enviesadas (MCNEISH, 2016); (SMID et al., 2019). Estudos apontam que, em muitos casos de amostras pequenas, as correções de viés aplicadas nos métodos frequentistas podem oferecer melhores estimativas do que abordagens bayesianas que utilizam prioris não informativas (MESSAM et al., 2008).

Na inferência bayesiana, a incerteza sobre os parâmetros desconhecidos é representada por uma distribuição de probabilidade (BOX; TIAO, 1992). Ao contrário da

abordagem frequentista, a abordagem bayesiana permite a incorporação de informações prévias (conhecimentos prévios ou crenças) juntamente com os dados observados para a estimação de parâmetros. A inferência Bayesiana constitui uma alternativa coerente e abrangente à inferência clássica, baseada na utilização da verosimilhança e na adoção de distribuições a priori que exprimem o conhecimento ou a incerteza relativamente aos parâmetros em estudo (MURTEIRA et al., 2018).

Sendo $X \sim f(x|\theta)$, em que, θ é um parâmetro do modelo a ser estimado, segundo Gelman et al. (2013a), a inferência bayesiana se destaca ao permitir que as incertezas sobre parâmetros desconhecidos, como θ , sejam representadas por meio de um modelo probabilístico. Esse modelo incorpora informações a priori no processo de inferência, obtidas antes da observação dos dados. Desta forma, a inferência bayesiana oferece uma abordagem mais flexível para tratar a incerteza, especialmente quando há conhecimento prévio disponível.

O modelo Poisson é amplamente utilizado para analisar dados de contagem e tem aplicação em áreas como biologia (LAIRD; WARE, 1982), epidemiologia (LAWSON; BROWNE; RODEIRO, 2003), finanças (BOUCHER; DENUIT; GUILLÉN, 2009) e engenharia (RIGDON; BASU, 2000). No contexto da modelagem bayesiana, a distribuição Gama é frequentemente utilizada como a priori conjugada para o parâmetro da Poisson, que representa a taxa de ocorrência de eventos no espaço ou no tempo (GELMAN et al., 2013b). Isso ocorre porque a conjugação entre a Gama e a Poisson simplifica os cálculos da distribuição a posteriori, mantendo-se na mesma família de distribuições (BOX; TIAO, 1992). Entretanto, em situações mais complexas, onde os dados apresentam variação significativa ou múltiplas fontes de incerteza, uma única distribuição Gama pode não ser suficiente. Nesses casos, a mistura finita de distribuições Gama surge como uma alternativa mais flexível para capturar melhor as variações e comportamentos complexos dos dados (RAMOS; LOUZADA; MOALA, 2016).

As misturas finitas de distribuições Gama são especialmente úteis como prioris conjugadas devido à sua flexibilidade para modelar diferentes cenários de dados. Essa abordagem permite acomodar melhor a variabilidade e complexidade dos dados, o que é uma vantagem em relação ao uso de uma única distribuição Gama. Além disso, os modelos de mistura atraem atenção na literatura devido à sua versatilidade, com aplicações práticas e teóricas (BERGOT, 2006). Al-Saleh e Agarwal (2007a) demonstram que a mistura finita de distribuições Gama não apenas amplia a flexibilidade das prioris, mas também simplifica o processo de inferência para dados com estrutura hierárquica, especialmente quando se utiliza métodos de amostragem eficientes, como o Markov Chain Monte Carlo (MCMC).

Neste contexto, a presente pesquisa propõe explorar a utilização de misturas finitas de distribuições Gama como uma abordagem flexível para modelar a incerteza a priori no parâmetro do modelo Poisson. As misturas finitas de distribuições capturam

padrões complexos nos dados, sendo úteis para modelar heterogeneidade populacional, distribuições multimodais, caudas pesadas e mudanças de regime. Elas também lidam com variáveis latentes, tornando-se ferramentas essenciais em inferência bayesiana. A estrutura subjacente refere-se a fatores ocultos ou padrões nos dados que influenciam seu comportamento. Assim, a pesquisa visa contribuir para o avanço teórico da estatística bayesiana e encontrar uma alternativa para a modelagem de incertezas referentes ao parâmetro da distribuição Poisson.

1.1 Objetivo Geral

O objetivo principal deste trabalho é propor distribuições alternativas a priori conjugada para o modelo Poisson baseadas em misturas finitas especiais da distribuição Gama. Além disso, pretendemos estudar diversas propriedades dessas distribuições e aplicar os resultados em dados reais.

1.2 Objetivos Específicos

- Apresentar distribuições a priori conjugadas alternativas para o modelo Poisson, baseadas em misturas específicas da distribuição Gama;
- Analisar as propriedades das distribuições a priori propostas;
- Aplicar as distribuições a priori propostas em um estudo de casos reais;
- Realizar uma análise comparativa entre as distribuições a priori propostas.

1.3 Organização do trabalho

A estrutura desta dissertação está organizada para abordar de maneira abrangente e detalhada os temas centrais relacionados à inferência bayesiana e às questões específicas que estão sendo investigadas.

No Capítulo 2, a dissertação inicia com uma introdução detalhada à inferência bayesiana. Inicialmente, no primeiro tópico, aborda-se o Teorema de Bayes, que é o alicerce da inferência bayesiana. O capítulo também inclui uma discussão sobre priori conjugada, explicando como certos tipos de distribuições a priori simplificam os cálculos e facilitam a inferência bayesiana. Finalmente, o capítulo aborda o conceito de mistura finita de distribuições, explicando como a combinação de diferentes distribuições pode ser usada para modelar algumas situações.

O Capítulo 3 é dedicado à apresentação e análise de distribuições estatísticas que surgem de misturas finitas de distribuições Gama. Neste capítulo, detalham-se três

distribuições específicas, começando com uma descrição das suas principais propriedades. Além disso, abordam-se as origens dessas distribuições e suas principais aplicações práticas, fornecendo uma compreensão completa de como e por que essas distribuições são usadas.

O Capítulo 4 examina como a combinação da densidade com a verossimilhança da Poisson pode ser utilizada para diferentes distribuições. Analisam-se três distribuições específicas: distribuição generalizada de Lindley 1, distribuição generalizada de Lindley 2 e distribuição generalizada de Lindley 3. O capítulo detalha os métodos utilizados para realizar essa análise, apresentando os resultados da posteriori e as densidades marginais obtidas para cada uma das distribuições estudadas.

O Capítulo 5 aplica a metodologia desenvolvida em um contexto prático. Este capítulo apresenta um estudo de caso que utiliza dados reais. A aplicação da metodologia a esse conjunto de dados ilustra como os conceitos teóricos são aplicados na prática, demonstrando a eficácia e a relevância do estudo desenvolvido.

Finalmente, no Capítulo 6, discutem-se algumas considerações finais da dissertação. Este capítulo revisa os principais resultados obtidos ao longo do trabalho e oferece uma reflexão sobre suas implicações. Além disso, exploram-se possíveis aplicações futuras dos resultados e metodologias desenvolvidas, destacando como este trabalho contribui para avanços na inferência bayesiana e modelagem estatística.

2 Inferência Bayesiana

Neste capítulo, realizaremos uma revisão abrangente da literatura estatística, com o intuito de apresentar e discutir os principais conceitos e metodologias que serão fundamentais ao longo deste trabalho. Focaremos em quatro tópicos centrais: o Teorema de Bayes, Função de Verossimilhança, Priori Conjugada e as Misturas de Distribuições. A seguir, discutiremos cada um desses tópicos em detalhe.

2.1 Teorema de Bayes

O Teorema de Bayes, nomeado em homenagem ao matemático inglês Thomas Bayes (1701–1761), é um dos pilares centrais da estatística bayesiana. Essa abordagem, que utiliza o teorema para atualizar o conhecimento sobre os parâmetros de um modelo à medida que novos dados são observados (GELMAN et al., 2013c). Segundo Berger (1985) o Teorema de Bayes estabelece uma estrutura formal para relacionar probabilidades condicionais, permitindo a atualização contínua de crenças sobre um parâmetro de interesse com a incorporação de novas evidências. Dessa forma, as probabilidades associadas a um parâmetro após a observação de dados são refinadas com base nas informações obtidas (BERNARDO; SMITH, 2000; ROBERT, 2007). Quanto maior a complexidade dos dados, mais relevante se torna essa abordagem, em grande parte devido à flexibilidade na incorporação de informações a priori. Essa característica torna a modelagem bayesiana adequada para uma ampla gama de contextos, como a análise de grandes volumes de dados ou a modelagem de situações em que o conhecimento prévio é valioso (MURPHY, 2012; RAIFFA, 1968).

A metodologia bayesiana tem sido amplamente aplicada em áreas como economia (LUCAS; SCHWAAB; ZHANG, 2015), medicina (BROEMELING, 2017) e genética (GELMAN et al., 2004). A flexibilidade da metodologia bayesiana resulta em boas estimativas pontuais e uma objetiva forma de estimação intervalar dos parâmetros via intervalos de credibilidade (FALCÃO et al., 2004; BOLSTAD, 2007; ROBERT; CASELLA, 2004).

Considere θ como um parâmetro de interesse desconhecido associado a distribuição de uma variável aleatória X . As informações sobre θ são inicialmente resumidas pela distribuição de probabilidade $p(\theta)$, distribuição a priori, que reflete o conhecimento prévio sobre θ (LEE, 2012). Já $p(x|\theta)$ representa a distribuição condicional de X dado θ . Quando observamos $X = x$, o conhecimento disponível sobre θ aumenta. No contexto da inferência bayesiana, o Teorema de Bayes pode ser formalmente expresso como

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(\theta|x)d\theta}, \quad (2.1)$$

em que

- $p(\theta|x)$: A distribuição a posteriori de θ dado x .
- $p(\theta)$: A distribuição a priori de θ , que representa o conhecimento prévio ou as crenças sobre o parâmetro θ antes de observar os dados x .
- $p(x | \theta)$ é a verossimilhança, indicando a plausibilidade dos dados x para um dado valor do parâmetro θ .
- $p(x)$: é definido como a densidade marginal das observações independente dos parâmetros, também conhecida como distribuição preditiva a priori ou constante de normalização.

Essa formulação tem implicações fundamentais, especialmente em cenários onde a inferência sobre eventos desconhecidos deve se basear em informações observadas (JAYNES, 2003). Conforme indicado na Equação (2.1), para um valor fixo de x , $p(x | \theta)$ define a verossimilhança para diferentes valores de θ , enquanto $p(\theta)$ reflete o conhecimento prévio com relação ao parâmetro θ . Ao combinar essas duas informações, é possível calcular a distribuição a posteriori de θ . A constante de normalização $p(x)$ não depende de θ e, portanto, podemos ignorá-la ao considerar a proporcionalidade. Assim, a forma simplificada do teorema é expressa como

$$p(\theta | x) \propto p(x | \theta)p(\theta),$$

em que $p(x | \theta)$ é a função de verossimilhança.

Essa expressão destaca que a posteriori é proporcional ao produto da verossimilhança e da priori (GELMAN et al., 2004; ROBERT, 2007), evidenciando como a informação sobre θ é refinada à medida que mais dados são analisados. A atualização das crenças ocorre de maneira sistemática e rigorosa com base nas evidências observadas.

2.2 Função de verossimilhança

A função de verossimilhança é um conceito amplamente utilizado na estatística inferencial utilizada na estimação de parâmetros de interesse. Considere uma amostra aleatória simples X_1, X_2, \dots, X_n , de tamanho n , composta por variáveis aleatórias independentes e identicamente distribuídas, provenientes da distribuição de uma variável aleatória X com função de densidade de probabilidade $f(x|\theta)$, onde θ é um vetor de parâmetros

populacionais desconhecidos. A função de verossimilhança, denotada por $p(x | \theta)$, é dada por

$$p(x | \theta) = \prod_{i=1}^n f(x_i | \theta), \quad \theta \in \Theta, \quad (2.2)$$

em que Θ é o espaço paramétrico dos possíveis valores de θ e $\mathbf{X} = (x_1, x_2, \dots, x_n)^\top$ é o vetor que representa a amostra aleatória.

Muitas vezes, pode ser mais conveniente trabalhar diretamente com a função de log-verossimilhança. A maximização da log-verossimilhança é equivalente à maximização da verossimilhança, pois o logaritmo é uma função estritamente crescente. Portanto, os pontos onde a log-verossimilhança atinge seu máximo são os mesmos pontos onde a função de verossimilhança também atinge seu máximo. A log-verossimilhança é dada por,

$$p(\theta; \mathbf{X}) = \log(p(x | \theta)) = \log\left(\prod_{i=1}^n f(x_i | \theta)\right) = \sum_{i=1}^n \log(f(x_i | \theta)).$$

2.3 Misturas de Distribuições

Na literatura estatística, encontramos diversos livros que tratam de modelos de misturas, como os de [McLachlan e Basford \(1988\)](#), [Lindsay \(1995\)](#), [Fruhwirth-Schnatter \(2006\)](#), [Mengersen, Robert e Titterington \(2011\)](#) e [McNicholas \(2017\)](#). Os modelos de misturas são famílias de distribuições que combinam várias distribuições de probabilidade para descrever de maneira mais flexível e realista a variabilidade dos dados observados. A mistura de distribuições são particularmente úteis em contextos onde os dados apresentam uma estrutura complexa ou comportamento multimodal, o que dificulta o ajuste adequado por meio de uma única distribuição. Nesse sentido, a modelagem por meio de uma combinação ponderada de distribuições permite maior flexibilidade e melhor representação da variabilidade observada. A mistura finita de distribuição resulta em uma densidade combinada ponderada das densidades que caracterizam essas distribuições mencionadas anteriormente. Esse conceito é amplamente empregado em diversas áreas para modelar fenômenos que demandam uma abordagem mais flexível do que a oferecida por uma única distribuição. Os modelos de misturas têm desempenhado um papel importante nas teorias das distribuições ([PATIL; RAO, 1978](#)).

[Everitt e Hand \(1981\)](#) destacam que as distribuições de mistura finita são amplamente aplicáveis na modelagem de dados que contêm subgrupos distintos, sendo particularmente úteis em análises de agrupamento e na estimação do número de componentes envolvidos.

Suponha que tenhamos k distribuições de probabilidade $f_1(x|\lambda_1), f_2(x|\lambda_2), \dots, f_k(x|\lambda_k)$, em que cada $f_i(x|\lambda_i)$ representa a função de densidade de uma variável aleatória X . Ou seja, assumimos que

$$X \sim f_i(x|\lambda_i), \quad i = 1, \dots, k.$$

Cada λ_i é um parâmetro associado à i -ésima distribuição, pertencente a um espaço paramétrico Θ adequado ao tipo de distribuição considerada. Essas distribuições podem modelar diferentes subgrupos da população ou capturar variações estruturais nos dados. Uma mistura finita de distribuições permite que os dados sejam modelados por uma combinação de diferentes distribuições, capturando a heterogeneidade dos dados ao permitir que eles provenham de diferentes subpopulações. A densidade da mistura finita de k distribuições pode ser representada pela seguinte forma geral

$$f(x) = \sum_{i=1}^k p_i f_i(x|\lambda_i),$$

em que p_i representa a proporção de mistura associada à i -ésima distribuição, com $\sum_{i=1}^k p_i = 1$ e $0 < p_i < 1$. Assim, λ_i são os vetores de parâmetros associados a cada distribuição $f_i(\cdot)$. O espaço paramétrico de $\lambda_1, \lambda_2, \dots, \lambda_k$ depende das distribuições envolvidas. Por exemplo, se $f_i(\cdot)$ for uma distribuição Gama, o espaço paramétrico será definido pelos seus parâmetros específicos, como forma e escala. Portanto, a escolha adequada desses parâmetros é essencial para capturar a estrutura subjacente dos dados.

[Dalal e Hall \(1983\)](#) demonstram que qualquer densidade a priori para uma família exponencial pode ser aproximada com alta precisão por uma mistura de distribuições conjugadas. No contexto de misturas finitas, como descrito por [Al-Saleh e Agarwal \(2007b\)](#), essa abordagem permite uma grande flexibilidade, especialmente ao modelar dados complexos com múltiplas fontes de variação. No artigo mencionado, os autores exploram a mistura finita de distribuições Gama como uma priori conjugada, para um número arbitrário de componentes, cada um com seu próprio conjunto de parâmetros. Os autores aplicam métodos como o MCMC e o amostrador de Gibbs para estimar a distribuição a posterior dos parâmetros em modelos hierárquicos, ilustrando a utilidade da misturas de distribuições Gama em contextos variados de inferência estatística.

A mistura de funções de densidade da distribuição Gama tem sido usada para descrever a heterogeneidade e dados ([MAYROSE; FRIEDMAN; PUPKO, 2005](#)). Elas oferecem uma maneira flexível de capturar múltiplos padrões e estruturas nos dados, permitindo uma análise mais precisa e uma modelagem mais realista. [Niekerk, Bekker e Arashi \(2017\)](#) propõe uma abordagem bayesiana subjetiva para estimar parâmetros em modelos normais, utilizando uma classe de misturas de distribuições Gama. Os resultados obtidos através de simulações e dados reais demonstram as vantagens práticas desse modelo na inferência bayesiana. [PÉREZ \(2009\)](#) discute uma análise bayesiana das famílias exponenciais naturais com função de variância quadrática, utilizando uma mistura de distribuições a priori para representar diferentes fontes de informação ([PÉREZ, 2009](#)). [Bochkina e Rousseau \(2016\)](#) propõem um estimador bayesiano hierárquico baseado em uma mistura de distribuições Gama para estimar densidades não-limitadas, demonstrando taxas de convergência adaptáveis conforme a suavidade da densidade subjacente.

2.4 Priori Conjugada

Na inferência bayesiana, segundo [Amir-Ahmadi, Matthes e Wang \(2016\)](#), a distribuição a priori representa o conhecimento prévio sobre o parâmetro de interesse θ , expresso por meio de uma distribuição de probabilidade, antes da observação dos dados. Nesse contexto, ao utilizar uma abordagem de priori conjugada, a ideia é que as distribuições a priori e a posteriori pertençam à mesma classe de distribuições, o que facilita a atualização do conhecimento sobre o parâmetro de interesse por meio da modificação dos hiperparâmetros. Os hiperparâmetros são parâmetros da distribuição a priori. Além disso, a escolha desses hiperparâmetros é crucial para influenciar a forma da distribuição a priori, conforme discutido por [Arnst e Soize \(2019\)](#), especialmente em problemas de alta dimensão, onde a atualização dos parâmetros se baseia em matrizes simétricas positivas-definidas. Essa mudança nos hiperparâmetros reflete em uma atualização na crença sobre a distribuição de θ , influenciando assim a forma e as propriedades da distribuição a priori ([GELMAN et al., 2013b](#)). Além disso, quando utilizamos a família conjugada, podemos aproveitar a estrutura conhecidas das densidades a priori e a posteriori para determinar a distribuição marginal dos dados.

Além da definição de distribuições a priori e a importância dos hiperparâmetros, é fundamental considerar que a escolha dos hiperparâmetros não é apenas técnica, mas também envolve decisões sobre o grau de confiança no conhecimento prévio. Segundo [Griffin e Brown \(2010\)](#), a seleção desses hiperparâmetros pode impactar significativamente o desempenho do modelo, ajustando a quantidade de suavização ou regularização aplicada aos parâmetros.

Definição 1 : *Seja F uma classe de distribuições de amostragem $f(x|\theta)$ e P uma classe de distribuições a priori $p(\theta)$. Dizemos que P é **conjugada** a F se, para qualquer $f(\cdot | \theta) \in F$ e $p(\cdot) \in P$, a distribuição posterior $p(\theta|x)$ também pertence a P .*

Suponha uma amostra aleatória X_1, \dots, X_n independentes e identicamente distribuída a partir de uma distribuição de Poisson(λ), em que, $\lambda > 0$, a média, é desconhecida. Sendo assim, a função de distribuição é,

$$f(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, x = 0, 1, \dots$$

A função de verossimilhança, Equação (2.2), da distribuição Poisson é dada por,

$$\begin{aligned} l(\lambda; x) &= f(x|\lambda) = \prod_{i=1}^n f(x_i|\lambda) = \prod_{i=1}^n \left(\frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \right) \\ &= \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \propto \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}. \end{aligned} \tag{2.3}$$

Para encontrar o estimador de máxima verossimilhança (EMV) de λ , tomamos o logaritmo da função de verossimilhança,

$$\log l(\lambda; x) = \sum_{i=1}^n x_i \log \lambda - n\lambda.$$

Derivando em relação a λ e igualando a zero,

$$\frac{d}{d\lambda} \log l(\lambda) = \frac{\sum_{i=1}^n x_i}{\lambda} - n = 0.$$

A raiz da equação anterior é o ponto de inflexão dado por

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}.$$

Agora, para garantir que $\hat{\lambda}$ seja de fato um máximo, analisamos a segunda derivada da log-verossimilhança,

$$\frac{d^2}{d\lambda^2} \log l(\lambda) = -\frac{\sum_{i=1}^n x_i}{\lambda^2}.$$

Como $\frac{d^2}{d\lambda^2} \log l(\lambda)$ é sempre negativa para $\lambda > 0$, concluímos que $\hat{\lambda}$ é um máximo local, confirmando que o EMV de λ é a média amostral.

Agora, vamos considerar uma distribuição a priori conjugada amplamente utilizada na modelagem de dados que seguem uma distribuição de Poisson, a distribuição Gama. A principal vantagem de escolher a distribuição Gama(α, β), em que, α é o parâmetro de forma e β é o parâmetro de taxa, como a priori conjugada é que ela mantém a mesma forma funcional tanto para a distribuição a priori quanto para a distribuição a posteriori. Isso significa que, após observar os dados, podemos atualizar nossas crenças sobre o parâmetro λ de maneira simples, facilitando o processo de inferência.

$$p(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x, \alpha, \beta > 0,$$

em que função $\Gamma(\alpha)$, conhecida como função gama, é definida por

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

Usando o Teorema de Bayes, obtemos

$$\begin{aligned} p(x | x_1, \dots, x_n) &= \prod_{i=1}^n \frac{x_i^{x_i} e^{-x_i}}{x_i!} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \\ &= \frac{x^{\sum_{i=1}^n x_i} e^{-nx}}{\prod_{i=1}^n x_i!} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \\ &= \frac{e^{-nx} x^{\sum_{i=1}^n x_i} \beta^\alpha x^{\alpha-1} e^{-\beta x}}{\prod_{i=1}^n x_i! \Gamma(\alpha)} \\ &= \frac{e^{-x(n+\beta)} x^{\sum_{i=1}^n x_i + \alpha - 1} \beta^\alpha}{\prod_{i=1}^n x_i! \Gamma(\alpha)} \\ &\propto e^{-x(n+\beta)} x^{\sum_{i=1}^n x_i + \alpha - 1} \end{aligned}$$

Portanto, pela Definição (1), a distribuição a priori Gama é conjugada à distribuição Poisson com parâmetros $\alpha' = \sum_{i=1}^n x_i + \alpha$ e $\beta' = n + \beta$.

Considere agora uma mistura de duas distribuições Gama como priori para o parâmetro λ , descrito por,

$$f(\lambda \mid \alpha_1, \beta_1, \alpha_2, \beta_2, p) = p f_1(\lambda \mid \alpha_1, \beta_1) + (1 - p) f_2(\lambda \mid \alpha_2, \beta_2),$$

em que f_1 e f_2 são densidades da distribuição Gama definidas como

$$f_i(\lambda \mid \alpha_i, \beta_i) = \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \lambda^{\alpha_i-1} e^{-\beta_i \lambda}, \quad i = 1, 2,$$

e $p \in [0, 1]$ representa o peso da mistura. Nesse contexto, $\alpha_1, \beta_1, \alpha_2, \beta_2$, em que $\alpha_1, \beta_1 > 0$ e $\alpha_2, \beta_2 > 0$, são os parâmetros dos componentes compõem a mistura. Combinando a priori acima com a verossimilhança (2.3) da distribuição de Poisson, temos que a distribuição de λ dado um conjunto de observações x_1, x_2, \dots, x_n , é proporcional a,

$$f(\lambda \mid x_1, \dots, x_n) \propto \left[p \lambda^{\alpha_1-1} e^{-\beta_1 \lambda} + (1 - p) \lambda^{\alpha_2-1} e^{-\beta_2 \lambda} \right] \cdot \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}.$$

Ao simplificar a expressão, obtemos,

$$f(\lambda \mid x_1, \dots, x_n) \propto p \lambda^{\alpha_1 + \sum_{i=1}^n x_i - 1} e^{-\lambda(\beta_1 + n)} + (1 - p) \lambda^{\alpha_2 + \sum_{i=1}^n x_i - 1} e^{-\lambda(\beta_2 + n)}.$$

Essa forma demonstra que a posteriori é uma mistura de duas distribuições Gama, cada uma com parâmetros atualizados de acordo com as observações. Os parâmetros das duas distribuições Gama na mistura posterior são,

Para f_1 :

$$\alpha'_1 = \alpha_1 + \sum_{i=1}^n x_i, \quad \beta'_1 = \beta_1 + n.$$

Para f_2 :

$$\alpha'_2 = \alpha_2 + \sum_{i=1}^n x_i, \quad \beta'_2 = \beta_2 + n.$$

Portanto, a mistura de duas distribuições Gama mantém a propriedade de conjugação no modelo Poisson, garantindo que a distribuição a posteriori pertença à mesma família funcional da distribuição a priori. Além disso, o parâmetro p , que representa a proporção entre as componentes da mistura, não requer estimação, pois sua definição é totalmente determinada pelos parâmetros da distribuição a priori, sendo fixado antes da observação dos dados.

3 Distribuições Lindley Generalizadas

Neste capítulo, exploraremos algumas distribuições de probabilidade que emergem das misturas de distribuições Gama. Especificamente, as distribuições pioneiramente introduzidas por pesquisadores como, [Abouammoh, Alshangiti e Ragab \(2015\)](#), [Ramos, Louzada e Moala \(2021\)](#) e [Zakerzadeh e Dolati \(2009\)](#). Ao longo deste capítulo, iremos não apenas explorar as aplicações fundamentais dessas distribuições, mas também apresentar suas propriedades. As distribuições podem ser adaptadas como distribuições a priori conjugadas para diferentes tipos de verossimilhança, como Poisson, exponencial, normal (com média conhecida), Pareto (com mínimo conhecido), Normal (com média μ conhecida), Log-normal (com média μ conhecida), Inversa Gama (com forma k conhecida), Weibull (com forma a conhecida) e Gaussiana Inversa (com média μ conhecida).

3.1 Distribuição Lindley generalizada 1

A distribuição Lindley generalizada 1 (GLD1) foi introduzida por [Abouammoh, Alshangiti e Ragab \(2015\)](#) como uma extensão da distribuição de Lindley original. Essa generalização amplia significativamente a flexibilidade do modelo ao incorporar parâmetros adicionais que permitem ajustar a assimetria, a curtose e o comportamento das caudas da distribuição, proporcionando melhor adequação a diferentes tipos de dados empíricos. A GLD1 é baseada na mistura de duas distribuições Gamas, a Gama(η, θ) e a Gama(τ, θ), com probabilidades de mistura $p = \frac{\theta^\phi}{\gamma + \theta^\phi}$ e $1 - p = \frac{\gamma}{\gamma + \theta^\phi}$, respectivamente.

Seja $T \sim \text{GLD1}(\gamma, \phi, \theta, \eta, \tau)$, então sua função densidade de probabilidade é dada por,

$$f(t \mid \gamma, \phi, \theta, \eta, \tau) = \frac{\gamma t^{\eta-1} \theta^\eta}{(\gamma + \theta^\phi) \Gamma(\eta)} e^{-\theta t} + \frac{\theta^{\phi+\tau} t^{\tau-1}}{(\gamma + \theta^\phi) \Gamma(\tau)} e^{-\theta t},$$

em que

- $\gamma \geq 1$ é um parâmetro de forma,
- $\phi \geq 1$ é outro parâmetro de forma,
- $\theta > 0$ é o parâmetro de taxa comum às distribuições Gama,
- $\eta > 0$ e $\tau > 0$ são os parâmetros de forma das duas distribuições Gama misturadas,
- $t \geq 0$ é a variável.

Para $\gamma = 1$, $\phi = 1$, $\eta = 1$ e $\tau = 2$, a GLD1 assume exatamente a forma da distribuição Lindley. A Figura (1) podemos observar três curvas de densidade com diferentes

combinações de parâmetros. O gráfico da função de densidade de probabilidade da distribuição GLD1 destaca a variabilidade de sua forma em resposta a diferentes configurações de parâmetros.

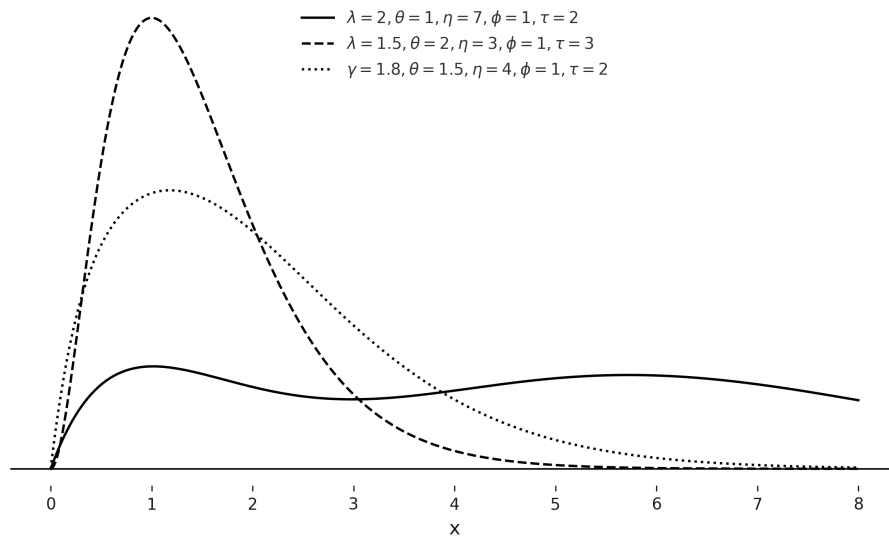


Figura 1 – Gráfico de densidade da GLD1.

Devido à mistura de duas distribuições Gama, a GLD1 é capaz de capturar vários comportamentos nos dados, o que a torna particularmente útil em aplicações como análise de sobrevivência e confiabilidade. Por exemplo, na análise de tempos de falha, a GLD1 pode modelar tanto cenários onde os eventos de falha ocorrem de forma rápida quanto cenários onde os eventos ocorrem de forma mais dispersa ao longo do tempo. Além disso, pela Figura (1) pode-se notar que a distribuição GLD1 pode assumir forma bimodal, diferentemente da distribuição Gama.

O valor esperado da variável aleatória T , que segue a GLD1, pode ser calculado como uma média ponderada dos valores esperados das distribuições componentes, com base nas probabilidades associadas. A expressão para o valor esperado é dada por

$$E(T) = pE(T_1) + (1 - p)E(T_2),$$

isso significa que $T_1 \sim \text{Gama}(\eta, \theta)$ e $T_2 \sim \text{Gama}(\tau, \theta)$, então $T \sim \text{GLD1}(\theta, \eta, \lambda, \tau)$, tal que $T = pT_1 + (1 - p)T_2$. Substituindo as expressões dos valores esperados das distribuições Gama, temos,

$$E(T) = \frac{\theta^\phi}{\gamma + \theta^\phi} \cdot \frac{\eta}{\theta} + \frac{\gamma}{\gamma + \theta^\phi} \cdot \frac{\tau}{\theta} = \frac{\eta\theta^\phi + \gamma\tau}{\theta(\theta^\phi + \gamma)}.$$

A variância de T pode ser expressa como,

$$\text{Var}(T) = \frac{\theta^\phi \eta}{(\gamma + \theta^\phi)\theta^2} + \frac{\gamma\tau}{(\gamma + \theta^\phi)\theta^2}.$$

Por fim, a moda da distribuição, que é o valor de T que maximiza a densidade de probabilidade, é dada por

$$\begin{aligned}
\frac{d}{dt}f(t; \gamma, \phi, \theta, \eta, \tau) &= \frac{d}{dt} \left(\frac{\gamma t^{\eta-1} \theta^\eta}{(\gamma + \theta^\phi) \Gamma(\eta)} e^{-\theta t} + \frac{\theta^\phi t^{\tau-1} \theta^\tau}{(\gamma + \theta^\phi) \Gamma(\tau)} e^{-\theta t} \right) = 0 \\
&= \frac{d}{dt} \left(\frac{\gamma t^{\eta-1} \theta^\eta}{(\gamma + \theta^\phi) \Gamma(\eta)} e^{-\theta t} \right) + \frac{d}{dt} \left(\frac{\theta^\phi t^{\tau-1} \theta^\tau}{(\gamma + \theta^\phi) \Gamma(\tau)} e^{-\theta t} \right) = 0 \\
&= \frac{\gamma \theta^\eta}{(\gamma + \theta^\phi) \Gamma(\eta)} \left((\eta - 1) t^{\eta-2} - \theta t^{\eta-1} \right) e^{-\theta t} + \frac{\theta^\phi \theta^\tau}{(\gamma + \theta^\phi) \Gamma(\tau)} \\
&\quad \left((\tau - 1) t^{\tau-2} - \theta t^{\tau-1} \right) e^{-\theta t} = 0 \\
&= e^{-\theta t} \left[\frac{\gamma \theta^\eta}{(\gamma + \theta^\phi) \Gamma(\eta)} \left((\eta - 1) t^{\eta-2} - \theta t^{\eta-1} \right) + \frac{\theta^\phi \theta^\tau}{(\gamma + \theta^\phi) \Gamma(\tau)} \right. \\
&\quad \left. \left((\tau - 1) t^{\tau-2} - \theta t^{\tau-1} \right) \right] = 0 \\
&= \frac{\gamma \theta^\eta}{(\gamma + \theta^\phi) \Gamma(\eta)} \left((\eta - 1) t^{\eta-2} - \theta t^{\eta-1} \right) + \frac{\theta^\phi \theta^\tau}{(\gamma + \theta^\phi) \Gamma(\tau)} \\
&\quad \left((\tau - 1) t^{\tau-2} - \theta t^{\tau-1} \right) = 0 \\
&= \gamma \theta^\eta \Gamma(\tau) \left((\eta - 1) t^{\eta-2} - \theta t^{\eta-1} \right) + \theta^\phi \theta^\tau \Gamma(\eta) \left((\tau - 1) t^{\tau-2} - \theta t^{\tau-1} \right) = 0 \\
&= \frac{1}{\theta} \ln \left(\frac{\Gamma(\eta) \Gamma(\tau) (\theta^\phi + \gamma)}{\theta \Gamma(\eta) + \gamma \Gamma(\tau)} + \theta^{\eta-1} \Gamma(\tau) \right).
\end{aligned}$$

Assim,

$$\text{Moda}(T) = \frac{1}{\theta} \ln \left(\frac{\Gamma(n) \Gamma(\tau) (\theta^\phi + \gamma)}{\theta \Gamma(n) + \gamma \Gamma(\tau)} + \theta^{n-1} \Gamma(\tau) \right).$$

3.1.1 Gerando amostras de uma distribuição GLD1

A geração de números aleatórios a partir da distribuição GLD1 é realizada por meio de uma mistura de duas distribuições Gama, na qual uma variável aleatória uniforme determina qual das componentes será utilizada para a geração da amostra. A variável aleatória X segue uma distribuição definida como a combinação de duas distribuições Gama com parâmetros distintos. A função densidade de probabilidade da GLD1 é dada pela soma ponderada das duas distribuições Gama, com as respectivas probabilidades dependentes dos parâmetros γ , θ e ϕ :

- A primeira componente segue uma distribuição Gama(α , θ).
- A segunda componente segue uma distribuição Gama(τ , θ).

Para simular uma amostra da distribuição GLD1, utiliza-se uma variável $U \sim \text{Uniforme}(0, 1)$, que define a escolha da componente de onde cada observação será extraída:

- Se $U \leq \frac{\gamma}{\gamma + \theta^\phi}$, então a amostra é gerada a partir da distribuição Gama(α, θ).
- Caso contrário, a amostra é gerada a partir da distribuição Gama(τ, θ).

A geração de cada amostra pode ser descrita pelo seguinte algoritmo:

Passo 1: Inicialize $i = 1$. Defina:

- O tamanho da amostra n ;
- Um vetor λ de tamanho n ;
- Os parâmetros $\alpha, \tau, \theta, \phi, \gamma$.

Passo 2: Calcule a probabilidade de escolha da primeira componente Gama:

$$p = \frac{\gamma}{\gamma + \theta^\phi}$$

Passo 3: Para cada $i = 1, 2, \dots, n$, repita:

- Gere $U \sim \text{Uniforme}(0, 1)$;
- Se $U \leq p$, então $\lambda_i \sim \text{Gama}(\alpha, \theta)$;
- Caso contrário, $\lambda_i \sim \text{Gama}(\tau, \theta)$;
- Armazene o valor gerado em $\lambda[i]$.

Passo 4: Após n repetições, o vetor λ conterá a amostra gerada da distribuição GLD1.

3.2 Distribuição Lindley generalizada 2

A distribuição Lindley generalizada 2 (GLD2), foi recentemente introduzida por [Ramos, Louzada e Moala \(2021\)](#) como uma extensão da distribuição Lindley. A principal motivação por trás dessa generalização é a necessidade de uma maior flexibilidade na modelagem de dados que apresentam uma função de taxa de risco crescente, incluindo aqueles com comportamentos assimétricos. Uma característica frequentemente observada em diversos contextos práticos, como na análise de sobrevivência e na confiabilidade de sistemas.

Seja $T \sim \text{GLD2}(\theta, \gamma)$ então sua densidade de probabilidade é dada por,

$$f(t; \theta, \gamma) = \frac{(\gamma - 2\theta)}{(\gamma - \theta)} \cdot \frac{\gamma^\theta}{\Gamma(\theta)} t^{\theta-1} e^{-\gamma t} + \frac{\theta}{(\gamma - \theta)} \cdot \frac{\gamma^{\theta+1}}{\Gamma(\theta + 1)} t^\theta e^{-\gamma t},$$

em que, $0 \leq \theta \leq \frac{\gamma}{2}$ é o parâmetro de forma e $\gamma > 0$ é o parâmetro de escala. A densidade é construída como uma mistura de duas distribuições do tipo Gama, a primeira com parâmetros (θ, γ) e a segunda com parâmetros $(\theta + 1, \gamma)$, isto é, uma Gama(θ, γ) e uma

Gama($\theta + 1, \gamma$). Os pesos da mistura, $\frac{(\gamma-2\theta)}{(\gamma-\theta)}$ e $\frac{\theta}{(\gamma-\theta)}$, garantem que a função densidade seja não-negativa e integrada a 1, assumindo $\theta < \gamma/2$. Quando $\theta = \gamma = 1$, a densidade da GLD2 coincide com a da distribuição de Lindley clássica.

Para ilustrar a forma e o comportamento da função de densidade da GLD2, a Figura (2) apresenta o gráfico correspondente, evidenciando picos mais pronunciados, caudas mais longas e maior flexibilidade na assimetria em comparação a distribuição Gama.

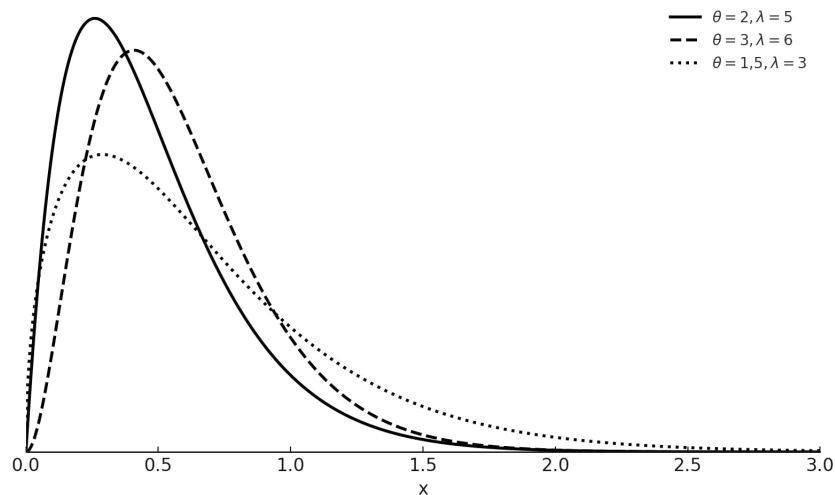


Figura 2 – Gráfico de densidade da GLD2.

Essa distribuição pode ser vista como uma mistura de duas distribuições Gama, o que lhe confere uma estrutura mais rica e a capacidade de capturar diferentes comportamentos dos dados. Especificamente, a mistura é composta por duas distribuições Gama, uma com parâmetros (θ, λ) e outra com parâmetros $(\theta + 1, \lambda)$, refletindo a generalização introduzida. A mistura é definida pelas probabilidades p e $1 - p$, dadas por

$$p = \frac{(\lambda - 2\theta)}{(\lambda - \theta)}, \quad 1 - p = \frac{\theta}{(\lambda - \theta)},$$

em que a probabilidade p corresponde à distribuição Gama com parâmetros (θ, λ) , enquanto $1 - p$ se refere à distribuição com parâmetros $(\theta + 1, \lambda)$. Essas probabilidades determinam a contribuição relativa de cada componentes da mistura, permitindo a modelagem em amplos cenários, incluindo aqueles com comportamentos assimétricos e variações complexas nos dados.

A GLD2 também possui expressões analíticas para seus momentos, o que é fundamental para entender as propriedades estatísticas da distribuição. O r -ésimo momento da distribuição é dado por

$$E(T^r) = \frac{\theta(\theta + 1) \cdots (\theta + r - 1)(\lambda - \theta + r)}{\lambda^r(\lambda - \theta)}, \quad r \in \mathbb{N}.$$

Além disso, a esperança matemática (ou valor esperado) da variável aleatória T , que segue a GLD2, pode ser obtida através da média ponderada dos valores esperados das distribuições Gama. A esperança é dada por,

$$E(T) = pE(T_1) + (1 - p)E(T_2),$$

isso significa que $T_1 \sim \text{Gama}(\theta, \lambda)$ e $T_2 \sim \text{Gama}(\theta + 1, \lambda)$, então $T \sim \text{GLD2}(\theta, \lambda)$, tal que $T = pT_1 + (1 - p)T_2$. Assim, $E(T_1)$ e $E(T_2)$ são os valores esperados das distribuições Gama subjacentes. Substituindo os valores das probabilidades e dos parâmetros, temos,

$$\begin{aligned} E(T) &= pE(T_1) + (1 - p)E(T_2) = p \cdot \frac{\theta}{\lambda} + (1 - p) \cdot \frac{(\theta + 1)}{\lambda} \\ &= \frac{(\lambda - 2\theta)}{(\lambda - \theta)} \cdot \frac{\theta}{\lambda} + \frac{\theta}{(\lambda - \theta)} \cdot \frac{(\theta + 1)}{\lambda} = \frac{\lambda\theta - \theta^2 + \theta}{(\lambda - \theta)\lambda}. \end{aligned}$$

A variância da GLD2, é dada por,

$$\text{Var}(T) = \frac{\theta(\theta + 1)(\lambda - \theta) + (\lambda - 2\theta)\theta}{(\lambda - \theta)^2\lambda^2}.$$

Por fim, a moda da distribuição, que indica o valor mais frequente ou provável da variável aleatória T , é calculada como:

$$\begin{aligned} \frac{d}{dt}f(t|\theta, \lambda) &= \frac{d}{dt} \left(\frac{\lambda^\theta}{(\lambda - \theta)\Gamma(\theta)} t^{\theta-1} (\lambda + \lambda t - 2\theta) e^{-\lambda t} \right) = 0 \\ &= \frac{\lambda^\theta}{(\lambda - \theta)\Gamma(\theta)} \left[\frac{d}{dt} \left(t^{\theta-1} (\lambda + \lambda t - 2\theta) e^{-\lambda t} \right) \right] = 0 \\ &= \frac{\lambda^\theta}{(\lambda - \theta)\Gamma(\theta)} \left[e^{-\lambda t} t^{\theta-2} \left(-2\theta^2 + \theta(2 + \lambda + 3t\lambda) - \lambda(1 + t\lambda + t^2\lambda) \right) \right] = 0 \\ &= \frac{\sqrt{\theta^2 - 2\lambda\theta + 8\theta + \lambda^2 - 4\lambda} + 3\theta - \lambda}{2\lambda}. \end{aligned}$$

Logo,

$$\text{Moda}(T) = \frac{\sqrt{\theta^2 - 2\lambda\theta + 8\theta + \lambda^2 - 4\lambda} + 3\theta - \lambda}{2\lambda}.$$

3.2.1 Gerando amostras de uma distribuição GLD2

A distribuição GLD2 é uma mistura de duas distribuições Gama, com a escolha da distribuição Gama a ser utilizada dependendo de uma variável aleatória uniforme. A GLD2 é composta por duas distribuições Gama, onde a primeira é $\text{Gama}(\theta, \lambda)$ e a segunda é $\text{Gama}(\theta + 1, \gamma)$. O mecanismo que utilizaremos para geração de variáveis aleatórias com distribuição GLD3 passa pela escolha de qual Gama irá determinar a variável aleatória em questão, para isso utilizamos uma variável aleatória com distribuição uniforme no intervalo $[0, 1]$.

A variável aleatória $U \sim \text{Uniforme}(0, 1)$ é gerada para cada amostra. Com base no valor de U , fazemos a escolha da distribuição Gama a ser utilizada para gerar a amostra:

- Se $U \leq \frac{\lambda-2\theta}{\lambda+\theta}$, a amostra será gerada a partir da distribuição Gama(θ, λ).
- Caso contrário, a amostra será gerada a partir da distribuição Gama($\theta + 1, \lambda$).

A geração de cada amostra pode ser descrita pelo seguinte algoritmo:

Passo 1: Inicialize $i = 1$. Defina:

- O tamanho da amostra n ;
- Um vetor λ de tamanho n ;
- Os parâmetros θ e γ , com $0 < \theta < \frac{\gamma}{2}$.

Passo 2: Calcule a probabilidade de escolha da primeira componente Gama:

$$p = \frac{\gamma - 2\theta}{\gamma + \theta}$$

Passo 3: Para cada $i = 1, 2, \dots, n$, execute:

- Gere $U \sim \text{Uniforme}(0, 1)$;
- Se $U \leq p$, então gere $\lambda_i \sim \text{Gama}(\theta, \gamma)$;
- Caso contrário, gere $\lambda_i \sim \text{Gama}(\theta + 1, \gamma)$;
- Armazene o valor gerado em $\lambda[i]$.

Passo 4: Após n repetições, o vetor λ conterá a amostra gerada da distribuição GLD2.

3.3 Distribuição Lindley generalizada 3

A distribuição Lindley Generalizada 3 (GLD3) foi introduzida por [Zakerzadeh e Dolati \(2009\)](#) como uma extensão da distribuição Lindley, com o objetivo de oferecer maior flexibilidade na modelagem de dados.

Seja $T \sim \text{GLD3}(\alpha, \theta, \gamma)$ então sua densidade de probabilidade é dada por,

$$f(t|\alpha, \theta, \gamma) = \frac{\theta^2(\theta t)^{\alpha-1}(\alpha + \gamma t)e^{-\theta t}}{(\gamma + \theta)\Gamma(\alpha + 1)},$$

em que $\theta \geq 0$ é o parâmetro de escala que controla dispersão da distribuição, $\alpha \geq 0$ é o parâmetro de forma que ajusta a inclinação da função de densidade de probabilidade, e $\gamma \geq 0$ é um parâmetro adicional que proporciona maior flexibilidade na forma da distribuição. A GLD3 se reduz à distribuição de Lindley quando $\alpha = 1$ e $\gamma = 0$. Com esses valores para os parâmetros, a função de densidade da GLD3 passa a ter a mesma forma da distribuição de Lindley, com θ sendo o único parâmetro da distribuição.

Para ilustrar a distribuição, a Figura (3) apresenta o gráfico da densidade de uma variável aleatória com distribuição GLD3 para diferentes valores dos parâmetros, evidenciando modas em diferentes posições, caudas com decaimentos variados. Esse comportamento ressalta como a forma da distribuição muda, para diferentes conjuntos de parâmetros, destacando a versatilidade da GLD3 em modelar diferentes padrões de dados.

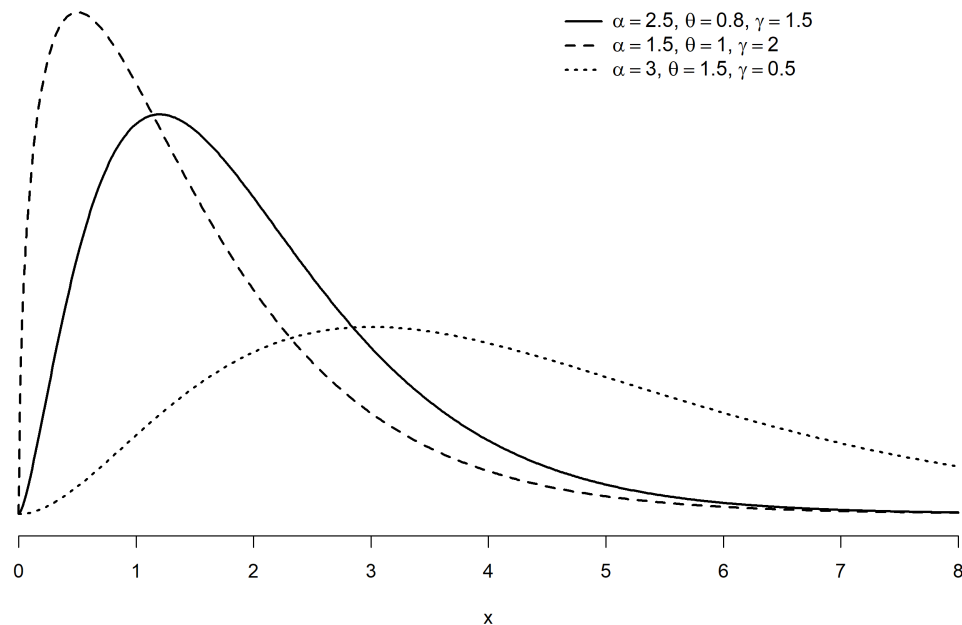


Figura 3 – Gráfico de densidade da GLD3.

Essa distribuição pode ser vista como uma mistura de duas distribuições Gama, especificamente as distribuições Gama (α, θ) e Gama $(\alpha+1, \theta)$. As probabilidades associadas a cada uma dessas componentes são $\frac{\theta}{\gamma+\theta}$ e $\frac{\gamma}{\gamma+\theta}$, respectivamente. Nessas distribuições, α e $\alpha+1$ representam os parâmetros de forma, enquanto θ continua a ser o parâmetro de escala.

A função de densidade de probabilidade pode ser reescrita utilizando a definição de mistura de distribuições, resultando na seguinte expressão,

$$f(x, \alpha, \theta, \gamma) = \frac{\theta}{\gamma + \theta} \cdot \frac{\theta^\alpha x^{\alpha-1} e^{-\theta x}}{\Gamma(\alpha)} + \frac{\gamma}{\gamma + \theta} \cdot \frac{\theta^{\alpha+1} x^\alpha e^{-\theta x}}{\Gamma(\alpha + 1)},$$

Essa formulação destaca a flexibilidade adicional da GLD3, permitindo uma modelagem mais precisa de dados com diferentes características de dispersão e assimetria.

A média da distribuição é dada por,

$$E(X) = \frac{\alpha\theta + \alpha\gamma + \gamma}{\theta(\gamma + \theta)}.$$

A variância da distribuição é dada por,

$$Var(X) = \frac{\alpha\theta^2 + 2\gamma\theta\alpha + \gamma^2\alpha + 2\gamma\theta + \gamma^2}{\theta^2(\gamma + \theta)^2}.$$

A moda da GLD3, que indica o valor mais frequente ou provável da variável aleatória X , é expressa por,

$$Moda(X) = \frac{-\alpha\theta + \alpha\gamma + \sqrt{\alpha^2\theta^2 + 2\alpha^2\theta\gamma + \alpha^2\gamma^2 - 4\alpha\theta\gamma}}{2\theta\gamma}.$$

O r -ésimo momento é dado por:

$$E(X^r) = \frac{\Gamma(r + \alpha)(\theta^{-r}\gamma\alpha + \theta^{-r}\gamma r + \theta^{-r+1}\alpha)}{(\gamma + \theta)\Gamma(\alpha + 1)}.$$

3.3.1 Gerando amostras de uma distribuição GLD3

A distribuição GLD3 é uma mistura de duas distribuições Gama. O mecanismo que utilizaremos para geração de variáveis aleatórias com distribuição GLD3 passa pela escolha de qual Gama irá determinar a variável aleatória em questão, para isso utilizaremos uma variável com distribuição Uniform no intervalo $[0,1]$. A GLD3 é composta por duas distribuições Gama, $Gama(\alpha, \theta)$ e $Gama(\alpha + 1, \theta)$.

A variável aleatória $U \sim \text{Uniforme}(0, 1)$ é gerada para cada amostra. Com base no valor de U , a distribuição de onde a amostra será retirada é escolhida:

- Se $U \leq \frac{\theta}{\gamma + \theta}$, a amostra será gerada a partir da distribuição $Gama(\alpha, \theta)$.
- Caso contrário, a amostra será gerada a partir da distribuição $Gama(\alpha + 1, \theta)$.

A geração de cada amostra pode ser descrita pelo seguinte algoritmo:

Passo 1: Inicialize $i = 1$. Defina:

- O tamanho da amostra n ;
- Um vetor λ de tamanho n ;
- Os parâmetros α , θ e γ , com $\alpha > 0$, $\theta > 0$ e $\gamma > 0$.

Passo 2: Calcule a probabilidade de escolha da primeira componente Gama:

$$p = \frac{\theta}{\gamma + \theta}$$

Passo 3: Para cada $i = 1, 2, \dots, n$, execute:

- Gere $U \sim \text{Uniforme}(0, 1)$;
- Se $U \leq p$, então gere $\lambda_i \sim \text{Gama}(\alpha, \theta)$;
- Caso contrário, gere $\lambda_i \sim \text{Gama}(\alpha + 1, \theta)$;
- Armazene o valor gerado em $\lambda[i]$.

Passo 4: Após n repetições, o vetor λ conterá a amostra gerada da distribuição GLD3.

4 Abordagem Bayesiana

A distribuição Poisson é amplamente empregada para modelar dados de contagem. Ao combinar a distribuição Poisson com prioris baseadas em misturas da distribuição Gama, como as Distribuições Lindley Generalizadas, buscamos explorar as propriedades dessas distribuições para aprimorar as estimativas do parâmetros de interesse.

4.1 Distribuição Lindley generalizada 1

Considere a densidade GLD1 como uma distribuição a priori para o parâmetro λ da distribuição Poisson. Para alinhar a análise à estrutura do modelo bayesiano em desenvolvimento, é necessário redefinir a variável de interesse como λ , conforme sugerido pela forma da densidade GLD1

$$\begin{aligned} f(\lambda|\gamma, \phi, \theta, \eta, \tau) &= \frac{\gamma\lambda^{\eta-1}\theta^\eta}{(\gamma + \theta^\phi)\Gamma(\eta)}e^{-\theta\lambda} + \frac{\theta^\phi\lambda^{\tau-1}\theta^\tau}{(\gamma + \theta^\phi)\Gamma(\tau)}e^{-\theta\lambda} \\ &\propto \lambda^{\eta-1}e^{-\theta\lambda} + \lambda^{\tau-1}e^{-\theta\lambda}. \end{aligned}$$

Considerando a verossimilhança da Poisson, abordada anteriormente na Equação (2.3) dada por,

$$f(x_1, \dots, x_n|\lambda) = \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \propto \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}.$$

Se combinarmos a densidade da GLD1 com a função de verossimilhança de Poisson, obtemos,

$$\begin{aligned} f(\lambda|x_1, \dots, x_n) &= \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \cdot \left(\frac{\gamma\lambda^{\eta-1}\theta^\eta}{(\gamma + \theta^\phi)\Gamma(\eta)}e^{-\theta\lambda} + \frac{\theta^\phi\lambda^{\tau-1}\theta^\tau}{(\gamma + \theta^\phi)\Gamma(\tau)}e^{-\theta\lambda} \right) \\ &= \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \gamma\lambda^{\eta-1}\theta^\eta e^{-\theta\lambda}}{\prod_{i=1}^n x_i! (\gamma + \theta^\phi)\Gamma(\eta)} + \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \theta^\phi\lambda^{\tau-1}\theta^\tau e^{-\theta\lambda}}{\prod_{i=1}^n x_i! (\gamma + \theta^\phi)\Gamma(\tau)} \\ &= \frac{\gamma\lambda^{\eta+\sum_{i=1}^n x_i-1} e^{-\lambda(n+\theta)} \theta^\eta}{\prod_{i=1}^n x_i! (\gamma + \theta^\phi)\Gamma(\eta)} + \frac{\lambda^{\tau+\sum_{i=1}^n x_i-1} e^{-\lambda(n+\theta)} \theta^{\phi+\tau}}{\prod_{i=1}^n x_i! (\gamma + \theta^\phi)\Gamma(\tau)} \\ &= \frac{\gamma\lambda^{\eta+\sum_{i=1}^n x_i-1} e^{-\lambda(n+\theta)} \theta^\eta}{\prod_{i=1}^n x_i! (\gamma + \theta^\phi)\Gamma(\eta)} + \frac{\lambda^{\tau+\sum_{i=1}^n x_i-1} e^{-\lambda(n+\theta)} \theta^{\phi+\tau}}{\prod_{i=1}^n x_i! (\gamma + \theta^\phi)\Gamma(\tau)} \\ &\propto \lambda^{\eta-1+\sum_{i=1}^n x_i} e^{-\lambda(n+\theta)} + \lambda^{\tau-1+\sum_{i=1}^n x_i} e^{-\lambda(n+\theta)}. \end{aligned}$$

Veja que, a distribuição resultante é uma mistura de distribuições Gama com parâmetros combinados. Cada termo da mistura pode ser visto como uma distribuição Gama, o primeiro termo segue uma distribuição Gama($\eta' = \eta + \sum_{i=1}^n x_i, \theta' = n + \theta$) e o

segundo termo segue uma distribuição Gama($\tau' = \tau + \sum_{i=1}^n x_i, \theta' = n + \theta$). Assim, podemos concluir que a distribuição resultante é uma GLD1, uma mistura dessas duas distribuições Gama, com os parâmetros τ' e θ' ajustados pela soma das observações $\sum_{i=1}^n x_i$, e a taxa θ' ajustada por $n + \theta$.

Para construir um intervalo de credibilidade para o parâmetro λ utilizando a distribuição GLD1 como priori no contexto bayesiano, seguimos um procedimento teórico fundamentado nos quantis da distribuição a posteriori. A GLD1, quando combinada com a verossimilhança da Poisson, resulta em uma distribuição a posteriori que é uma mistura de duas distribuições Gama.

Seja $f_{\text{Gama}(a,b)}(\lambda)$ a densidade da distribuição Gama com parâmetros a e b , dada por:

$$f_{\text{Gama}(a,b)}(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda}, \quad \lambda > 0.$$

A distribuição a posteriori de λ , denotada por $p(\lambda | \mathbf{x})$, é dada por uma mistura de duas distribuições Gama,

$$p(\lambda | \mathbf{x}) = w_1 \cdot f_{\text{Gama}(\eta+S, n+\theta)}(\lambda) + w_2 \cdot f_{\text{Gama}(\tau+S, n+\theta)}(\lambda),$$

em que, w_1 e w_2 são os pesos da mistura dados por

$$w_1 = \frac{\gamma\theta^\eta}{\gamma\theta^\eta + \theta^{\phi+\tau}}, \quad w_2 = 1 - w_1,$$

e $S = \sum_{i=1}^n x_i$ é a soma das observações.

Como a posteriori é uma mistura de distribuições Gama, o intervalo de credibilidade é determinado a partir dos quantis dessa mistura. Teoricamente, isso pode ser feito das seguintes formas, como a mistura envolve duas distribuições Gama, os quantis não podem ser obtidos diretamente de uma fórmula fechada. Em vez disso, eles são calculados considerando a combinação ponderada dos quantis de cada componente da mistura. Especificamente, o quantil de ordem α é obtido resolvendo-se:

$$Q_\alpha = F^{-1}(\alpha) = \inf \left\{ \lambda : \int_0^\lambda p(\lambda' | \mathbf{x}) d\lambda' \geq \alpha \right\}.$$

De forma prática, isso pode ser feito numericamente, utilizando métodos como a inversa da função empírica ou amostragem numérica.

Uma abordagem alternativa e amplamente utilizada é a amostragem a partir da distribuição a posteriori, devido à sua forma de mistura. Isso é feito da seguinte maneira

1. Fixe o número de repetições N , os pesos da mistura w_1 e $w_2 = 1 - w_1$, e os parâmetros das distribuições: η , τ , S , n e θ .

2. Para $i = 1, 2, \dots, N$, repita:

a) Gere um número aleatório u de uma distribuição uniforme no intervalo $(0, 1)$.

b) Se $u < w_1$, então:

• Amostre $\lambda_i \sim \text{Gama}(\eta + S, n + \theta)$.

c) Caso contrário:

• Amostre $\lambda_i \sim \text{Gama}(\tau + S, n + \theta)$.

3. Após obter as N amostras, ordene os valores gerados: $\lambda_{(1)} \leq \lambda_{(2)} \leq \dots \leq \lambda_{(N)}$.

4. Estime os quantis da distribuição a posteriori:

• $Q_{2.5\%} = \lambda_{([0.025N])}$,

• $Q_{97.5\%} = \lambda_{([0.975N])}$.

5. O intervalo de credibilidade de 95% para λ é:

$$[Q_{2.5\%}, Q_{97.5\%}], \quad \text{tal que} \quad P(Q_{2.5\%} \leq \lambda \leq Q_{97.5\%} \mid \mathbf{x}) = 0,95.$$

O modelo de Poisson para uma única observação, x , tem a seguinte distribuição preditiva a priori,

$$\begin{aligned} f(x) &= \int_0^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} \cdot \left(\frac{\gamma \lambda^{\eta-1} \theta^{\eta}}{(\gamma + \theta^{\phi}) \Gamma(\eta)} e^{-\theta \lambda} + \frac{\theta^{\phi} \lambda^{\tau-1} \theta^{\tau}}{(\gamma + \theta^{\phi}) \Gamma(\tau)} e^{-\theta \lambda} \right) d\lambda \\ &= \left(\frac{\gamma \theta^{\eta}}{x! (\gamma + \theta^{\phi}) \Gamma(\eta)} + \frac{\theta^{\phi}}{x! (\gamma + \theta^{\phi}) \Gamma(\tau)} \right) \int_0^{\infty} \lambda^x e^{-\lambda} \gamma \lambda^{\eta-1} \theta^{\eta} e^{-\theta \lambda} + \lambda^x e^{-\lambda} \theta^{\phi} \lambda^{\tau-1} \theta^{\tau} e^{\theta \lambda} d\lambda \\ &= \left(\frac{\gamma \theta^{\eta}}{x! (\gamma + \theta^{\phi}) \Gamma(\eta)} + \frac{\theta^{\phi}}{x! (\gamma + \theta^{\phi}) \Gamma(\tau)} \right) \int_0^{\infty} \lambda^x e^{-\lambda} \gamma \lambda^{\eta-1} \theta^{\eta} e^{-\theta \lambda} d\lambda + \\ &\quad \int_0^{\infty} \lambda^x e^{-\lambda} \theta^{\phi} \lambda^{\tau-1} \theta^{\tau} e^{\theta \lambda} d\lambda \\ &= \left(\frac{\gamma \theta^{\eta}}{x! (\gamma + \theta^{\phi}) \Gamma(\eta)} + \frac{\theta^{\phi}}{x! (\gamma + \theta^{\phi}) \Gamma(\tau)} \right) \int_0^{\infty} \gamma \lambda^{x+\eta-1} e^{-\lambda(1+\theta)} \theta^{\eta} d\lambda + \\ &\quad \int_0^{\infty} \lambda^{x+\tau-1} \theta^{\tau+\phi} e^{-\lambda(1+\theta)} d\lambda \\ &= \left(\frac{\gamma \theta^{\eta}}{x! (\gamma + \theta^{\phi}) \Gamma(\eta)} + \frac{\theta^{\phi}}{x! (\gamma + \theta^{\phi}) \Gamma(\tau)} \right) \left(\frac{\gamma \Gamma(x + \eta)}{(1 + \theta)^{x+\eta}} + \frac{\Gamma(x + \tau)}{(1 + \theta)^{x+\tau}} \right) \\ &= \left(\frac{\gamma \theta^{\eta}}{x! (\gamma + \theta^{\phi}) \Gamma(\eta)} + \frac{\theta^{\phi}}{x! (\gamma + \theta^{\phi}) \Gamma(\tau)} \right) \cdot \left(\frac{\Gamma(x + \eta)}{(1 + \theta)^{x+\eta}} \right) \\ &\quad + \left(\frac{\gamma \theta^{\eta}}{x! (\gamma + \theta^{\phi}) \Gamma(\eta)} + \frac{\theta^{\phi}}{x! (\gamma + \theta^{\phi}) \Gamma(\tau)} \right) \cdot \left(\frac{\gamma \Gamma(x + \tau)}{(1 + \theta)^{x+\tau}} \right) \\ &= \frac{\gamma \theta^{\eta}}{\gamma \theta^{\eta} + \theta^{\phi}} \cdot \frac{\Gamma(x + \eta)}{x! \Gamma(\eta)} \left(\frac{1}{1 + \theta} \right)^{\eta} \left(\frac{\theta}{1 + \theta} \right)^x + \frac{\theta^{\phi}}{\gamma \theta^{\eta} + \theta^{\phi}} \cdot \frac{\Gamma(x + \tau)}{x! \Gamma(\tau)} \left(\frac{1}{1 + \theta} \right)^{\tau} \left(\frac{\theta}{1 + \theta} \right)^x \end{aligned}$$

para $x = 0, 1, 2, \dots$, obtemos uma mistura de duas distribuições binomiais negativas. Os pesos da mistura são dados por $p_1 = \frac{\gamma\theta^\eta}{\gamma\theta^\eta + \theta^\phi}$ e $p_2 = \frac{\theta^\phi}{\gamma\theta^\eta + \theta^\phi}$, enquanto os parâmetros das distribuições binomiais negativas são $(\eta, \frac{\theta}{1+\theta})$ para o primeiro componente e $(\tau, \frac{\theta}{1+\theta})$ para o segundo.

4.2 Distribuição Lindley generalizada 2

Considere a seguinte densidade da GLD2 como priori para o parâmetro λ da distribuição Poisson. Também mudaremos a variável de interesse para λ , conforme indicado pela densidade da GLD2,

$$\begin{aligned} f(\lambda|\theta, \gamma) &= \frac{(\gamma - 2\theta)}{(\gamma - \theta)} \cdot \frac{\gamma^\theta}{\Gamma(\theta)} \lambda^{\theta-1} e^{-\gamma\lambda} + \frac{\theta}{(\gamma - \theta)} \frac{\gamma^{\theta+1}}{\Gamma(\theta + 1)} \lambda^\theta e^{-\gamma\lambda} \\ &\propto \lambda^{\theta-1} e^{-\gamma\lambda} + \lambda^\theta e^{-\gamma\lambda}. \end{aligned}$$

Dada a verossimilhança da Poisson, Equação (2.3), obtemos

$$\begin{aligned} f(\lambda|x_1, \dots, x_n) &= \frac{(\gamma - 2\theta)}{(\gamma - \theta)} \cdot \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \cdot \frac{\gamma^\theta}{\Gamma(\theta)} \lambda^{\theta-1} e^{-\gamma\lambda} + \frac{\theta}{(\gamma - \theta)} \cdot \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \\ &\quad \cdot \frac{\gamma^{\theta+1}}{\Gamma(\theta + 1)} \lambda^\theta e^{-\gamma\lambda} \\ &= \frac{(\gamma - 2\theta)}{(\gamma - \theta)} \cdot \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \gamma^\theta \lambda^{\theta-1} e^{-\gamma\lambda}}{\prod_{i=1}^n x_i! \Gamma(\theta)} + \frac{\theta}{(\gamma - \theta)} \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \gamma^{\theta+1} \lambda^\theta e^{-\gamma\lambda}}{\prod_{i=1}^n x_i! \Gamma(\theta + 1)} \\ &= \frac{(\gamma - 2\theta)}{(\gamma - \theta)} \cdot \frac{\lambda^{\theta + \sum_{i=1}^n x_i - 1} e^{-\lambda(n+\gamma)} \gamma^\theta}{\prod_{i=1}^n x_i! \Gamma(\theta)} + \frac{\theta}{(\gamma - \theta)} \frac{\lambda^{\theta + \sum_{i=1}^n x_i} e^{-\lambda(n+\gamma)} \gamma^{\theta+1}}{\prod_{i=1}^n x_i! \Gamma(\theta + 1)} \\ &\propto \lambda^{\theta + \sum_{i=1}^n x_i - 1} e^{-\lambda(n+\gamma)} + \lambda^{\theta + \sum_{i=1}^n x_i} e^{-\lambda(n+\gamma)}. \end{aligned}$$

Assim, a densidade obtida corresponde a uma mistura de duas distribuições Gama com os parâmetros $\lambda, \gamma', \theta'$, onde $\gamma' = n + \gamma$ e $\theta' = \theta + \sum_{i=1}^n x_i$, o que caracteriza a distribuição GLD2.

Assim como no caso da GLD1, a construção de um intervalo de credibilidade para λ baseia-se na distribuição posteriori, que neste caso também resulta em uma mistura de distribuições Gama.

Seja $f_{\text{Gama}(a,b)}(\lambda)$ a densidade da distribuição Gama com parâmetros a (forma) e b (taxa), definida por

$$f_{\text{Gama}(a,b)}(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda}, \quad \lambda > 0.$$

No contexto em que a distribuição a priori de λ é dada pela GLD2, a distribuição a posteriori de λ , denotada por $p(\lambda | \mathbf{x})$, assume a forma de uma mistura de duas densidades Gama

$$p(\lambda | \mathbf{x}) = v_1 \cdot f_{\text{Gama}(\theta+S, n+\gamma)}(\lambda) + v_2 \cdot f_{\text{Gama}(\tau+S, n+\phi)}(\lambda),$$

em que os pesos da mistura são agora definidos por

$$v_1 = \frac{1}{1 + \theta}, \quad v_2 = 1 - v_1.$$

Como a distribuição posteriori continua sendo uma mistura de duas Gamas, os quantis necessários para definir o intervalo de credibilidade não podem ser obtidos diretamente. Assim, o quantil de ordem α é determinado resolvendo:

$$Q_\alpha = \inf \left\{ \lambda : \int_0^\lambda p(\lambda' | \mathbf{x}) d\lambda' \geq \alpha \right\}.$$

Dado que não há uma solução analítica fechada, esse cálculo deve ser feito numericamente. Tal como foi feito para a GLD1, uma abordagem prática e eficiente para determinar o intervalo de credibilidade é via amostragem. O procedimento é semelhante,

1. Fixe o número de repetições N , os pesos da mistura v_1 e $v_2 = 1 - v_1$, e os parâmetros das distribuições: θ , τ , S , n , γ e ϕ .
2. Para $i = 1, 2, \dots, N$, repita:
 - a) Gere um número aleatório u de uma distribuição uniforme no intervalo $(0, 1)$.
 - b) Se $u < v_1$, então:
 - Amostre $\lambda_i \sim \text{Gama}(\theta + S, n + \gamma)$.
 - c) Caso contrário:
 - Amostre $\lambda_i \sim \text{Gama}(\tau + S, n + \phi)$.
3. Após obter as N amostras, ordene os valores gerados: $\lambda_{(1)} \leq \lambda_{(2)} \leq \dots \leq \lambda_{(N)}$.
4. Estime os quantis da distribuição a posteriori:
 - $Q_{2.5\%} = \lambda_{(\lfloor 0.025N \rfloor)}$,
 - $Q_{97.5\%} = \lambda_{(\lfloor 0.975N \rfloor)}$.
5. O intervalo de credibilidade de 95% para λ é:

$$[Q_{2.5\%}, Q_{97.5\%}], \quad \text{tal que} \quad P(Q_{2.5\%} \leq \lambda \leq Q_{97.5\%} | \mathbf{x}) = 0,95.$$

Considere a densidade da GLD2,

$$f(\lambda | \theta, \gamma) = \frac{\gamma^\theta}{(\gamma - \theta)\Gamma(\theta)} \lambda^{\theta-1} (\gamma + \gamma\lambda - 2\theta) e^{-\gamma\lambda},$$

o modelo de Poisson para uma única observação x , tem a seguinte distribuição preditiva a priori,

$$\begin{aligned}
f(x) &= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \int_0^\infty \lambda^{x+\theta-1} e^{-(1+\gamma)\lambda} (\gamma + \gamma\lambda - 2\theta) d\lambda \\
&= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \left[\gamma \int_0^\infty \lambda^{x+\theta-1} e^{-(1+\gamma)\lambda} d\lambda + \gamma \int_0^\infty \lambda^{x+\theta} e^{-(1+\gamma)\lambda} d\lambda \right. \\
&\quad \left. - 2\theta \int_0^\infty \lambda^{x+\theta-1} e^{-(1+\gamma)\lambda} d\lambda \right] \\
&= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \left[(\gamma - 2\theta) \int_0^\infty \lambda^{x+\theta-1} e^{-(1+\gamma)\lambda} d\lambda + \gamma \int_0^\infty \lambda^{x+\theta} e^{-(1+\gamma)\lambda} d\lambda \right] \\
&= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \left[(\gamma - 2\theta) \frac{\Gamma(x + \theta)}{(1 + \gamma)^{x+\theta}} + \gamma \frac{\Gamma(x + \theta + 1)}{(1 + \gamma)^{x+\theta+1}} \right] \\
&= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \left[\frac{(\gamma - 2\theta)\Gamma(x + \theta)}{(1 + \gamma)^{x+\theta}} + \frac{\gamma(x + \theta)\Gamma(x + \theta)}{(1 + \gamma)^{x+\theta+1}} \right] \\
&= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \cdot \frac{\Gamma(x + \theta)}{(1 + \gamma)^{x+\theta}} \left[(\gamma - 2\theta) + \frac{\gamma(x + \theta)}{1 + \gamma} \right] \\
&= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \cdot \frac{\Gamma(x + \theta)}{(1 + \gamma)^{x+\theta}} \left[\frac{(\gamma - 2\theta)(1 + \gamma) + \gamma(x + \theta)}{1 + \gamma} \right] \\
&= \frac{\gamma^\theta}{x!(\gamma - \theta)\Gamma(\theta)} \cdot \frac{\Gamma(x + \theta)}{(1 + \gamma)^{x+\theta}} \left[\frac{\gamma(1 + \gamma - 2\theta + x)}{1 + \gamma} \right] \\
&= \frac{\gamma^{\theta+1}}{x!(\gamma - \theta)\Gamma(\theta)} \cdot \frac{\Gamma(x + \theta)}{(1 + \gamma)^{x+\theta}} \cdot \frac{1 + \gamma - 2\theta + x}{1 + \gamma} \\
&= \frac{\gamma^{\theta+1}}{(\gamma - \theta)(1 + \gamma)^{x+\theta}} \cdot \frac{\Gamma(x + \theta)}{x!\Gamma(\theta)} \cdot (1 + \gamma - 2\theta + x) \\
&= \frac{\gamma}{(\gamma - \theta)} \cdot \frac{\Gamma(x + \theta)}{x!\Gamma(\theta)} \cdot \left(\frac{\gamma - 2\theta}{1 + \gamma} \right)^\theta \cdot \left(\frac{\gamma^\theta}{\gamma - \theta} \right)^x + \frac{\theta}{(\gamma - \theta)} \cdot \frac{\Gamma(x + \theta)}{x!\Gamma(\theta)} \\
&\quad \cdot \left(\frac{\gamma}{1 + \gamma} \right)^{\theta+1} \left(\frac{x + \theta}{\gamma - \theta} \right)^x
\end{aligned}$$

também para $x = 0, 1, 2, \dots$, obtemos uma mistura de duas distribuições binomiais negativas. Os pesos da mistura são $p_1 = \frac{\gamma}{\gamma - \theta}$ e $p_2 = \frac{\theta}{\gamma - \theta}$. Os parâmetros das distribuições binomiais negativas associadas são $(\theta, \frac{\gamma^\theta}{\gamma - \theta})$ para o primeiro componente e $(\theta, \frac{x + \theta}{\gamma - \theta})$ para o segundo.

4.3 Distribuição Lindley generalizada 3

Considere a seguinte densidade da GLD3 como priori para o parâmetro λ da distribuição Poisson. Também mudaremos a variável de interesse para λ , conforme indicado pela densidade da Distribuição GLD3,

$$\begin{aligned}
f(\lambda|\alpha, \theta, \gamma) &= \frac{\theta}{\gamma + \theta} \cdot \frac{\theta^\alpha \lambda^{\alpha-1} e^{-\theta\lambda}}{\Gamma(\alpha)} + \frac{\gamma}{\gamma + \theta} \cdot \frac{\theta^{\alpha+1} \lambda^\alpha e^{-\theta\lambda}}{\Gamma(\alpha + 1)} \\
&\propto \lambda^{\alpha-1} e^{-\theta\lambda} + \lambda^\alpha e^{-\theta\lambda}.
\end{aligned}$$

Dada a verossimilhança da Poisson Equação (2.3), obtemos

$$\begin{aligned}
f(\lambda|x_1, \dots, x_n) &= \frac{\theta}{\gamma + \theta} \cdot \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \cdot \frac{\theta^\alpha \lambda^{\alpha-1} e^{-\theta\lambda}}{\Gamma(\alpha)} + \frac{\gamma}{\gamma + \theta} \cdot \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda}}{\prod_{i=1}^n x_i!} \cdot \frac{\theta^{\alpha+1} \lambda^\alpha e^{-\theta\lambda}}{\Gamma(\alpha + 1)}, \\
&= \frac{\theta}{\gamma + \theta} \cdot \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \theta^\alpha \lambda^{\alpha-1} e^{-\theta\lambda}}{\prod_{i=1}^n x_i! \Gamma(\alpha)} + \frac{\gamma}{\gamma + \theta} \cdot \frac{\lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \theta^{\alpha+1} \lambda^\alpha e^{-\theta\lambda}}{\prod_{i=1}^n x_i! \Gamma(\alpha + 1)}, \\
&= \frac{\theta}{\gamma + \theta} \cdot \frac{\lambda^{\sum_{i=1}^n x_i + \alpha - 1} e^{-\lambda(n+\theta)} \theta^\alpha}{\prod_{i=1}^n x_i! \Gamma(\alpha)} + \frac{\gamma}{\gamma + \theta} \cdot \frac{\lambda^{\sum_{i=1}^n x_i + \alpha} e^{-\lambda(n+\theta)} \theta^{\alpha+1}}{\prod_{i=1}^n x_i! \Gamma(\alpha + 1)}, \\
\propto & \lambda^{\sum_{i=1}^n x_i + \alpha - 1} e^{-\lambda(n+\theta)} + \lambda^{\sum_{i=1}^n x_i + \alpha} e^{-\lambda(n+\theta)}.
\end{aligned}$$

Portanto, a densidade obtida pode ser interpretada como uma mistura de duas distribuições Gama, caracterizando uma distribuição GLD3 com os parâmetros $\alpha' = \alpha + \sum_{i=1}^n x_i$ e $\theta' = n + \theta$.

Assim como já foi feito para as prioris GLD1 e GLD2, a construção do intervalo de credibilidade para λ no contexto da GLD3 segue a mesma abordagem bayesiana, utilizando os quantis da distribuição a posteriori. Ao se adotar a priori GLD3, a distribuição a posteriori de λ continua sendo uma mistura de distribuições Gama, cuja forma específica é dada por

$$p(\lambda | \mathbf{x}) = u_1 \cdot f_{\text{Gama}}(\lambda; \theta + S, n + \theta) + u_2 \cdot f_{\text{Gama}}(\lambda; S + \alpha + 1, n + \theta),$$

em que $f_{\text{Gama}}(\lambda; a, b)$ representa a densidade da distribuição Gama com parâmetro de forma a e parâmetro de taxa b .

Os pesos da mistura são definidos como

$$u_1 = \frac{\theta}{\gamma + \theta} \quad u_2 = 1 - u_1.$$

Como a distribuição posteriori ainda é uma mistura de Gamas, o intervalo de credibilidade é obtido a partir dos quantis da distribuição, resolvendo:

$$Q_\alpha = \inf \left\{ \lambda : \int_0^\lambda p(\lambda' | \mathbf{x}) d\lambda' \geq \alpha \right\}.$$

Dado que não há uma solução analítica fechada, essa integral deve ser resolvida numericamente. Assim como fizemos para GLD1 e GLD2, a maneira mais prática de obter o intervalo de credibilidade é por amostragem da posteriori.

Para $i = 1, 2, \dots, N$, repita:

1. Gere um número aleatório u de uma distribuição uniforme no intervalo $(0, 1)$.
2. Se $u < u_1$, então:
 - Amostre $\lambda_i \sim \text{Gama}(\theta + S, n + \theta)$.

3. Caso contrário:

- Amostre $\lambda_i \sim \text{Gama}(S + \alpha + 1, n + \theta)$.
- Após obter as N amostras, ordene os valores gerados: $\lambda_{(1)} \leq \lambda_{(2)} \leq \dots \leq \lambda_{(N)}$.
- Estime os quantis da distribuição a posteriori:
 - $Q_{2,5\%} = \lambda_{(\lfloor 0.025N \rfloor)}$,
 - $Q_{97,5\%} = \lambda_{(\lfloor 0.975N \rfloor)}$.
 - O intervalo de credibilidade de 95% para λ é

$$[Q_{2,5\%}, Q_{97,5\%}], \quad \text{tal que} \quad P(Q_{2,5\%} \leq \lambda \leq Q_{97,5\%} \mid \mathbf{x}) \approx 0,95.$$

Com famílias conjugadas, podemos utilizar a forma conhecida das densidades a priori e posteriori para encontrar a distribuição marginal,

$$\begin{aligned} f(x) &= \int_0^\infty \frac{\lambda^x e^{-\lambda}}{x!} \cdot \frac{\theta^2 (\theta \lambda)^{\alpha-1} (\alpha + \gamma \lambda) e^{-\theta \lambda}}{(\gamma + \theta) \Gamma(\alpha + 1)} d\lambda \\ &= \frac{\theta^{\alpha+1}}{x! (\gamma + \theta) \Gamma(\alpha + 1)} \int_0^\infty \lambda^x e^{-\lambda} \lambda^{\alpha-1} (\alpha + \gamma \lambda) e^{-\theta \lambda} d\lambda \\ &= \frac{\theta^{\alpha+1}}{x! (\gamma + \theta) \Gamma(\alpha + 1)} \int_0^\infty \lambda^{x+\alpha-1} (\alpha + \gamma \lambda) e^{-\lambda(1+\theta)} d\lambda \\ &= \frac{\theta^{\alpha+1}}{x! (\gamma + \theta) \Gamma(\alpha + 1)} + \int_0^\infty \gamma \lambda^{x+\alpha} e^{-\lambda(1+\theta)} d\lambda + \int_0^\infty \alpha \lambda^{x+\alpha-1} e^{-\lambda(1+\theta)} d\lambda \\ &= \frac{\theta^{\alpha+1}}{x! (\gamma + \theta) \Gamma(\alpha + 1)} \left[\alpha \int_0^\infty \lambda^{x+\alpha-1} e^{-\lambda(1+\theta)} d\lambda + \gamma \int_0^\infty \lambda^{x+\alpha} e^{-\lambda(1+\theta)} d\lambda \right] \\ &= \frac{\theta^{\alpha+1}}{x! (\gamma + \theta) \Gamma(\alpha + 1)} \left[\frac{\alpha \Gamma(x + \alpha)}{(1 + \theta)^{x+\alpha}} + \frac{\gamma \Gamma(x + \alpha + 1)}{(x + \alpha)^{x+\alpha+1}} \right] \\ &= \frac{\theta}{\gamma + \theta} \cdot \frac{\Gamma(x + \alpha)}{x! \Gamma(\alpha)} \left(\frac{\theta}{1 + \theta} \right)^\alpha \left(\frac{1}{1 + \theta} \right)^x + \frac{\gamma}{\gamma + \theta} \cdot \frac{\Gamma(x + \alpha + 1)}{x! \Gamma(\alpha + 1)} \left(\frac{\theta}{1 + \theta} \right)^{\alpha+1} \left(\frac{1}{1 + \theta} \right)^x \end{aligned}$$

para $x = 0, 1, 2, \dots$, obtemos uma mistura de duas distribuições binomiais negativas. Os pesos da mistura são $p_1 = \frac{\theta}{\gamma + \theta}$ e $p_2 = \frac{\gamma}{\gamma + \theta}$. Os parâmetros das distribuições binomiais negativas associadas são $(\alpha, \frac{\theta}{1+\theta})$ para o primeiro componente e $(\alpha + 1, \frac{\theta}{1+\theta})$ para o segundo.

5 Aplicação

Neste capítulo, aplicamos nossa metodologia de estudo a um conjunto de dados reais: o número de ataques de bombas voadoras no sul de Londres durante a Segunda Guerra Mundial (FELLER, 1957). Durante esse período, o intenso bombardeio afetou consideravelmente a população e a infraestrutura da cidade. A escolha desse conjunto de dados para a aplicação da metodologia é, portanto, motivada não apenas pela sua relevância histórica, mas também pelas propriedades estatísticas que ele apresenta, como a equidispersão, ou seja, a variância é aproximadamente igual à média, o que os torna adequados para a distribuição de Poisson.

5.1 O número de ataques de bombas em Londres durante a Segunda Guerra Mundial

Durante a Segunda Guerra Mundial, Londres foi alvo de diversos ataques de bombas voadoras, particularmente na fase conhecida como "Blitz" (BROOKS, 2011), que inclui o "Sábado Negro" (GEOGHEGAN, 2010). Esses ataques causaram destruição em larga escala e tiveram um impacto duradouro na história da cidade. Londres foi bombardeada durante 57 noites consecutivas, de 7 de setembro a 2 de novembro (BROOKS, 2011). A cidade foi dividida em 576 pequenas áreas de um quarto de quilômetro quadrado cada, e o número de áreas atingidas exatamente z vezes foi registrado. Este conjunto de dados históricos não apenas oferece um vislumbre dos desafios enfrentados pelos londrinos, mas também serve como um excelente exemplo para a aplicação de técnicas estatísticas modernas. Esses dados estão disponíveis na Tabela (1).

Tabela 1 – Ataques de bombas voadoras em Londres durante a Segunda Guerra Mundial.

Ataques	0	1	2	3	4	5+
Frequência absoluta	229	211	93	35	7	1

Utilizando o software R (TEAM, 2017), realizamos uma análise descritiva inicial dos dados. A média amostral foi calculada como 0,9323, enquanto a variância amostral foi 0,9711. A razão entre a variância e a média da amostra, conhecida como índice de dispersão empírica, foi encontrada como 1,0416. Este valor, próximo de 1, sugere que os dados são equidispersos, ou seja, a variância é aproximadamente igual à média. Esta característica é um indicativo de que a distribuição Poisson é um modelo apropriado para os dados em questão, já que a distribuição Poisson assume que a média e a variância são iguais.

No contexto da análise bayesiana do parâmetro λ , consideramos as distribuições descritas neste trabalho. Cada uma dessas distribuições é empregada como priori conjugada para λ , permitindo que a densidade posteriori também siga uma forma conhecida, facilitando os cálculos da média e do desvio padrão a posteriori. Aqui, abordamos cinco tipos de distribuições a priori: a distribuição Lindley ponderada (WL), GLD1, GLD2, GLD3 e Gama.

A distribuição Lindley ponderada (WL) foi introduzida por [Ghitany et al. \(2011\)](#) e posteriormente explorada por [Bourguignon \(2018\)](#), que analisou sua aplicação como uma priori conjugada. Nesse contexto, o autor destacou a flexibilidade e a utilidade da WL em relação à distribuição Gama na inferência bayesiana, especialmente em análises de sobrevivência e estudos de confiabilidade. Essa distribuição representa uma generalização da distribuição Lindley original, proposta por [Lindley \(1958\)](#).

Para a distribuição WL, a escolha dos valores $\alpha = 0.001$ e $\beta = 0.001$ representa uma abordagem não informativa. Uma abordagem não informativa para os parâmetros de um modelo estatístico é uma maneira de modelar as incertezas sobre os parâmetros sem assumir nenhuma informação prévia forte sobre seus valores. Dessa forma, a inferência é conduzida essencialmente pelos dados observados. Após a atualização bayesiana, a posteriori de λ continua seguindo uma distribuição WL, mas com parâmetros ajustados de acordo com as observações, refletindo a informação incorporada a partir dos dados. Essa abordagem permite calcular diretamente a média e o desvio padrão da distribuição a posteriori, resultando em uma média a posteriori de $\hat{\lambda} = 0.9286$.

Para a GLD1, o parâmetro λ é modelado como seguindo uma distribuição GLD1 com parâmetros não informativos, definidos como $\gamma = 0.001$, $\phi = 0.001$, $\eta = 0.001$, $\tau = 0.001$ e $\theta = 0.001$. A escolha desses valores reflete a ausência de conhecimento prévio significativo sobre λ , permitindo que a inferência seja guiada principalmente pelos dados observados, de maneira análoga à abordagem utilizada para a distribuição WL. Após a atualização bayesiana, a posteriori de λ passa a seguir uma distribuição GLD1 ajustada de acordo com as observações. Com essa configuração, a média posteriori de λ é calculada como $\hat{\lambda} = 0.9370$.

A GLD2 é utilizada como priori conjugada com valores não informativos de $\gamma = 0.001$ e $\theta = 0.001$. Para esta distribuição, a densidade posterior também assume a forma de uma distribuição GLD2, permitindo o cálculo direto da média posterior e do desvio padrão. A média posterior obtida com a GLD2 é $\hat{\lambda} = 0.9301$.

Por fim, a GLD3 é considerada com uma priori não informativa para os parâmetros $\alpha = 0.001$, $\theta = 0.001$ e $\gamma = 0.001$. Neste caso, a densidade posterior também segue uma distribuição GLD3 ajustada, e a média posterior calculada é $\hat{\lambda} = 0.9296$.

Na Tabela 2, apresentada acima, é realizada uma comparação entre os modelos

Tabela 2 – Comparação das Estimativas dos Modelos GLD1, GLD2, GLD3, Gama e WL: Parâmetros Teóricos, Amostrais, EMV, Erro Padrão, AIC, BIC e Intervalos de Credibilidade

Modelo		$\hat{\lambda}$	$ep(\hat{\lambda})$	AIC	BIC	IC
	EMV	0.9322	0.0402	1459.42	1463.78	[0.8577, 1.0132]
GLD1	Teórico	0.9305	0.0401	1459.42	1463.78	[0.8509; 1.0077]
	Amostral	0.9370	0.0416	1459.46	1463.82	[0.8531; 1.0097]
GLD2	Teórico	0.9514	0.0416	1459.73	1464.09	[0.8561; 1.0090]
	Amostral	0.9301	0.0403	1459.42	1463.78	[0.8515; 1.0156]
GLD3	Teórico	0.9288	0.0401	1459.46	1463.78	[0.8521; 1.0136]
	Amostral	0.9296	0.0408	1459.42	1463.78	[0.8461; 1.0106]
Gama	Teórico	0.9288	0.0401	1459.42	1463.78	[0.8464; 1.0096]
	Amostral	0.9279	0.0402	1459.42	1463.78	[0.8554; 1.0107]
WL	Teórico	0.9296	0.0401	1459.42	1463.78	[0.8592; 1.0083]
	Amostral	0.9286	0.0404	1459.42	1463.78	[0.8538; 1.0038]

estatísticos propostos: GLD1, GLD2, GLD3, Gama e WL. A tabela inclui diversas medidas de qualidade de ajuste, como o valor estimado de $\hat{\lambda}$, o erro padrão do estimador $ep(\hat{\lambda})$, o Critério de Informação de Akaike (AIC), o Critério de Informação Bayesiano (BIC) e o Intervalo de Credibilidade (IC).

O AIC (Akaike Information Criterion) foi desenvolvido em 1973 pelo estatístico japonês Hirotugu Akaike, com o objetivo de selecionar modelos que melhor expliquem os dados com o menor número possível de parâmetros. Ele é baseado na teoria da informação e busca minimizar a divergência entre o modelo estimado e o modelo verdadeiro. Sua fórmula penaliza a complexidade do modelo, evitando sobreajuste.

Já o BIC (Bayesian Information Criterion), proposto em 1978 pelo matemático Gideon E. Schwarz, parte de uma abordagem bayesiana. Assim como o AIC, penaliza modelos com muitos parâmetros, porém de forma mais intensa, pois leva em conta o tamanho da amostra. Em contextos com grandes volumes de dados, o BIC tende a favorecer modelos mais parcimoniosos.

O Intervalo de Credibilidade (IC), por sua vez, surge no contexto da estatística bayesiana, fundamentada nos trabalhos do matemático e teólogo inglês Thomas Bayes (1702–1761). Ao contrário dos intervalos de confiança frequentistas, os intervalos de credibilidade permitem interpretações probabilísticas diretas: por exemplo, um intervalo de 95% indica que, dados os dados observados e o modelo especificado, há 95% de probabilidade de que o parâmetro de interesse esteja dentro do intervalo.

O intervalo de credibilidade teórico de 95% para o parâmetro λ foi obtido por meio de uma abordagem Bayesiana, na qual a distribuição a posteriori de λ é modelada como uma

mistura de duas distribuições Gama. A partir dos dados observados, os hiperparâmetros da distribuição a priori foram atualizados, resultando em novos valores para os parâmetros. Com esses parâmetros, foram geradas amostras da distribuição a posteriori utilizando uma função que simula valores com base na estrutura da mistura e nos pesos definidos pelos próprios parâmetros. Em seguida, os quantis de ordem 2,5% e 97,5% foram calculados a partir das amostras, fornecendo o intervalo de credibilidade de 95%. Esse procedimento corresponde a uma aproximação de Monte Carlo, frequentemente empregada na inferência Bayesiana para a obtenção de estimativas e intervalos quando a forma analítica da posteriori não é diretamente manipulável.

Para cada modelo, foram calculadas duas estimativas de $\hat{\lambda}$ no software R: o valor teórico, obtido a partir da expressão analítica da esperança da distribuição posterior de λ , considerando os parâmetros ajustados; e o valor amostral, estimado por meio de simulação, utilizando amostras geradas da distribuição a posteriori de λ . Essa abordagem permitiu avaliar a variabilidade do estimador na prática. Além disso, os erros padrão foram calculados para ambas as abordagens, possibilitando uma análise comparativa da precisão dos modelos.

O erro padrão $ep(\hat{\lambda})$ indica a precisão ou incerteza da estimativa de $\hat{\lambda}$. De maneira geral, menores valores de erro padrão sugerem maior precisão da estimativa. Ao compararmos os erros padrão teóricos e amostrais das estimativas de $\hat{\lambda}$, observa-se que os valores teóricos, calculados a partir das expressões analíticas da variância da distribuição a posteriori, tendem a ser ligeiramente menores do que os erros padrão obtidos via simulação (amostrais). Essa diferença é esperada, pois o erro padrão teórico parte de pressupostos matemáticos ideais, enquanto o erro padrão amostral reflete a variabilidade real das amostras geradas, incorporando flutuações inerentes ao processo de simulação. Já os critérios AIC e BIC foram utilizados para avaliar e comparar a qualidade dos modelos propostos aqui, os menores valores desses critérios indicam um melhor ajuste do modelo aos dados.

É importante notar que os valores estimados para λ são bastante próximos entre si, tanto para as estimativas teóricas quanto para as amostrais, em todos os modelos considerados. Isso sugere uma boa concordância entre as abordagens teórica e empírica. No entanto, pequenas variações nos erros padrão $ep(\hat{\lambda})$ entre as estimativas teóricas e amostrais indicam diferenças sutis na precisão dos estimadores, dependendo do modelo utilizado.

Quanto aos critérios AIC e BIC, observa-se que os resultados são muito semelhantes entre os modelos GLD1, GLD2, GLD3, Gama e WL.

6 Conclusão

No presente trabalho, propusemos a utilização de misturas da distribuição Gama como alternativas a priori conjugadas para o modelo Poisson, explorando suas propriedades e aplicando-as em um estudo de caso com dados reais. A pesquisa mostrou que as distribuições propostas, baseadas em misturas de distribuição Gama, não só reproduzem os comportamentos observados com a distribuição Gama tradicional, como também oferecem maior flexibilidade, assumindo, por exemplo, formas bimodais.

A aplicação prática no estudo de caso sobre os ataques de bombas voadoras em Londres durante a Segunda Guerra Mundial demonstrou a eficácia das distribuições propostas em capturar padrões nos dados. A natureza dispersa dos eventos observados indicou que o modelo Poisson era adequado para representar o comportamento dos dados.

As contribuições desta pesquisa são significativas tanto do ponto de vista teórico quanto prático. Teoricamente, ampliamos o entendimento sobre o uso de misturas de distribuições Gama como a priori em modelos bayesianos, proporcionando novas ferramentas para modelar a incerteza em parâmetros do modelo Poisson.

Para avanços futuros, sugere-se a investigação de novas combinações de distribuições a priori para criar modelos ainda mais robustos, no sentido de aumentar a flexibilidade e a capacidade preditiva das inferências feitas. A ideia é testar combinações que possam capturar melhor as características específicas de diferentes conjuntos de dados, de modo a melhorar a precisão dos resultados. Além disso, é fundamental explorar a aplicação dessas técnicas em uma variedade de contextos e conjuntos de dados distintos, de modo a avaliar sua eficácia em diferentes cenários.

A análise de sensibilidade dos parâmetros utilizados na modelagem pode fornecer informações importantes sobre como as variações nos parâmetros a priori afetam as inferências, contribuindo para um entendimento mais profundo da estabilidade do modelo. Também seria interessante explorar outras metodologias bayesianas, como técnicas de amostragem mais avançadas ou modelos hierárquicos, que podem oferecer novas perspectivas e melhorar a robustez das conclusões.

Este trabalho, portanto, não só propõe alternativas para a modelagem de incertezas em modelos Poisson, como também abre novas possibilidades para pesquisas futuras. Tais estudos podem expandir ainda mais a aplicação da inferência bayesiana, ampliando seu uso e influência na estatística.

Referências Bibliográficas

- ABOUAMMOH, A.; ALSHANGITI, A. M.; RAGAB, I. E. A new generalized lindley distribution. *Journal of Statistical Computation and Simulation*, Informa UK Limited, v. 85, n. 18, p. 3662–3678, 2015. Published on 2 Jan. 2015. Disponível em: <<http://dx.doi.org/10.1080/00949655.2014.995101>>. Citado 3 vezes nas páginas 7, 8 e 25.
- AL-SALEH, J. A.; AGARWAL, S. K. Finite mixture of gamma distributions: a conjugate prior. *Computational Statistics Data Analysis*, v. 51, p. 4369–4378, 2007. Citado na página 15.
- AL-SALEH, M.; AGARWAL, S. Finite mixture of gamma distributions with applications. *Journal of Statistical Theory and Applications*, v. 6, n. 1, p. 57–71, 2007. Citado na página 21.
- AMIR-AHMADI, P.; MATTHES, C.; WANG, M. Choosing prior hyperparameters. *ERN: Bayesian Analysis (Topic)*, 2016. Citado na página 22.
- ARNST, M.; SOIZE, C. Identification and sampling of bayesian posteriors of high-dimensional symmetric positive-definite matrices for data-driven updating of computational models. *Computer Methods in Applied Mechanics and Engineering*, 2019. Citado na página 22.
- BERGER, J. O. *Statistical Decision Theory and Bayesian Analysis*. [S.l.]: Springer-Verlag, 1985. Citado na página 18.
- BERGOT, M. *Finite Mixture Distributions*. Tese (Doutorado) — Université Bordeaux, Ecole d'Ingénieurs en Modélisation Mathématique et Mécanique, 2006. Citado na página 15.
- BERNARDO, J. M.; SMITH, A. F. *Bayesian Theory*. [S.l.]: Wiley, 2000. Citado na página 18.
- BOCHKINA, N.; ROUSSEAU, J. Adaptive density estimation based on a mixture of gammas. *arXiv: Statistics Theory*, 2016. Citado na página 21.
- BOLSTAD, W. M. *Understanding Computational Bayesian Statistics*. [S.l.]: Wiley, 2007. Citado na página 18.
- BOUCHER, J. P.; DENUIT, M.; GUILLÉN, M. Number of accidents or number of claims? an approach with zero-inflated poisson models for panel data. *Journal of Risk and Insurance*, v. 76, n. 4, p. 821–846, 2009. Citado na página 15.
- BOURGUIGNON, M. An alternative conjugate prior distribution for positive parameters. *Annals of Data Science*, Springer Science and Business Media LLC, v. 6, n. 2, p. 237–243, 2018. Published on 16 Aug. 2018. Disponível em: <<http://dx.doi.org/10.1007/s40745-018-0174-z>>. Citado na página 44.
- BOX, G. E. P.; TIAO, G. C. *Bayesian Inference in Statistical Analysis*. New York: J. Wiley Interscience, 1992. Citado na página 22. Citado 2 vezes nas páginas 14 e 15.

- BROEMELING, L. D. Bayesian biostatistics and diagnostic medicine. *CRC Press*, 2017. Discusses Bayesian inference applied to medical diagnosis and clinical trials. Citado na página 18.
- BROOKS, A. *Londres em guerra: relíquias da frente doméstica das Guerras Mundiais*. Barnsley, South Yorkshire: Wharncliffe Books, 2011. Pág. 25. ISBN 978-1-84563-139-0. Citado na página 43.
- DALAL, S. R.; HALL, W. J. Approximating priors by mixtures of natural conjugate priors. *Journal of the Royal Statistical Society: Series B*, v. 45, p. 278–286, 1983. Citado na página 21.
- EVERITT, B. S.; HAND, D. J. Finite mixture distributions. 1981. Citado na página 20.
- FALCÃO, A. J. S. et al. Aplicação de métodos reml e bayesiano via amostrador de gibbs na estimação de componentes de variância para produção de leite no estado do paran . In: SBMA. *Anais do Simp sio da Sociedade Brasileira de Melhoramento Animal*. Pirassununga, 2004. Citado na página 18.
- FELLER, W. *Uma introdu o   teoria da probabilidade e suas aplica es*. Nova Jersey: Wiley, 1957. v. 1. Citado na p gina 43.
- FRUHWIRTH-SCHNATTER, S. *Finite Mixture and Markov Switching Models*. New York: Springer, 2006. ISBN 0387329099. Citado na p gina 20.
- GELMAN, A. et al. Bayesian data analysis in genetics research. *Genetics Research*, Oxford University Press, v. 43, n. 3, p. 179–206, 2004. Citado 2 vezes nas p ginas 18 e 19.
- GELMAN, A. et al. *Bayesian Data Analysis*. [S.l.]: CRC Press, 2013. Citado na p gina 15.
- GELMAN, A. et al. *Bayesian Data Analysis*. 3rd. ed. [S.l.]: CRC Press, 2013. ISBN 978-1439840955. Citado 2 vezes nas p ginas 15 e 22.
- GELMAN, A. et al. *Bayesian Data Analysis*. [S.l.]: Chapman and Hall/CRC, 2013. Citado na p gina 18.
- GEOGHEGAN, T. O blitz realmente unificou a gr -bretanha. *BBC News*, setembro 2010. Recuperado em 2 de fevereiro de 2023. Citado na p gina 43.
- GHITANY, M. E. et al. A two-parameter weighted lindley distribution and its applications to survival data. *Mathematics and Computers in Simulation*, v. 81, p. 1190–1201, 2011. Citado na p gina 44.
- GRIFFIN, J.; BROWN, P. Inference with normal-gamma prior distributions in regression problems. *Bayesian Analysis*, v. 5, p. 171–188, 2010. Citado na p gina 22.
- JAYNES, E. T. *Probability Theory: The Logic of Science*. [S.l.]: Cambridge University Press, 2003. Citado na p gina 19.
- LAIRD, N. M.; WARE, J. H. Random-effects models for longitudinal data. *Biometrics*, JSTOR, v. 38, n. 4, p. 963–974, 1982. Citado na p gina 15.
- LAWSON, A. B.; BROWNE, W. J.; RODEIRO, C. L. V. *Disease mapping with WinBUGS and MLwiN*. [S.l.]: John Wiley Sons, 2003. ISBN 978-0-471-48900-4. Citado na p gina 15.

- LEE, P. M. *Bayesian Statistics: An Introduction*. 4th. ed. [S.l.]: Wiley, 2012. Citado na página 18.
- LINDLEY, D. Fiducial distributions and bayes theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, v. 9, n. 1, p. 102–107, 1958. Citado na página 44.
- LINDSAY, B. G. *Mixture Models: Theory, Geometry and Applications*. Hayward, California: Institute of Mathematical Statistics, 1995. v. 5. (NSF-CBMS Regional Conference Series in Probability and Statistics, v. 5). ISBN 0940600313. Citado na página 20.
- LUCAS, A.; SCHWAAB, B.; ZHANG, X. Bayesian econometrics in macroeconomic forecasting. *Journal of Applied Econometrics*, Wiley Online Library, v. 30, n. 1, p. 46–71, 2015. Citado na página 18.
- MAYROSE, I.; FRIEDMAN, N.; PUPKO, T. A gamma mixture model better accounts for amongsite rate heterogeneity. *Bioinformatics*, v. 21, n. suppl_2, p. ii151–ii158, 2005. Citado na página 21.
- MCHUGH, M. Descriptive statistics. part i: level of measurement. *Journal of Spec Pediatric Nursing*, v. 8, p. 35–47, 2003. Citado na página 14.
- MCLACHLAN, G. J.; BASFORD, K. E. *Mixture Models: Inference and Applications to Clustering*. New York: Marcel Dekker Inc., 1988. ISBN 0824777460. Citado na página 20.
- MCNEISH, D. M. On using bayesian methods to address small sample problems. *Structural Equation Modeling: A Multidisciplinary Journal*, v. 23, p. 750 – 773, 2016. Citado na página 14.
- MCNICHOLAS, P. D. *Mixture Model-Based Classification*. Boca Raton: Chapman and Hall/CRC, 2017. ISBN 9781498725855. Citado na página 20.
- MENGERSEN, K. L.; ROBERT, C. P.; TITTERINGTON, D. M. *Mixtures: Estimation and Applications*. Chichester, UK: John Wiley & Sons, 2011. ISBN 9780470743820. Citado na página 20.
- MESSAM, L. L. M. et al. Frequentist and bayesian approaches to prevalence estimation using examples from johnes disease. *Animal Health Research Reviews*, v. 9, p. 1 – 23, 2008. Citado na página 14.
- MURPHY, K. P. *Machine Learning: A Probabilistic Perspective*. [S.l.]: MIT Press, 2012. Citado na página 18.
- MURTEIRA, B. et al. *Estatística Bayesiana*. 2. ed. [S.l.]: Fundação Calouste Gulbenkian, 2018. Citado na página 15.
- NIEKERK, J. V.; BEKKER, A.; ARASHI, M. A gamma-mixture class of distributions with bayesian application. *Communications in Statistics Part B: Simulation and Computation*, v. 46, n. 10, p. 8152–8165, 2017. ISSN 0361-0918. Citado na página 21.
- PATIL, G.; RAO, C. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, v. 34, n. 2, p. 179–189, 1978. Citado na página 20.

PÉREZ, M. R. e JACINTO MARTÍN e C. Inferência sobre famílias exponenciais com mistura de distribuições anteriores. *Comput. Stat. Data Anal.*, v. 53, p. 3271–3280, 2009. Citado na página 21.

RAIFFA, H. *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. [S.l.]: Addison-Wesley, 1968. Citado na página 18.

RAMOS, P. L.; LOUZADA, F.; MOALA, F. A. A two-parameter distribution with increasing and bathtub hazard rate. *Journal of Data Science*, School of Statistics, Renmin University of China, v. 18, n. 4, p. 813–827, 2021. Disponível em: <[http://dx.doi.org/10.6339/jds.202010_18\(4\).0014](http://dx.doi.org/10.6339/jds.202010_18(4).0014)>. Citado 4 vezes nas páginas 7, 8, 25 e 28.

RAMOS, P. L.; LOUZADA, F.; MOALA, P. A new class of lindley distributions applied to lifetime data. *Journal of Statistical Computation and Simulation*, v. 86, p. 3065–3075, 2016. Citado na página 15.

RIGDON, S. E.; BASU, A. P. *Statistical Methods for the Reliability of Repairable Systems*. [S.l.]: Wiley-Interscience, 2000. ISBN 978-0-471-18616-3. Citado na página 15.

ROBERT, C. P. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. 2nd. ed. [S.l.]: Springer, 2007. Citado 2 vezes nas páginas 18 e 19.

ROBERT, C. P.; CASELLA, G. *Monte Carlo Statistical Methods*. 2nd. ed. [S.l.]: Springer, 2004. Citado na página 18.

SMID, S. C. et al. Bayesian versus frequentist estimation for structural equation models in small sample contexts: A systematic review. *Structural Equation Modeling: A Multidisciplinary Journal*, v. 27, p. 131 – 161, 2019. Citado na página 14.

TEAM, R. C. *Uma linguagem e ambiente para computação estatística*. Viena, 2017. Citado na página 43.

TWYXCROSS, A.; SHIELDS, L. Statistics made simple. part 1: mean, medians, and mode. *Paediatric Nursing*, v. 16, p. 32, 2004. Citado na página 14.

YOKOO, M. J. I. et al. O uso da estatística bayesiana no melhoramento genético animal: uma breve explicação. *Scientia Agraria Paranaensis*, v. 12, n. 4, p. 247–257, 2013. Citado na página 14.

ZAKERZADEH, H.; DOLATI, A. Generalized lindley distribution. *Journal of Mathematical Extension*, v. 3, p. 13–25, 2009. Citado 4 vezes nas páginas 7, 8, 25 e 31.