



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 Binomiais com priori conjugada baseada em generalizações da distribuição beta

Luana Mayara Lucas Leite

Natal – RN

Março/2025

Luana Mayara Lucas Leite

**Inferência Bayesiana para Modelos Binomiais com
priori conjugada baseada em generalizações da
distribuição beta**

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

Março/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

Leite, Luana Mayara Lucas.

Inferência bayesiana para modelos binomiais com priori conjugada baseada em generalizações da distribuição beta / Luana Mayara Lucas Leite. - 2025.

61 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. Distribuição beta - Dissertação. 3. Modelo binomial - Dissertação. 4. Posteriori - Dissertação. 5. Priori - 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 Binomiais com priori conjugada baseada em generalizações da distribuição beta* apresentada por Luana Mayara Lucas Leite 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
Instituição

Professor

Departamento
Instituição

Professor

Departamento
Instituição

Natal – RN
Março/2025

Agradecimentos

Agradeço, primeiramente, a Deus, por ser a força que me sustenta e pelas oportunidades que me ofertou ao longo da vida.

À minha mãe, Telma, grande responsável por eu chegar até aqui. Ao meu pai, Joailton, por me incentivar e apresentar a Universidade pela primeira vez.

Aos meus irmãos, Amanda e Deyvid, por serem fontes de inspiração, e à pequena Ana Júlia, minha sobrinha, pelo carinho e inúmeros momentos de alegria que proporcionou, mesmo em meio às dificuldades que surgiram.

Ao meu noivo, por todo o apoio e completo incentivo nessa trajetória.

Ao meu orientador, pela disponibilidade e ricos momentos de aprendizado.

À UFRN, por oferecer a estrutura necessária ao desenvolvimento deste trabalho.

*“Persistence is the path to success.”
(Charles Chaplin)*

Resumo

A presente dissertação objetiva propor distribuições alternativas a priori conjugada para o modelo binomial baseadas em generalizações da distribuição beta. Nesse contexto, a metodologia desenvolvida busca estimar o parâmetro de proporção π da distribuição binomial por meio da abordagem bayesiana, utilizando-se como priori generalizações da distribuição beta, de tal forma que a distribuição a posteriori também pertença à mesma classe de distribuição da priori. Além disso, serão estudadas em detalhes as propriedades dessas distribuições, com a realização de simulações a partir da geração de números aleatórios, sendo verificadas as suas vantagens em relação à distribuição beta, visto que em termos de ajuste, busca-se alcançar melhores resultados aos obtidos pela distribuição beta, quando comparados.

Palavras-chave: Inferência Bayesiana. Distribuição Beta. Modelo Binomial. Posteriori. Priori.

Abstract

The present dissertation aims to propose alternative conjugate prior distributions for the binomial model based on generalizations of the beta distribution. In this context, the developed methodology seeks to estimate the proportion parameter π of the binomial distribution using a Bayesian approach, employing generalizations of the beta distribution as priors in such a way that the posterior distribution also belongs to the same class as the prior. Additionally, the properties of these distributions will be studied in detail, with simulations conducted through random number generation. Their advantages over the beta distribution will be assessed, as the aim is to achieve better fitting results than those obtained by the beta distribution when compared.

Keywords: Bayesian Inference. Beta distribution. Binomial Model. Prior. Posterior.

Lista de Figuras

Figura 1 – Densidade da distribuição Kummer beta com parâmetros (a) $a = 3$, $b = 5$, $c = -3$, (b) $a = 3$, $b = 5$, $c = 4$, (c) $a = 4$, $b = 5$, $c = -8$, (d) $a = 1$, $b = 1$, $c = 0$	31
Figura 2 – Densidade da distribuição Kummer beta com parâmetros (a) $a = 0.5$, $b = 1$, $c = 8$, (b) $a = 0.5$, $b = 1$, $c = -8$, (c) $a = 0.9$, $b = 6$, $c = -7$ e (d) $a = 0.9$, $b = 9$, $c = -9$	32
Figura 3 – Densidade da distribuição Kummer beta com parâmetros (a) $a = 0.1$, $b = 2$, $c = -8$, (b) $a = 0.3$, $b = 3$, $c = -8$, (c) $a = 0.4$, $b = 3$, $c = -6$ e (d) $a = 0.7$, $b = 3$, $c = -6$	33
Figura 4 – Densidade da distribuição Hipergeométrica de Gauss com parâmetros (a) $a = 6$, $b = 5$, $c = 4$ e $d = 0.1$, (b) $a = 4$, $b = 5$, $c = 6$ e $d = 0.9$, (c) $a = 2$, $b = 2$, $c = 2$ e $d = 0.5$, (d) $a = 1$, $b = 1$, $c = 0$ e $d = 0.5$	34
Figura 5 – Densidade da distribuição Hipergeométrica de Gauss com parâmetros (a) $a = 1$, $b = 2$, $c = 5$ e $d = -0.4$, (b) $a = 1$, $b = 3$, $c = 5$ e $d = -0.5$, (c) $a = 1.5$, $b = 3$, $c = 5$ e $d = -0.6$, (d) $a = 1$, $b = 3$, $c = 5$ e $d = -0.7$	35
Figura 6 – Histogramas da distribuição Kummer beta com parâmetros $a = 0.1$, $b = 2$ e $c = -8$, para tamanhos de amostra m iguais a: (a) 100, (b) 250, (c) 500, (d) 1000.	40
Figura 7 – Histogramas da distribuição Hipergeométrica de Gauss com parâmetros $a = 2$, $b = 2$, $c = 2$ e $d = 0.5$, para tamanhos de amostra m iguais a: (a) 100, (b) 250, (c) 500, (d) 1000.	41
Figura 8 – Gráficos da posteriori Kummer beta com parâmetros $a = 1.5$, $b = 3.5$ e $c = 3$, para total de peças e número de defeitos iguais a: (a) 30 e 8, (b) 50 e 6, (c) 70 e 12, (d) 90 e 9, respectivamente.	47
Figura 9 – Gráficos da posteriori Hipergeométrica de Gauss com parâmetros $a = 1.5$, $b = 3$, $c = 3.5$, $d = 0.9$, para total de peças e número de defeitos iguais a: (a) 30 e 8, (b) 50 e 6, (c) 70 e 12, (d) 90 e 9, respectivamente.	48
Figura 10 – Gráficos da posteriori Kummer beta com parâmetros $a = 1.5$, $b = 5$ e $c = 5.5$, para total de peças e número de defeitos iguais a: (a) 40 e 2, (b) 60 e 7, (c) 80 e 16, (d) 100 e 7, respectivamente.	49
Figura 11 – Gráficos da posteriori Hipergeométrica de Gauss com parâmetros $a = 1.1$, $b = 2$, $c = 1.8$, $d = 0.4$, para total de peças e número de defeitos iguais a: (a) 40 e 2, (b) 60 e 7, (c) 80 e 16, (d) 100 e 7, respectivamente.	50

Lista de Tabelas

Tabela 1	– Priori conjugada para o modelo binomial baseada em generalizações da distribuição beta, suas respectivas posteriores e alguns casos especiais que, por aproximação, resultam na distribuição beta.	30
Tabela 2	– Resultado da média teórica e amostral para a distribuição Kummer beta com parâmetros $a = 0.9$, $b = 6$, $c = -7$ e a distribuição Hipergeométrica de Gauss com parâmetros $a = 3$, $b = 4$, $c = 2$, $d = 0.5$, para diferentes tamanhos de amostras.	39
Tabela 3	– Resultado da média teórica e amostral para a distribuição Kummer beta com parâmetros $a = 0.7$, $b = 5$, $c = -8$, e a distribuição Hipergeométrica de Gauss com parâmetros $a = 4$, $b = 3$, $c = 1$, $d = -0.9$, para diferentes tamanhos de amostras.	39
Tabela 4	– Resultado para a distribuição Beta com parâmetros $\alpha = 1$, $\beta = 1$, Kummer beta com parâmetros $a = 1.5$, $b = 3.5$, $c = 3$ e a distribuição Hipergeométrica de Gauss com parâmetros $a = 1.5$, $b = 3$, $c = 3.5$, $d = 0.9$, para diferentes tamanhos de amostras.	45
Tabela 5	– Resultado para a distribuição Beta com parâmetros $\alpha = 1$, $\beta = 1$, Kummer beta com parâmetros $a = 1.5$, $b = 5$, $c = 5.5$ e a distribuição Hipergeométrica de Gauss com parâmetros $a = 1.1$, $b = 2$, $c = 1.8$, $d = 0.4$, para diferentes tamanhos de amostras.	45

Lista de abreviaturas e siglas

AIC	Critério de Informação de Akaike
BIC	Critério de Informação de Bayes
CDF	Função de Distribuição Cumulativa
FDP	Função Densidade de Probabilidade
MCMC	Monte Carlo por Cadeias de Markov

Sumário

1	INTRODUÇÃO	13
1.1	Objetivos	14
1.2	Organização do trabalho	15
2	REFERENCIAL TEÓRICO	16
2.1	Conceitos fundamentais na Inferência Bayesiana	16
2.2	Estimação do parâmetro no modelo binomial	18
3	DISTRIBUIÇÕES BETA GENERALIZADAS	22
3.1	Distribuição Kummer beta	22
3.2	Distribuição Hipergeométrica de Gauss	23
3.3	Distribuição Apell beta	24
3.4	Distribuição beta Bessel I	25
3.5	Distribuição beta Bessel II	26
3.6	Distribuição beta Bessel III	26
3.7	Distribuição Beta Trigonométrica Cos I	27
3.8	Distribuição Beta Trigonométrica Sen I	27
3.9	Distribuição Kummer do tipo II	28
3.10	Distribuição Beta Generalizada do tipo II	29
3.11	Distribuição Kummer beta II	29
4	INFERÊNCIA PARA O MODELO BINOMIAL	30
4.1	Geração de números pseudoaleatórios	37
5	APLICAÇÃO COM DADOS SIMULADOS	42
5.1	Critério de Informação de Akaike (AIC)	42
5.2	Critério de Informação Bayesiano (BIC)	42
5.3	Critério de Informação Deviance (DIC)	43
5.4	Simulações	43
5.5	Resultados e discussões	45
5.6	Gráficos da distribuição a posteriori	47
6	CONSIDERAÇÕES FINAIS	51
	REFERÊNCIAS BIBLIOGRÁFICAS	52

APÊNDICES

56

APÊNDICE A – CÁLCULO PARA OBTENÇÃO DA POSTERIORI 57

A.0.1	Priori Kummer beta	57
A.0.2	Priori Hipergeométrica de Gauss	57
A.0.3	Priori Apell beta	58
A.0.4	Priori BB I	58
A.0.5	Priori BB II	58
A.0.6	Priori BB III	59
A.0.7	Priori BT Cos I	59
A.0.8	Priori BT Sen I	59
A.0.9	Priori Kummer do tipo II	60
A.0.10	Priori Beta Generalidaza do tipo II	60
A.0.11	Priori Kummer beta II	61

1 Introdução

Fundamentada nos princípios estabelecidos pelo reverendo matemático britânico Thomas Bayes, no século XVIII, cujas ideias foram postumamente publicadas em “An Essay towards solving a Problem in the Doctrine of Chances”, originalmente em [Bayes \(1763\)](#), a Inferência Bayesiana compreende um método estatístico que usa teoria das probabilidades para atualizar crenças prévias com base em novas evidências observadas. Embora o trabalho de Bayes tenha sido inicialmente recebido com pouca atenção, a formulação e desenvolvimento posterior dos princípios bayesianos foram fundamentais para a evolução da estatística moderna ([STIGLER, 1986](#)).

No século XX, a Inferência Bayesiana ressurgiu com os trabalhos de estatísticos como [Jeffreys \(1939\)](#), que destacou a vantagem da abordagem Bayesiana na combinação de informações prévias e dados observados, permitindo o desenvolvimento dos métodos bayesianos para inferência, de maneira coerente e formal, e aplicou-os em diversas áreas científicas, tais como geofísica e astrofísica, nas quais a incorporação de conhecimento prévio era crucial para obtenção de estimativas mais precisas. Além disso, a partir da década de 1950, com o advento da computação moderna, se tornou possível a implementação prática de modelos bayesianos complexos que permitiu à Inferência Bayesiana experimentar um crescimento significativo. Nesse sentido, a aplicação da Inferência Bayesiana tem sido, desde então, amplamente explorada em diversos campos como biologia ([GILKS; RICHARDSON; SPIEGELHALTER, 1996; DAVISON, 2003](#)), medicina ([SPIEGELHALTER; ABRAMS; MYLES, 1999](#)), engenharia ([BERGER, 1985; CHEN; SHAO; IBRAHIM, 2004](#)), economia ([ZELLNER, 1971; GEWEKE, 1999](#)), ecologia ([HOOTEN; HOBBS, 2015](#)), entre outros.

Nessa perspectiva, a Inferência Bayesiana oferece uma base teórica que se destaca como uma poderosa ferramenta prática para análise de dados em diferentes contextos, se diferenciando da abordagem frequentista ou clássica, que se baseia na frequência relativa de eventos, em que o parâmetro de interesse θ é desconhecido, mas considerado fixo. No entanto, essa abordagem pode não ser adequada em muitas situações, tornando-se útil fazer inferências sobre o parâmetro θ sob o ponto de vista bayesiano ([LEE, 2012](#)).

Ao contrário da abordagem clássica, a Inferência Bayesiana considera a probabilidade como uma medida de incerteza subjetiva, incorporando informações prévias e atualizando-as conforme novos dados são obtidos. Esta abordagem permite a atualização de crenças a priori sobre parâmetros desconhecidos usando informações observadas, combinando de maneira rigorosa probabilidade subjetiva e evidências empíricas ([JAYNES, 2003; GELMAN et al., 2014a; ROBERT, 2007](#)).

O Teorema de Bayes é a ferramenta fundamental da Inferência Bayesiana, o qual

fornece uma maneira de atualizar as crenças sobre um parâmetro de interesse θ com base em novas evidências observadas. Esse parâmetro de interesse θ pode ser pensado como a realização de uma variável aleatória cuja variabilidade pode ser descrita por uma distribuição de probabilidade que representa a crença ou informação disponível a respeito do parâmetro antes de realizar o experimento, sendo denominada distribuição a *priori*. Nesse contexto, surge também o conceito de distribuição a *posteriori*, que refere-se às crenças atualizadas sobre os parâmetros do modelo após a observação dos dados, refletindo nossa incerteza atualizada sobre os parâmetros, levando em consideração as informações prévias e as novas evidências (GELMAN et al., 2014b).

Neste trabalho, por meio da metodologia bayesiana, desenvolveu-se uma investigação sobre distribuições alternativas a priori conjugada para o modelo binomial, baseando-se em generalizações da distribuição beta. Com a estimação do parâmetro de proporção π da distribuição binomial, utilizando a abordagem bayesiana, que garantiu que a distribuição a posteriori das generalizações de beta pertence à mesma classe da priori escolhida, foi realizado um estudo detalhado das propriedades dessas novas distribuições, em especial com as distribuições Kummer beta e Hipergeométrica de Gauss.

À vista do exposto, em outros trabalhos na área da Inferência Bayesiana com foco em modelos binomiais geralmente utilizam a distribuição beta como priori, devido à sua capacidade de conjugação. Esses trabalhos normalmente se concentram em analisar como essa distribuição beta se aplica a diferentes cenários, estimando o parâmetro de proporção π da distribuição binomial. No entanto, a metodologia bayesiana proposta nesta dissertação visa introduzir distribuições alternativas, baseadas em generalizações da distribuição beta, que possuam a capacidade de conjugação, e além disso, ofereçam melhores resultados quando comparados à distribuição beta, em termos de ajuste. Nesse sentido, será realizada uma avaliação do desempenho dessas distribuições com relação ao ajuste de dados, com simulações que mostrarão suas vantagens em relação à distribuição beta.

1.1 Objetivos

Este trabalho busca propor generalizações da distribuição beta como priori conjugada para o parâmetro π da distribuição binomial, que acontece quando a distribuição a posteriori pertence à mesma família de distribuições da priori, e comparar essas generalizações com a distribuição beta usual, quando esta é utilizada como priori.

Alguns objetivos são estabelecidos para explorar e avaliar o impacto dessas generalizações, tais como: promover o estudo das propriedades dessas generalizações, a forma da função de densidade, momentos; e avaliar como as generalizações de beta se comparam com a distribuição beta usual em termos de ajuste em Inferência Bayesiana.

1.2 Organização do trabalho

Inicialmente, no Capítulo 2, serão abordados os conceitos fundamentais da Inferência Bayesiana, com ênfase na apresentação do Teorema de Bayes como uma ferramenta para atualizar crenças sobre parâmetros de interesse com base em novas evidências, a definição de priori conjugada, especialmente no contexto do modelo binomial, e a escolha da distribuição beta como uma opção usual para a priori devido à sua conjugação. O capítulo também destaca a capacidade da Inferência Bayesiana de quantificar incertezas e adaptar-se a novos dados, tornando-a uma ferramenta valiosa em diversas aplicações práticas.

Na sequência, o Capítulo 3 será dedicado à apresentação das distribuições beta generalizadas, que são propostas como alternativas à distribuição beta usual, no contexto da Inferência Bayesiana, discutindo suas propriedades matemáticas e a forma de suas funções densidade de probabilidade (FDP), que via Teorema de Bayes, em combinação com a verossimilhança da binomial, resultará em uma posteriori pertencente à mesma classe de distribuição da priori utilizada – estes resultados podem ser vistos no Apêndice A. Cada uma dessas distribuições é analisada em termos de sua capacidade de atuar como priori conjugada para o modelo binomial, mantendo a propriedade de conjugação, que facilita a aplicação da Inferência Bayesiana.

Mais adiante, o Capítulo 4 abordará a aplicação da Inferência Bayesiana no contexto do modelo binomial. Nele, será descrito o método de geração de números pseudoaleatórios que foi utilizado para as simulações. O capítulo inclui, também, uma apresentação de histogramas que representam as distribuições simuladas, permitindo uma visualização clara da adequação dos modelos aos dados.

Em seguida, o Capítulo 5, se concentra na implementação prática da Inferência Bayesiana utilizando dados simulados, com o objetivo de avaliar o desempenho das distribuições Kummer beta e Hipergeométrica de Gauss em comparação com a distribuição beta usual, utilizando os critérios AIC, BIC e DIC. Nesse contexto, tem-se destaque para a importância da escolha do modelo adequado e suas implicações.

Finalmente, o Capítulo 6 apresenta os resultados, os quais indicam que a distribuição beta pode não ser a mais adequada para representar os dados no contexto da aplicação realizada, concluindo que é sábio considerar modelos alternativos para uma descritiva mais precisa dos dados e que a Inferência Bayesiana apresenta uma abordagem valiosa na análise estatística, especialmente quando lidamos com incertezas em contextos práticos.

2 Referencial Teórico

Nesse capítulo, inicialmente, serão abordados conceitos fundamentais da Inferência Bayesiana, em especial, a apresentação do Teorema de Bayes e a definição de priori conjugada, especificamente no contexto do modelo binomial. Além disso, será mostrado como se dá a estimação dos parâmetros do modelo binomial pelo método de máxima verossimilhança, e depois, pela abordagem bayesiana.

2.1 Conceitos fundamentais na Inferência Bayesiana

Na abordagem bayesiana, fundamentada nos trabalhos de [Bayes \(1763\)](#) e [Jeffreys \(1939\)](#), assume-se que o parâmetro de interesse, definido como uma medida numérica que descreve alguma característica de uma população, segue uma distribuição de probabilidade, também definida pela verossimilhança dos valores observados, conforme apresentado por [Box e Tiao \(1973\)](#).

Na inferência bayesiana, a probabilidade é uma medida de incerteza a respeito da ocorrência de um evento (ou observação de um valor). A probabilidade pode ser definida de forma subjetiva (o quanto você acredita que tal evento ocorrerá?), mas é atualizada pelos dados, conforme comenta [Gamerman e Migon \(1993\)](#) ao dizer que é desenvolvida na presença de observações, cujos valores, inicialmente, vêm de distribuições incertas e são descritos como uma função de densidade.

Para uma variável aleatória, a função que representa a distribuição de probabilidade é chamada de função de probabilidade (FP) no caso de variáveis discretas e FDP no caso de variáveis contínuas. A FP atribui probabilidades específicas a valores discretos, enquanto a FDP descreve a forma da curva de distribuição da probabilidade de ocorrência de cada valor para uma variável aleatória contínua. Uma forma alternativa de poder visualizar os dados é mediante uma função de distribuição cumulativa (CDF), que indica a probabilidade de uma variável aleatória assumir um valor menor ou igual a um valor selecionado. Para variáveis discretas, a CDF é a soma das probabilidades acumuladas da função massa de probabilidade, enquanto para variáveis contínuas, é a integral da função de densidade de probabilidade, que pode ser visto em [DeGroot e Schervish \(2012\)](#).

Tal como na Inferência Clássica, na Inferência Bayesiana a distribuição de probabilidade de uma variável aleatória possui forma funcional indexada por um parâmetro desconhecido. No entanto, na abordagem clássica, os parâmetros são fixos, enquanto na bayesiana são considerados aleatórios. Seja θ um vetor de parâmetros a ser estimado, assumindo valores em Θ , que é denominado espaço paramétrico e representa o conjunto de

todos os valores que o parâmetro pode assumir. Segundo Nascimento (2012), a abordagem bayesiana incorpora o conhecimento sobre θ por meio de uma função de probabilidade $P(\theta)$ ou função densidade de probabilidade $f(\theta)$, denominada distribuição a priori, pois se trata da distribuição de probabilidade de θ antes que se observem os dados amostrais representados pelo vetor y . Aqui, θ será tratado como um vetor de variáveis contínuas, por isso sua distribuição a priori é denotada por $f(\theta)$.

Na Inferência Bayesiana, os dados Y_1, Y_2, \dots, Y_n , possuem uma distribuição de probabilidade descrita pela forma funcional $f(y_i|\theta)$, $i = 1, \dots, n$, indexada por um vetor de parâmetros θ aleatório e desconhecido (não observável). Uma vez observada uma amostra y_1, y_2, \dots, y_n , define-se a função de verossimilhança:

$$L(\theta|y) = \prod_{i=1}^n f(y_i|\theta).$$

Neste contexto, a Inferência Bayesiana trabalha na presença de observações y cujos valores inicialmente incertos são descritos através de uma distribuição de probabilidade com densidade ou função de probabilidade $f(y|\theta)$, sendo função de θ dada a amostra y efetivamente observada, da qual deriva a função de verossimilhança $L(\theta|y)$. De acordo com Nascimento (2012), a função de verossimilhança e a distribuição a priori são os dois ingredientes necessários para estimar o comportamento dos parâmetros. Com isso, tem-se então a distribuição do vetor de parâmetros, após a inserção da informação contida nos dados observados em combinação com a informação a priori, denominada distribuição a posteriori.

Assim, toda a inferência estatística (sob a teoria bayesiana) sobre θ é baseada na sua função densidade de probabilidade depois que os dados são observados, representada por $f(\theta|y)$ e obtida por meio do Teorema de Bayes, dado por:

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)},$$

em que $f(y|\theta) = L(\theta|y)$ é a verossimilhança advinda dos dados observados e $f(y) = \int f(y|\theta)f(\theta)d\theta$ a densidade marginal de Y .

Assim, o Teorema de Bayes pode ser reescrito:

$$f(\theta|y) = \frac{L(\theta|y)f(\theta)}{\int f(y|\theta)f(\theta)d\theta}.$$

Nota-se que $1/f(y)$, que não depende de θ , funciona como uma constante normalizadora de $f(\theta|y)$. Dessa forma, em problemas de estimação de parâmetros, pode-se omitir a constante de normalização $f(y)$ no cálculo da posteriori, gerando, portanto, a forma usual do Teorema de Bayes:

$$f(\theta|y) \propto L(\theta|y)f(\theta),$$

em que \propto representa o símbolo de proporcionalidade, permitindo escrever a função a posteriori $f(\theta|y)$ de uma forma mais simplificada.

A utilização de informação a priori requer a especificação de uma distribuição a priori para a quantidade de interesse θ . É possível definir uma família paramétrica de densidades a partir do conhecimento que se tem sobre θ . Dessa forma, a distribuição a priori é representada por uma forma funcional aos quais os parâmetros devem ser especificados de acordo com este conhecimento. Estes parâmetros indexadores da família de distribuições a priori são chamados de hiperparâmetros para distingui-los dos parâmetros de interesse θ . Nesse contexto, um importante conceito que se destaca é o de priori conjugada (GELMAN et al., 2014b).

Trata-se de uma escolha especial para a distribuição a priori $f(\theta)$, que, quando combinada com uma função de verossimilhança específica, resulta em uma distribuição a posteriori que pertence à mesma família da distribuição a priori. Sendo útil, pois muitas vezes, o cálculo da distribuição a posteriori quando utilizamos algumas prioris pode ser exaustivo ou resulta em distribuições não conhecidas. Ao fazer uso das prioris conjugadas cuja ideia é de que a priori e a posteriori pertençam à mesma classe de distribuições, a atualização do conhecimento se dá apenas com a mudança dos hiperparâmetros (O'HAGAN, 1994).

2.2 Estimação do parâmetro no modelo binomial

A estimação de parâmetros é um dos aspectos fundamentais da inferência estatística, permitindo que modelos sejam ajustados aos dados observados. Nesta seção, inicialmente, focaremos na estimação de parâmetro do modelo binomial utilizando o método da máxima verossimilhança, e depois, pela Inferência Bayesiana.

O modelo binomial é apropriado para descrever a probabilidade de um determinado número de sucessos em uma série de experimentos de Bernoulli independentes e identicamente distribuídos (CASELLA; BERGER, 2002).

Diz-se que a variável aleatória Y tem distribuição binomial, com parâmetros n e π , denotado por $Y \sim Bin(n, \pi)$, se sua função de probabilidade é dada por:

$$f(y|\pi) = \binom{n}{y} \pi^y (1 - \pi)^{n-y}, y = 0, 1, \dots, n,$$

em que $0 < \pi < 1$ e $n \in \mathbb{N}$.

Na Estatística, trabalha-se frequentemente com grandes populações, em que π , entendida como uma proporção da população, possui um determinado atributo. Ao selecionar uma amostra aleatória dessa população e definir Y como o número de indivíduos na amostra que possuem o atributo, é possível contar o número total de “sucessos” em n

tentativas independentes, onde cada tentativa tem dois resultados possíveis: “sucesso” e “fracasso”. Um sucesso em uma tentativa significa que o item selecionado possui o atributo. Mais detalhes podem ser encontrados no trabalho de Bolstad (2004).

A probabilidade de sucesso em qualquer tentativa individual é π , a proporção da população que possui o atributo. Essa proporção permanece constante em todas as tentativas porque a população é grande. Assim, a distribuição condicional da observação Y , dado o parâmetro π , segue uma distribuição $Bin(n, \pi)$.

Nota-se que a distribuição binomial é caracterizada por dois parâmetros: n e π . Aqui, o interesse será na estimação do parâmetro π dado um número fixo de ensaios n .

O método da máxima verossimilhança é um procedimento utilizado para estimar os parâmetros de um modelo ao maximizar a função de verossimilhança (HOGG; TANIS, 2015). Seja y_1, y_2, \dots, y_m uma amostra aleatória de tamanho m , provinda de uma distribuição $Bin(n, \pi)$, ou seja, a variável aleatória Y que conta o número de sucessos em n tentativas segue uma distribuição binomial com parâmetros n e π , em que cada experimento resulta em y_i sucessos em n ensaios, a função de verossimilhança do modelo binomial pode ser expressa como:

$$L(n, \pi) = \prod_{i=1}^m P(Y = y_i) = \prod_{i=1}^m \binom{n}{y_i} \pi^{y_i} (1 - \pi)^{n - y_i}. \quad (2.1)$$

Para simplificar a maximização, toma-se o logaritmo da função de verossimilhança, resultando na função de log-verossimilhança:

$$\log(L(n, \pi)) = \sum_{i=1}^m \log \left(\binom{n}{y_i} \right) + \sum_{i=1}^m y_i \log(\pi) + \sum_{i=1}^m (n - y_i) \log(1 - \pi).$$

A primeira soma, que envolve o coeficiente binomial, é uma constante em relação a π e, portanto, pode ser ignorada na maximização.

Para encontrar a estimativa de π , deriva-se a função de log-verossimilhança em relação a π e iguala-se a zero:

$$\frac{\partial}{\partial \pi} \log L(n, \pi) = \frac{\sum_{i=1}^m y_i}{\pi} - \frac{mn - \sum_{i=1}^m y_i}{1 - \pi} = 0,$$

em que $\sum_{i=1}^m y_i$ representa o número total de sucessos observados na amostra. Rearranjando a equação, obtemos:

$$(1 - \pi) \sum_{i=1}^m y_i = (mn - \sum_{i=1}^m y_i) \pi.$$

Resolvendo esta equação, encontramos a estimativa de máxima verossimilhança para π :

$$\hat{\pi} = \frac{\sum_{i=1}^m y_i}{mn}.$$

Essa estimativa representa a proporção de sucessos observados em relação ao número total de ensaios, refletindo a média amostral. No contexto bayesiano, consideraremos o caso $m = 1$.

A abordagem para estimar a probabilidade de sucesso em experimentos binomiais, por meio da Inferência Bayesiana, é um método que combina princípios de probabilidade com a teoria bayesiana e tem como objetivo fornecer uma estimativa detalhada e informada sobre o parâmetro de interesse π , que é a probabilidade de sucesso em uma série de ensaios.

Na Inferência Bayesiana, o parâmetro π não é tratado como uma constante desconhecida, mas como uma variável aleatória com uma distribuição de probabilidade. A abordagem bayesiana combina o conhecimento prévio sobre π com dados observados para atualizar as crenças sobre o valor de π (GELMAN et al., 2013).

Para usar o Teorema de Bayes, faz-se necessário uma distribuição a priori $g(\pi)$ que dê a nossa crença os possíveis valores do parâmetro π antes de obter os dados. Assim, a primeira etapa na estimação bayesiana é definir uma distribuição a priori para o parâmetro π . Em posse da priori $g(\pi)$ escolhida e da verossimilhança, por meio do Teorema de Bayes, utilizando a fórmula:

$$g(\pi|y) \propto g(\pi) \times f(y|\pi),$$

obtém-se a densidade posteriori.

Dependendo da priori $g(\pi)$ escolhida, pode não necessariamente haver uma forma fechada para a integral da função resultante do produto entre a função de verossimilhança e a priori sobre o intervalo $(0,1)$, então pode ser necessário fazer a integração numericamente. Desse modo, a distribuição Beta é uma escolha bastante comum para essa função a priori devido à sua conjugação com o modelo binomial.

A metodologia bayesiana quando aplicada ao modelo binomial, oferece uma maneira de analisar dados e inferir parâmetros que pode ser mais informativa quando comparada com abordagens tradicionais.

Para estimar o parâmetro π da distribuição binomial, utilizando a Inferência Bayesiana, usualmente, utiliza-se a distribuição beta como a priori, conforme mencionado anteriormente. Essa prática é comum na Inferência Bayesiana, devido à sua propriedade de conjugação. Isso significa que, ao atualizar uma distribuição beta com dados provenientes de uma distribuição binomial, o resultado é outra distribuição beta, o que simplifica os cálculos e a interpretação.

A distribuição beta permite que uma variedade de incertezas possa ser modelada por ela de forma útil. A sua FDP é dada por:

$$f(\pi; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} \pi^{\alpha-1} (1 - \pi)^{\beta-1}, \quad (2.2)$$

para $0 < \pi < 1$, em que $\alpha > 0$, $\beta > 0$ e $B(\alpha, \beta)$ denota a função beta definida por:

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1 - t)^{\beta-1} dt.$$

Os parâmetros α e β são relacionados simetricamente por $f(\pi; \alpha, \beta) = f(1 - \pi; \beta, \alpha)$. Isso implica que se uma variável aleatória X tem distribuição beta com parâmetros α e β , então $1 - X$ tem distribuição beta com parâmetros β e α . Mais detalhes podem ser encontrados no trabalho de [Nadarajah e Kotz \(2007a\)](#).

Assim, utilizando-a como priori e combinando-a com a distribuição binomial, pode-se obter a distribuição a posteriori, via Teorema de Bayes, da seguinte forma:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = \frac{\pi^{\alpha-1} (1 - \pi)^{\beta-1}}{B(\alpha, \beta)} \times \binom{n}{y} \pi^y (1 - \pi)^{n-y} \\ &= \binom{n}{y} \frac{\pi^{y+\alpha-1} (1 - \pi)^{n-y+\beta-1}}{B(\alpha, \beta)} \\ &\propto \pi^{y+\alpha-1} (1 - \pi)^{n-y+\beta-1}, \end{aligned}$$

o que segue que a função resultante é também a densidade de uma distribuição $Beta(\alpha', \beta')$, com parâmetros atualizados: $\alpha' = \alpha + y$ e $\beta' = \beta + n - y$.

O uso da distribuição beta como uma distribuição a priori para o modelo binomial tem algumas vantagens e desvantagens. Entre as vantagens está o fato de que a distribuição a posteriori sempre existe na mesma forma funcional simples e fechada. Esta forma é facilmente avaliada e suas propriedades, como momentos, também estão disponíveis na forma fechada. As desvantagens da classe de distribuições beta a priori são várias, entre as quais, tais distribuições são de rápido crescimento com caudas muito finas, causando alguns problemas no ajuste de alguns conjuntos de crenças a priori.

Além disso, o grau de assimetria é altamente dependente da média. Por exemplo, uma média acima de 0.5 implica assimetria negativa, e uma média abaixo de 0.5 implica assimetria positiva. Em geral, o grau de assimetria pode ser muito pequeno ou muito grande em uma aplicação específica. Isso também causa problemas de ajuste. Mais detalhes podem ser verificados no trabalho de [James e Melvin \(1984\)](#).

3 Distribuições Beta generalizadas

As distribuições beta generalizadas, que incluem a distribuição beta usual como um caso particular, oferecem uma flexibilidade maior na modelagem de dados. Este trabalho sugere que essas distribuições podem melhorar a precisão das inferências estatísticas em comparação com a distribuição beta, quando se utiliza a abordagem bayesiana. A seguir, serão definidas algumas das generalizações da distribuição beta propostas como distribuições alternativas a priori.

3.1 Distribuição Kummer beta

Ng e Kotz (1995) propuseram a distribuição Kummer beta no intervalo $(0, 1)$ com FDP dada por:

$$f(x; a, b, c) = K(a, b, c) x^{a-1} (1-x)^{b-1} \exp(-cx), \quad 0 < x < 1, \quad (3.1)$$

com $a > 0$, $b > 0$ e $-\infty < c < \infty$, em que:

$$K(a, b, c)^{-1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} {}_1F_1(a; a+b; -c)$$

e

$${}_1F_1(a; a+b; -c) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x^{a-1} (1-x)^{b-1} \exp(-cx) dx = \sum_{k=0}^{\infty} \frac{(a)_k (-c)^k}{(a+b)_k k!}$$

é a função hipergeométrica confluyente proposta por Abramowitz e Stegun (1968), $\Gamma(\cdot)$ é a função gama dada por:

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

com $\alpha > 0$ e $(d)_k = d(d+1) \dots (d+k-1)$ denota o fatorial ascendente.

De maneira independente, Gordy (1998) também definiu a distribuição Kummer beta em relação ao problema de leilão de valor comum. Note que a função densidade de probabilidade da beta surge como o caso particular de (3.1) para $c = 0$.

A média $E(X)$ é dada por:

$$E(X) = \int_0^1 x f(x; a, b, c) dx = \frac{a}{a+b} \cdot \frac{{}_1F_1(a+1; a+b+1; -c)}{{}_1F_1(a; a+b; -c)}.$$

A moda é dada por:

$$m_o = \begin{cases} \frac{\sqrt{(a+b+c-2)^2 + 4(1-a)c} + a + b + c - 2}{2c}, & c \neq 0; \\ \frac{a-1}{a+b-2}, & c = 0. \end{cases}$$

3.2 Distribuição Hipergeométrica de Gauss

Armero e Bayarri (1994) e Gordy (1998) introduziram independentemente a distribuição Hipergeométrica de Gauss, dada pela seguinte FDP:

$$f(x) = \frac{Kx^{a-1}(1-x)^{b-1}}{(1+dx)^c}, \quad (3.2)$$

para $0 < x < 1$, $a > 0$, $b > 0$, $-\infty < c < \infty$ e $d > -1$, com:

$$\frac{1}{K} = B(a, b) {}_2F_1(c, a; a+b; -d),$$

em que $B(a, b)$ e ${}_2F_1(a, b; c; x)$ são a função beta e a função Gauss Hipergeométrica, respectivamente, definidas por:

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt$$

e

$${}_2F_1(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tx)^{-a} dt = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!},$$

com $(a)_k = a(a+1)\dots(a+k-1)$ um fatorial ascendente.

A distribuição beta é o caso particular da Equação (3.2) para $c = 0$ ou $d = 0$. A distribuição beta de Libby e Novick (1982) é o caso particular da Equação (3.2) para $c = a + b$. Se X é uma variável aleatória com distribuição Hipergeométrica de Gauss com parâmetros (a, b, c, d) , então $1 - X$ também é uma variável aleatória com distribuição Hipergeométrica de Gauss com parâmetros $(b, a, c, d/(1+d))$. Portanto, X é simétrica, se e somente se $a = b$ e $d = 0$.

Houve apenas alguns artigos estudando a distribuição Hipergeométrica de Gauss. Nadarajah (2008) propôs uma extensão bivariada de (3.2). Nagar e Valencia (2011) derivaram as distribuições do produto e do quociente de duas variáveis aleatórias beta Hipergeométricas de Gauss independentes. Nagar e Valencia (2011) também derivaram expressões para a função de distribuição acumulada, momentos, entropia de Shannon (1948) e entropia de Rényi (1960) correspondentes a (3.2). Gupta e Nagar (2012) propuseram uma extensão matricial de (3.2). Nagar, Bedoya-Valencia e Nadarajah (2015) propuseram uma extensão multivariada de (3.2).

De acordo com a equação proposta no trabalho de Prudnikov, Brychkov e Marichev (1986), a função geradora de momentos de X é dada por:

$$M_X(t) = E[e^{tX}] = K \int_0^1 e^{tx} x^{a-1} (1-x)^{b-1} (1+dx)^c dx = KB(a, b) \Phi_1(a, -c, a+b; -d, -t),$$

em que:

$$\Phi_1(a, b, c, x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a)_{m+n} (b)_n x^m y^n}{(c)_{m+n} m! n!}.$$

A função característica correspondente é:

$$\phi_X(t) = E[e^{itX}] = KB(a, b)\Phi_1(a, -c, a + b; -d, -it),$$

em que $i = \sqrt{-1}$.

O n -ésimo momento de X é dado por:

$$\begin{aligned} E(X^n) &= K \int_0^1 x^{n+a-1}(1-x)^{b-1}(1+dx)^c dx \\ &= \frac{\Gamma(a+b)\Gamma(a)}{\Gamma(n+a)\Gamma(n+a+b)} {}_2F_1(c, n+a; n+a+b; -d) {}_2F_1(c, a; a+b; -d). \end{aligned}$$

Em particular, para $n = 1$, segue que $E(X)$ é dada por:

$$E(X) = \frac{a}{a+b} {}_2F_1(c, 1+a; 1+a+b; -d) {}_2F_1(c, a; a+b; -d).$$

3.3 Distribuição Apell beta

[Nadarajah e Kotz \(2006b\)](#) introduziram mais uma das generalizações da distribuição beta, dada pela seguinte FDP:

$$f(x) = \frac{Cx^{\alpha-1}(1-x)^{\beta-1}}{(1-ux)^\rho(1-vx)^\lambda}, \quad (3.3)$$

para $0 < x < 1$, $\alpha > 0$, $\beta > 0$, $\rho > 0$, $\lambda > 0$, $-1 < u < 1$ e $-1 < v < 1$, em que C denota a constante de normalização, dada por:

$$\frac{1}{C} = B(\alpha, \beta)F_1(\alpha, \rho, \lambda, \alpha + \beta; u, v),$$

e F_1 denota a função Appell do tipo I, dada por:

$$F_1(a, b, c; d; x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a)_{m+n}(b)_m(c)_n x^m y^n}{(d)_{m+n} m! n!}.$$

Nota-se que para $u = 0$ e $v = 0$ a FDP se simplifica para:

$$f(x) = Cx^{\alpha-1}(1-x)^{\beta-1},$$

e ao escolher C como a constante de normalização apropriada, tem-se a forma usual da função de densidade da distribuição beta.

O n -ésimo momento de uma variável aleatória X com FDP dada em (3.3) é dado por:

$$E(X^n) = CB(\alpha + n, \beta)F_1(\alpha + n, \rho, \lambda, \alpha + \beta + n; u, v).$$

Para $n = 1$, segue que a média $E(X)$ de uma variável aleatória X com FDP (3.3) é dada por:

$$E(X) = CB(\alpha + 1, \beta)F_1(\alpha + 1, \rho, \lambda, \alpha + \beta + 1; u, v).$$

3.4 Distribuição beta Bessel I

Gupta e Nadarajah (2006) introduziram as três primeiras generalizações de (2.2), que serão apresentadas na sequência, envolvendo a função de Bessel, definida por:

$$I_m(x) = \frac{x^m}{\sqrt{\pi}2^m\Gamma(m + 1/2)} \int_{-1}^1 (1 - t^2)^{m-1/2} \exp(xt) dt,$$

referindo-se a elas como as distribuições beta Bessel (BB).

A primeira generalização de beta introduzida em Gupta e Nadarajah (2006) é intitulada beta Bessel I (BB I) e dada pela FDP:

$$g(x) = Cx^{\alpha-1}(1 - x)^{\beta-1}I_\nu(cx) \tag{3.4}$$

para $0 < x < 1$, $\nu > 0$, $\alpha > 0$, $\beta > 0$, e $c \geq 0$, em que C denota a constante normalizadora, dada por:

$$\frac{1}{C} = \frac{c^\nu\Gamma(\alpha + \nu)\Gamma(\beta)}{2^\nu\Gamma(\alpha + \beta + \nu)\Gamma(\nu + 1)} {}_2F_3\left(\frac{\alpha + \nu}{2}, \frac{\alpha + \nu + 1}{2}; \nu + 1, \frac{\alpha + \nu + \beta}{2}, \frac{\alpha + \nu + \beta + 1}{2}, \frac{c^2}{4}\right)$$

e ${}_2F_3$ a função hipergeométrica definida por:

$${}_2F_3(a, b; c, d, e; x) = \sum_{k=0}^{\infty} \frac{(a)_k(b)_k}{(c)_k(d)_k(e)_k} \frac{x^k}{k!}.$$

A função densidade de probabilidade da beta surge como o caso particular de (3.4) para $c = 0$ e $\nu = 0$. Vários outros casos particulares de (3.4) podem ser obtidos usando propriedades especiais de $I_\nu(\cdot)$. Nota-se que:

$$I_{3/2}(x) = \sqrt{\frac{2}{\pi}} \frac{\cosh x \sinh x}{x^{3/2}},$$

$$I_{5/2}(x) = \sqrt{\frac{2}{\pi}} \frac{(x^2 + 3) \sinh x - 3 \cosh x}{x^{5/2}},$$

$$I_{7/2}(x) = \sqrt{\frac{2}{\pi}} \frac{x(x^2 + 15) \cosh x - 3(2x^2 + 5) \sinh x}{x^{7/2}},$$

$$I_{9/2}(x) = \sqrt{\frac{2}{\pi}} \frac{(x^4 + 45x^2 + 105) \sinh x - 5x(2x^2 + 21) \cosh x}{x^{9/2}}.$$

De forma mais geral, se $\nu - 1/2 \geq 1$ é um número inteiro, então:

$$I_\nu(x) = \sqrt{2}\sqrt{x\pi} \exp\left\{\frac{\pi i}{2}\left(\frac{1}{2} - \nu\right)\right\} \times \\ \left[\sinh\left(\frac{\pi x}{2}\left(\frac{1}{2} - \nu\right) - x\right) \times \sum_{k=0}^{[(2|\nu|-1)/4]} \frac{(|\nu| + 2k - 1/2)!}{(2k)! (|\nu| - 2k - 1/2)! (2x)^{2k}} + \right. \\ \left. \cosh\left(\frac{\pi x}{2}\left(\frac{1}{2} - \nu\right) - x\right) \sum_{k=0}^{[(2|\nu|-3)/4]} \frac{(|\nu| + 2k + 1/2)! (2x)^{-2k-1}}{(2k+1)! (|\nu| - 2k - 3/2)!} \right].$$

3.5 Distribuição beta Bessel II

A segunda generalização de beta introduzida no trabalho de [Gupta e Nadarajah \(2006\)](#), intitulada de beta Bessel II (BB II), é dada pela seguinte FDP:

$$g(x) = Cx^{\alpha-1}(1-x)^{\beta-1} \exp(cx)I_\nu(cx), \quad (3.5)$$

para $0 < x < 1$, $\nu > 0$, $\alpha > 0$, $\beta > 0$, e $c \geq 0$, em que C denota a constante normalizadora, dada por:

$$\frac{1}{C} = \frac{c^\nu \Gamma(\alpha + \nu) \Gamma(\beta)}{2^\nu \Gamma(\alpha + \beta + \nu) \Gamma(\nu + 1)} {}_2F_2\left(\nu + \frac{1}{2}, \alpha + \nu; 2\nu + 1, \alpha + \beta + \nu; 2c\right),$$

com ${}_2F_2$ função hipergeométrica, dada por:

$${}_2F_2(a, b; c, d; x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k (d)_k} \frac{x^k}{k!}.$$

Nota-se que a função densidade de probabilidade da beta surge como o caso particular de (3.5) para $c = 0$ e $\nu = 0$.

3.6 Distribuição beta Bessel III

A terceira generalização de beta introduzida em [Gupta e Nadarajah \(2006\)](#) é intitulada beta Bessel III (BB III) e dada pela FDP:

$$g(x) = Cx^{\alpha-1}(1-x)^{\beta-1} \exp(-cx)I_\nu(cx), \quad (3.6)$$

para $0 < x < 1$, $\nu > 0$, $\alpha > 0$, $\beta > 0$, e $c \geq 0$, em que C denota a constante normalizadora, dada pela seguinte equação:

$$\frac{1}{C} = \frac{c^\nu \Gamma(\alpha + \nu) \Gamma(\beta)}{2^\nu \Gamma(\alpha + \beta + \nu) \Gamma(\nu + 1)} {}_2F_2\left(\nu + \frac{1}{2}, \alpha + \nu; 2\nu + 1, \alpha + \beta + \nu; -2c\right)$$

com ${}_2F_2$ função hipergeométrica dada por:

$${}_2F_2(a, b; c, d; x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k (d)_k} \frac{x^k}{k!}.$$

A função densidade de probabilidade da beta surge como o caso particular de (3.6) para $c = 0$ e $\nu = 0$.

3.7 Distribuição Beta Trigonométrica Cos I

A generalização de beta chamada Beta Trigonométrica Cos I é dada pela FDP:

$$f(x) = Cx^{\nu-1}(1-x)^{\mu-1} \cos(ax), \quad (3.7)$$

para $0 < x < 1$, $\nu > 0$, $\mu > 0$ e $0 \leq a < \pi/2$, em que a constante C é dada por:

$$\frac{1}{C} = \frac{1}{2}B(\nu, \mu) \{ {}_1F_1(\nu; \nu + \mu; ia) + {}_1F_1(\nu; \nu + \mu; -ia) \},$$

em que $i = \sqrt{-1}$ é um número complexo. Para mais detalhes sobre essa distribuição, consultar o trabalho de [Nadarajah e Kotz \(2006a\)](#). Note que a Equação (3.7) é um número real:

$$\begin{aligned} {}_1F_1(\nu; \nu + \mu; ia) + {}_1F_1(\nu; \nu + \mu; -ia) &= \sum_{k=0}^{\infty} \frac{(\nu)_k}{(\nu + \mu)_k} \frac{a^k}{k!} \{ i^k + (-i)^k \} \\ &= 2 \sum_{k=0}^{\infty} \frac{(\nu)_k}{(\nu + \mu)_{2k}} \frac{(-1)^k a^{2k}}{(2k)!}. \end{aligned}$$

Por aproximação, a distribuição beta é um caso especial da distribuição Beta Trigonométrica Cos I, pois quando a se aproxima de 0, considera-se $\cos(ax) = \cos(0)$ e como $\cos(0) = 1$, a FDP (3.7) é simplificada para a forma usual da função de densidade da distribuição beta ao escolher a constante de normalização C apropriada.

Como pode ser visto no trabalho de [Gradshteyn e Ryzhik \(2000\)](#), o n -ésimo momento associado a BT Cos I pode ser descrito como:

$$E(X^n) = \frac{B(\mu, n + \nu) \{ {}_1F_1(n + \nu; n + \nu + \mu; ia) + {}_1F_1(n + \nu; n + \nu + \mu; -ia) \}}{B(\mu, \nu) \{ {}_1F_1(\nu; \nu + \mu; ia) + {}_1F_1(\nu; \nu + \mu; -ia) \}},$$

para $n \geq 1$. Desse mesmo modo, para $n = 1$, obtemos a média, dada por:

$$E(X) = \frac{B(\mu, 1 + \nu) \{ {}_1F_1(1 + \nu; 1 + \nu + \mu; ia) + {}_1F_1(1 + \nu; 1 + \nu + \mu; -ia) \}}{B(\mu, \nu) \{ {}_1F_1(\nu; \nu + \mu; ia) + {}_1F_1(\nu; \nu + \mu; -ia) \}}.$$

3.8 Distribuição Beta Trigonométrica Sen I

A generalização de beta, chamada de BT Sen I, tem a seguinte FDP senoidal complementar associada ao BT Cos I:

$$g(x) = Cx^{\nu-1}(1-x)^{\mu-1} \text{sen}(ax),$$

para $0 < x < 1$, $\nu > 1$ ($\nu \neq 0$), $\mu > 0$ e $0 < a < \pi$ (numeral), em que a constante C é dada por:

$$\frac{1}{C} = -\frac{i}{2}B(\nu, \mu) \{ {}_1F_1(\nu; \nu + \mu; ia) - {}_1F_1(\nu; \nu + \mu; -ia) \}.$$

Para mais detalhes sobre essa distribuição, consultar o trabalho de [Nadarajah e Kotz \(2006a\)](#).

O n -ésimo momento é dado por:

$$E(X^n) = \frac{B(\mu, n + \nu) \{ {}_1F_1(n + \nu; n + \nu + \mu; ia) - {}_1F_1(n + \nu; n + \nu + \mu; -ia) \}}{B(\mu, \nu) \{ {}_1F_1(\nu; \nu + \mu; ia) - {}_1F_1(\nu; \nu + \mu; -ia) \}},$$

para $n \geq 1$. Esse resultado pode ser encontrado no trabalho de [Gradshteyn e Ryzhik \(2000\)](#). Para $n = 1$ obtemos a média, que é dada por:

$$E(X) = \frac{B(\mu, 1 + \nu) \{ {}_1F_1(1 + \nu; 1 + \nu + \mu; ia) - {}_1F_1(1 + \nu; 1 + \nu + \mu; -ia) \}}{B(\mu, \nu) \{ {}_1F_1(\nu; \nu + \mu; ia) - {}_1F_1(\nu; \nu + \mu; -ia) \}}.$$

3.9 Distribuição Kummer do tipo II

[Chamayou e Letac \(1991\)](#) definiram a distribuição $f_\delta(a, b, c)$, cuja FDP é dada por:

$$f_\delta(a, b, c) = N_\delta(a, b, c)x^{a-1}(1-x)^{b-1}(1+(\delta-1)x)^c I_{[0,1]}(p), I_{(0,1)}(\pi)$$

em que $I_{[0,1]}(x)$ é a função indicadora, que assume o valor 1 se $x \in [0, 1]$ e 0 caso contrário, e $N_\delta(a, b, c)$ é a constante de normalização que garante que a integral da FDP sobre o intervalo $[0, 1]$ seja 1.

O parâmetro a de $f_\delta(a, b, c)$ afeta a forma da distribuição na região de x próxima de 0 e influencia o comportamento da distribuição para valores pequenos de x . O parâmetro b afeta a forma da distribuição na região de x próxima de 1 e influencia o comportamento da distribuição para valores grandes de x . Já δ , é um parâmetro adicional que introduz uma modulação da função $1 + (\delta - 1)x$, esse parâmetro ajusta o “peso” da parte linear da função. Quando $\delta = 1$, a FDP se reduz à forma usual da distribuição beta. O parâmetro c , pode ser um parâmetro adicional de forma ou ajuste, mas seu papel específico depende do contexto de aplicação da distribuição.

Com $f_\delta(a, b, c)$, $a, b, \delta > 0$, $c \in \mathbb{R}$ e $N_\delta(a, b, c)$ a constante normalizadora. Nota-se que $f_\delta(a, b, c) = H(-c, a; a + b; 1 - \delta)$, em que $H(a, b; c; z)$ representa a distribuição hipergeométrica, conforme definido no trabalho de [Chamayou e Letac \(1991\)](#). [Armero e Bayarri \(1994\)](#) consideraram a distribuição hipergeométrica de Gauss $GH(\alpha, \beta, \gamma, z)f_{1+z}(\alpha, \beta, -\gamma)$ como uma distribuição a priori marginal. No entanto, conservamos as notações de [Koudou e Vallois \(2011\)](#).

3.10 Distribuição Beta Generalizada do tipo II

Nadarajah e Kotz (2007b) definiram uma nova generalização de beta, intitulada Beta Generalizada do tipo II, cuja FDP é dada por:

$$f(x) = \frac{Dx^{\alpha-1}(1-x)^{\beta-1}\exp(-px)}{(x+z)^\rho}, \quad (3.8)$$

para $0 < x < 1$, $\alpha > 0$, $\beta > 0$, $\rho \geq 0$, $p \geq 0$ e $-\infty < z < \infty$; em que D denota a constante normalizadora dada por:

$$\frac{1}{D} = B(\alpha, \beta)z^{-\rho}\Phi_1(\alpha, \rho, \alpha + \beta; \frac{-1}{z}, p),$$

em que:

$$\Phi_1(a, b, c, x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a)_{m+n}(b)_n x^m y^n}{(c)_{m+n} m! n!}$$

denota a série hipergeométrica degenerada de duas variáveis.

O n -ésimo momento de uma variável aleatória X com FDP (3.8) pode ser calculado como:

$$E(X^n) = D \int_0^1 \frac{x^{n+\alpha-1}(1-x)^{\beta-1}\exp(-px)}{(x+z)^\rho} dx = DB(\alpha, \beta)z^{-\rho}\Phi_1(\alpha + n, \rho, \alpha + \beta + n; \frac{-1}{z}, p).$$

Desse modo, segue por (3.9) que a média $E(X)$ é dada por:

$$E(X) = D \int_0^1 \frac{x^\alpha(1-x)^{\beta-1}\exp(-px)}{(x+z)^\rho} dx = DB(\alpha, \beta)z^{-\rho}\Phi_1(\alpha + 1, \rho, \alpha + \beta + 1; \frac{-1}{z}, p).$$

3.11 Distribuição Kummer beta II

Ng e Kotz (1995) e Gordy (1998) propuseram a distribuição Kummer beta II com densidade dada por:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}e^{\lambda x}}{B(\alpha, \beta) {}_1F_1(\alpha, \alpha + \beta, \lambda)},$$

com $0 < x < 1$ e ${}_1F_1$ a função hipergeométrica confluyente de Kummer dada por:

$${}_1F_1(a; b; z) = \sum_{n=0}^{\infty} \frac{(a)_n z^n}{(b)_n n!},$$

em que $(a)_n$ é o símbolo de Pochhammer, que representa o produto $a(a+1)(a+2)\dots(a+n-1)$ e $(a)_0 = 1$.

A Distribuição Kummer beta II é caracterizada pelos seguintes parâmetros: $\alpha > 0$ e $\beta > 0$ que controlam a forma da distribuição e $\lambda \in \mathbb{R}$ que modula o fator exponencial na função de densidade, influenciando a forma e a assimetria da distribuição.

4 Inferência para o modelo binomial

Gelman e Hill (2007) afirmam que na Inferência Bayesiana para o modelo binomial, a abordagem consiste em usar a distribuição binomial como a função de verossimilhança e especificar uma distribuição a priori para o parâmetro de interesse, que é a probabilidade de sucesso em um único ensaio. A distribuição a posteriori é obtida aplicando o Teorema de Bayes, combinando a informação da amostra com a priori.

Ao utilizar cada uma das distribuições anteriormente definidas no Capítulo 2 como priori para o modelo binomial, em substituição à distribuição beta, nota-se que todas são conjugadas para o modelo binomial – vide Apêndice A, de tal forma que a distribuição a posteriori segue a mesma distribuição da priori, isto é, pertence à mesma classe de distribuições da priori utilizada.

A seguir, será apresentada uma tabela que sintetiza cada generalização da distribuição beta utilizada como priori, com seus parâmetros explicitados, e a respectiva posteriori pertencendo à mesma classe, porém com os parâmetros atualizados.

Priori	Posteriori	Distribuição beta
Kummer beta (a, b, c)	Kummer beta ($a + y, n + b - y, c$)	Kummer beta ($a, b, 0$)
Hipergeométrica de Gauss (a, b, c, d)	Hipergeométrica de Gauss ($a + y, n + b - y, c, d$)	Hipergeométrica de Gauss ($a, b, 0, d$)
Beta Não-central (a, b, λ)	Beta Não-central ($a + y, n + b - y, \lambda$)	Beta Não-central ($a, 0, 0$)
Apell beta ($\alpha, \beta, \rho, \lambda, u, v$)	Apell beta ($\alpha + y, n + \beta - y, \rho, \lambda, u, v$)	Apell beta ($\alpha, \beta, \rho, \lambda, 0, 0$)
BB I (α, β, c, v)	BB I ($\alpha + y, n + \beta - y, c, v$)	BB I ($\alpha, \beta, 0, 0$)
BB II (α, β, c, v)	BB II ($\alpha + y, n + \beta - y, c, v$)	BB II ($\alpha, \beta, 0, 0$)
BB III (α, β, c, v)	BB III ($\alpha + y, n + \beta - y, c, v$)	BB III ($\alpha, \beta, 0, 0$)
BT Cos I (v, μ, a)	BT Cos I ($v + y, n + \mu - y, a$)	BT Cos I ($v, \mu, 0$)
BT Sen I (v, μ, a)	BT Sen I ($y + v, n - y + \mu, a$)	BT Sen I ($v, \mu, 0$)
Euler (b, c, a, z)	Euler ($b + y, n - y + c, a, z$)	Euler ($b, c, 0, z$)
Beta Hipergeométrica (a, b, a')	Beta Hipergeométrica ($a + y, n - y + b, a'$)	Beta Hipergeométrica ($a, b, 0$)
Kummer do tipo II (a, b, c, δ)	Kummer do tipo II ($a + y, n - y + b, c, \delta$)	Kummer do tipo II ($a, b, c, 0$)
Beta generalizada tipo II ($\alpha, \beta, \rho, p, z$)	Beta generalizada tipo II ($\alpha + y, n + \beta - y, \rho, p, z$)	Beta generalizada tipo II ($\alpha, \beta, 0, 0, z$)
Kummer beta II (α, β, λ)	Kummer beta II ($\alpha + y, n + \beta - y, \lambda$)	Kummer beta II ($\alpha, \beta, 0$)

Tabela 1 – Priori conjugada para o modelo binomial baseada em generalizações da distribuição beta, suas respectivas posterioris e alguns casos especiais que, por aproximação, resultam na distribuição beta.

Dentre as diversas generalizações de beta apresentadas que são conjugadas para o modelo binomial, faremos um estudo mais detalhado para a Kummer beta e Hipergeométrica de Gauss, mas que pode ser estendido às demais generalizações aqui apresentadas de forma equivalente. A seguir, tem-se os gráficos de densidade de ambas as distribuições.

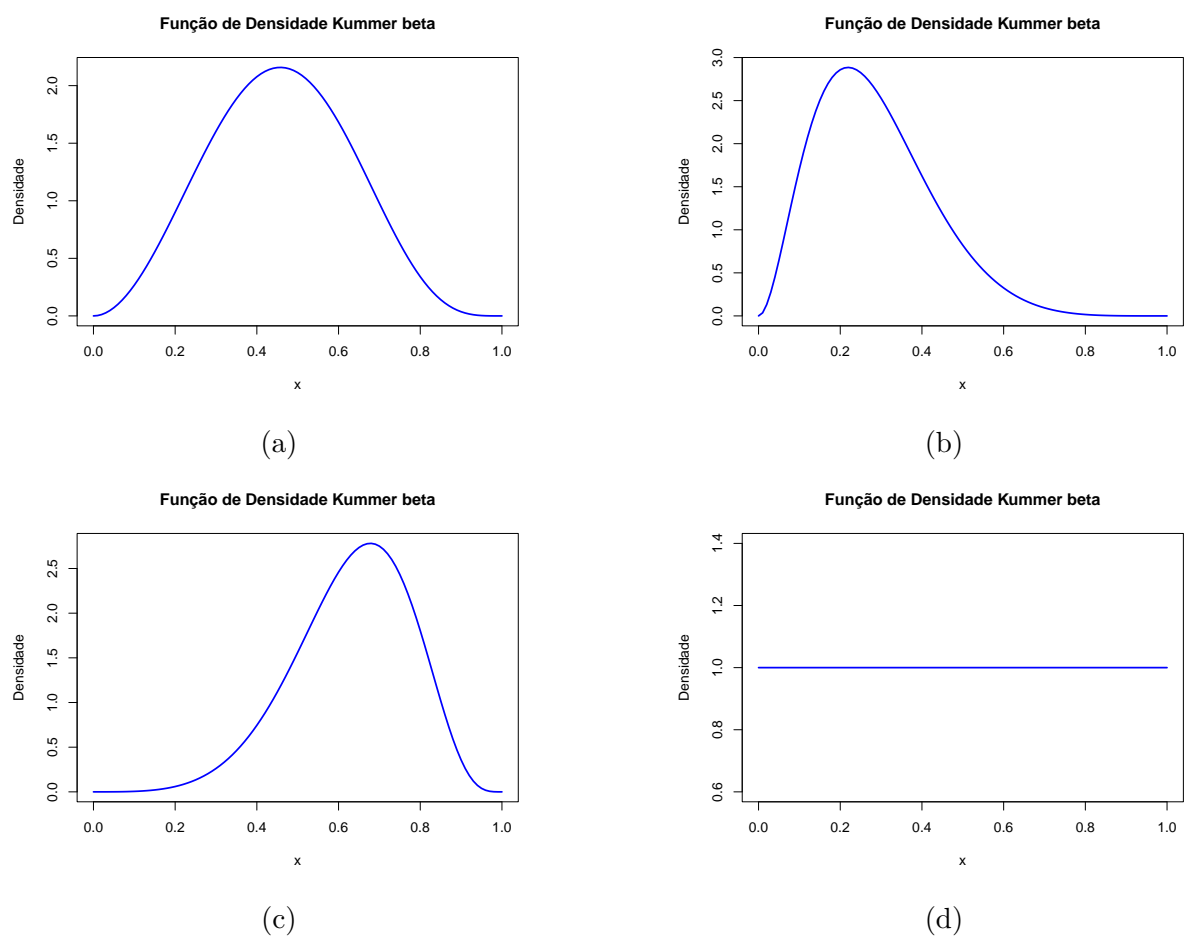


Figura 1 – Densidade da distribuição Kummer beta com parâmetros (a) $a = 3, b = 5, c = -3$, (b) $a = 3, b = 5, c = 4$, (c) $a = 4, b = 5, c = -8$, (d) $a = 1, b = 1, c = 0$.

A Kummer beta é uma extensão da distribuição beta e produz distribuições bimodais em faixa finita para $a < 1$ (e certos valores do parâmetro c).

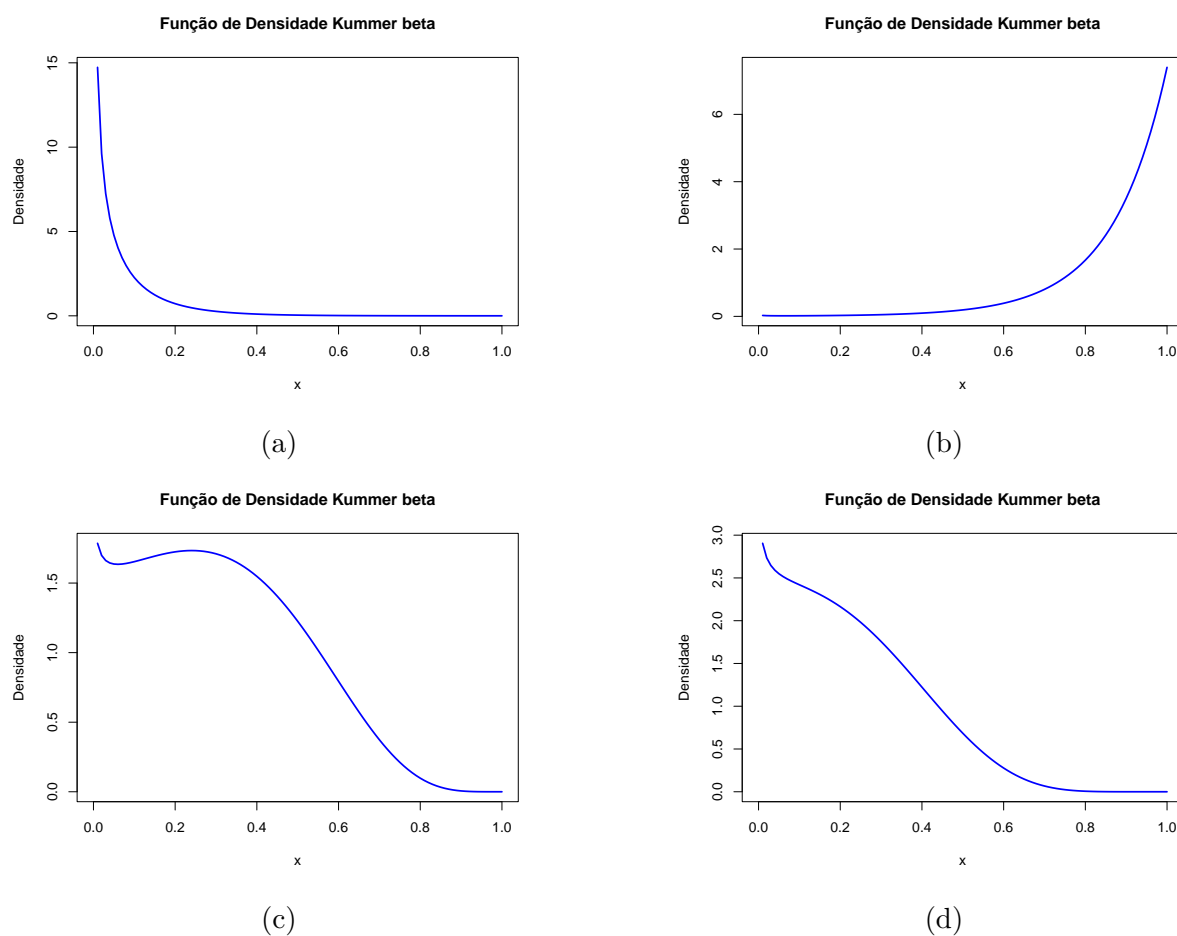


Figura 2 – Densidade da distribuição Kummer beta com parâmetros (a) $a = 0.5, b = 1, c = 8$, (b) $a = 0.5, b = 1, c = -8$, (c) $a = 0.9, b = 6, c = -7$ e (d) $a = 0.9, b = 9, c = -9$.

De modo mais específico, observou-se que quando $0 < a < 1$ e $c < 0$, tal que $2b \leq |c|$, é possível obter a densidade da Kummer beta bimodal em faixa finita, que pode ser mostrado através dos gráficos de densidade a seguir.

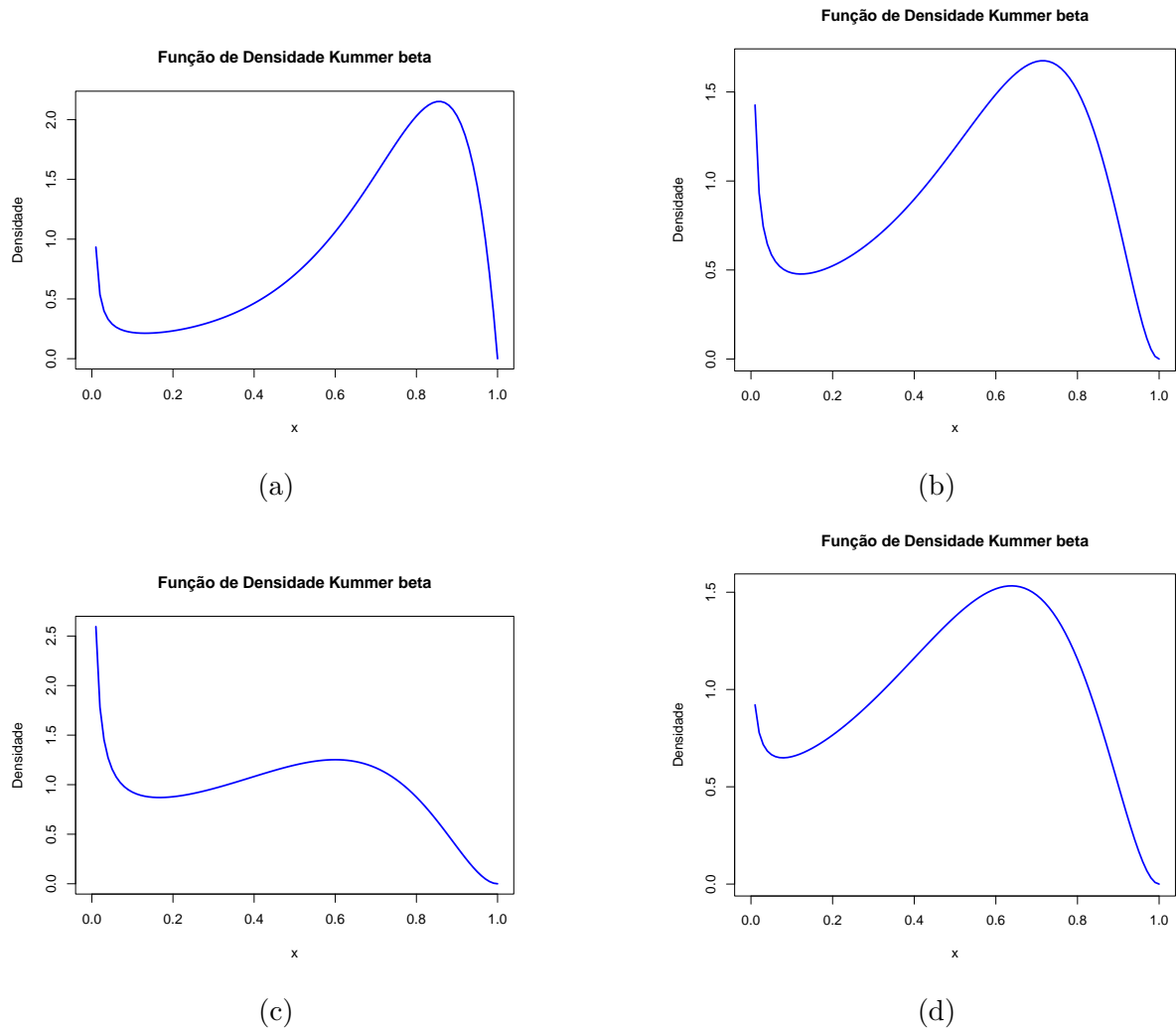


Figura 3 – Densidade da distribuição Kummer beta com parâmetros (a) $a = 0.1, b = 2, c = -8$, (b) $a = 0.3, b = 3, c = -8$, (c) $a = 0.4, b = 3, c = -6$ e (d) $a = 0.7, b = 3, c = -6$.

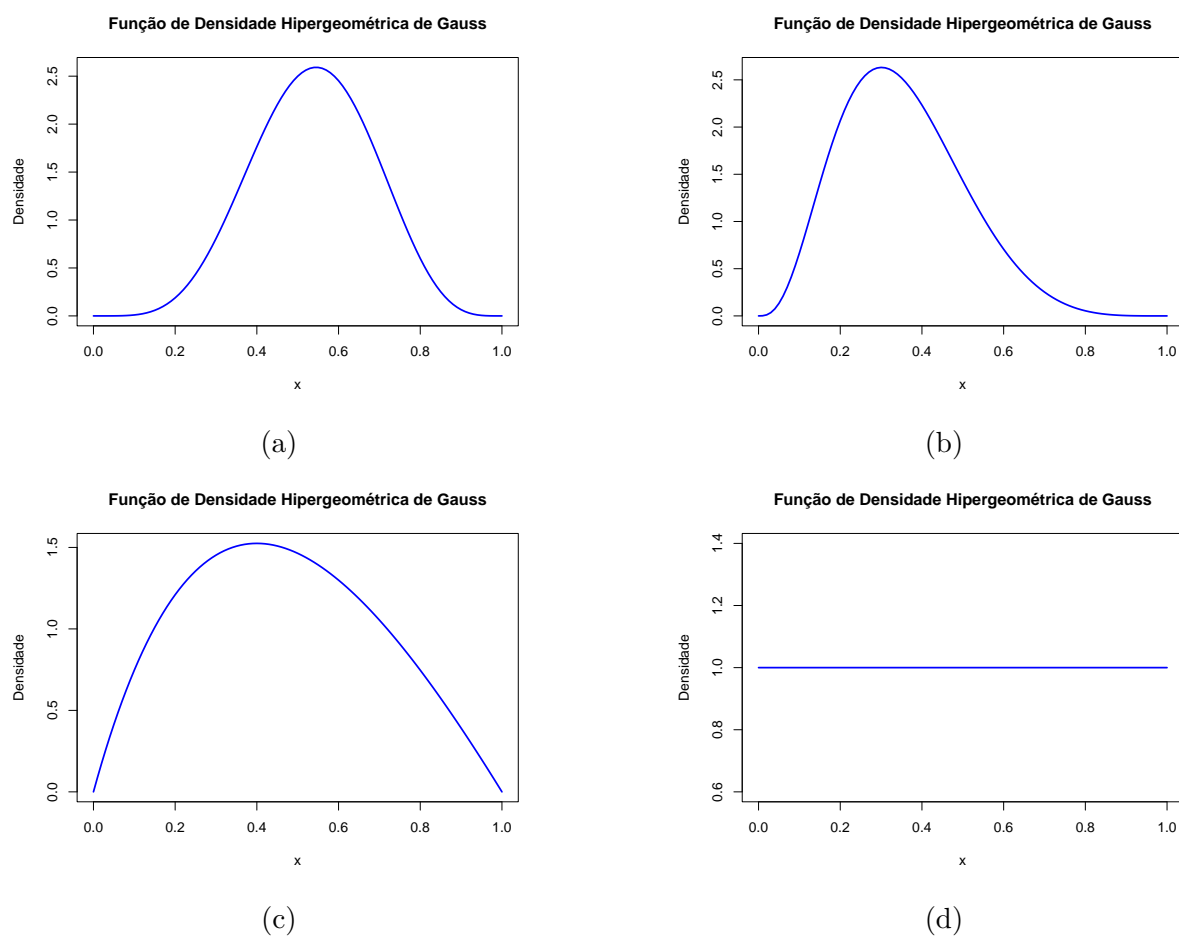


Figura 4 – Densidade da distribuição Hipergeométrica de Gauss com parâmetros (a) $a = 6, b = 5, c = 4$ e $d = 0.1$, (b) $a = 4, b = 5, c = 6$ e $d = 0.9$, (c) $a = 2, b = 2, c = 2$ e $d = 0.5$, (d) $a = 1, b = 1, c = 0$ e $d = 0.5$.

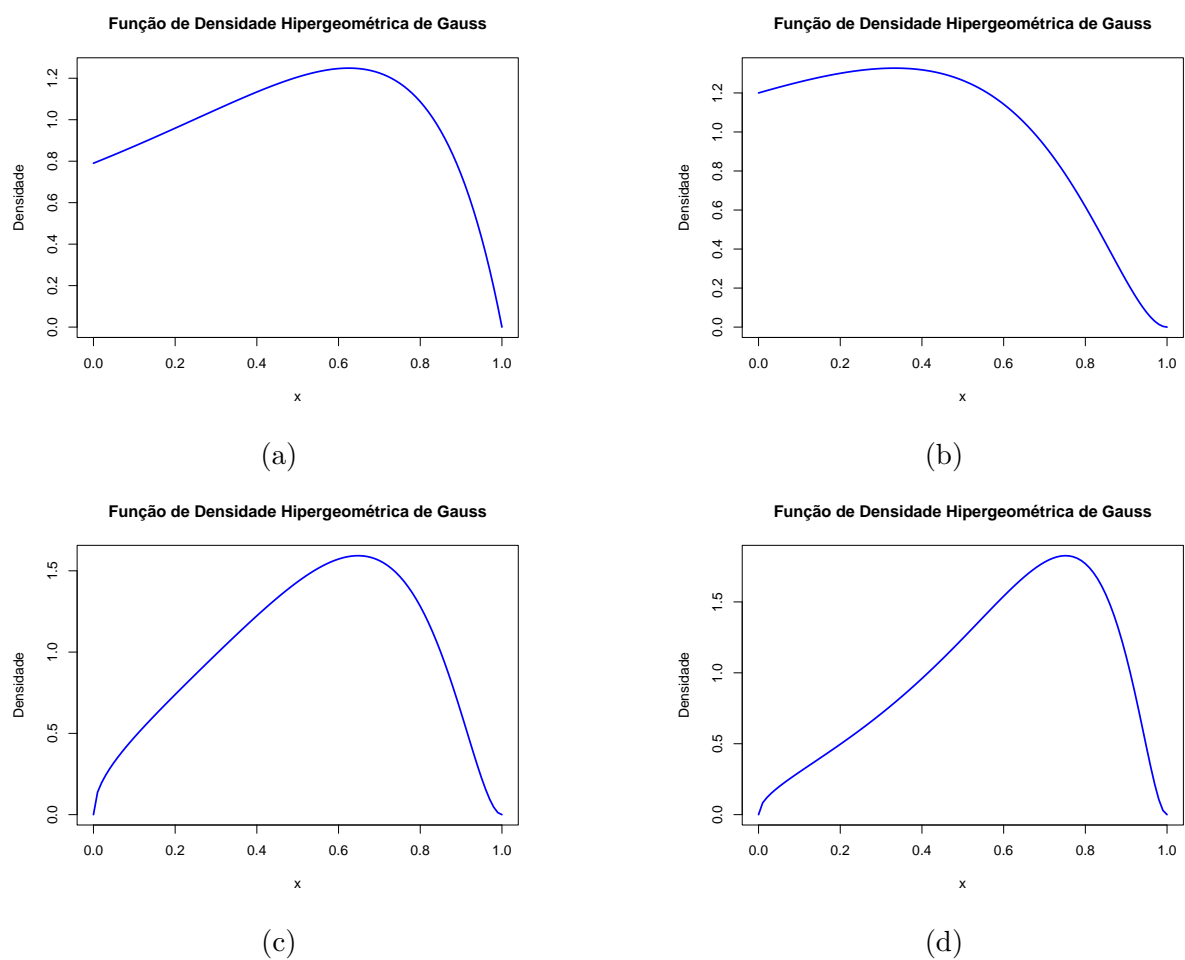


Figura 5 – Densidade da distribuição Hipergeométrica de Gauss com parâmetros (a) $a = 1, b = 2, c = 5$ e $d = -0.4$, (b) $a = 1, b = 3, c = 5$ e $d = -0.5$, (c) $a = 1.5, b = 3, c = 5$ e $d = -0.6$, (d) $a = 1, b = 3, c = 5$ e $d = -0.7$.

Os gráficos representam a FDP da distribuição Kummer beta e da distribuição Hipergeométrica de Gauss, permitindo uma comparação visual da forma e das características dessas distribuições. Cada gráfico ilustra como a escolha dos parâmetros influencia a forma da densidade. A importância dos parâmetros na modelagem das distribuições é um ponto central. Os gráficos mostram que variações nos parâmetros resultam em mudanças significativas nas FDPs, destacando a flexibilidade das distribuições Kummer beta e Hipergeométrica de Gauss para se ajustarem a diferentes contextos e padrões de dados. Para a Kummer beta, por exemplo, as diferentes combinações de parâmetros alteram a altura e a largura da distribuição, o que é crucial para capturar a variabilidade dos dados observados.

Os gráficos também podem refletir resultados obtidos a partir de simulações, indicando como esses modelos se comportam em situações práticas. O uso de histogramas nas simulações permite visualizar a adequação das distribuições aos dados simulados, tornando evidente a eficácia da Kummer beta e Hipergeométrica de Gauss em capturar a dinâmica dos dados quando observados e comparados ao comportamento da linha de densidade.

4.1 Geração de números pseudoaleatórios

O método de aceitação-rejeição, desenvolvido por [Neumann \(1951\)](#), foi um avanço importante no contexto da introdução de métodos estocásticos para gerar amostras de distribuições de probabilidade, oferecendo uma abordagem sistemática para gerar amostras a partir de distribuições complexas, usando uma combinação de amostras de uma distribuição mais simples e a técnica de aceitação-rejeição.

Um outro método de geração de números pseudoaleatórios bastante conhecido e utilizado é o método da inversão, descrito nos trabalhos de [Press et al. \(2007\)](#) e [Gentle \(2003\)](#), no entanto, ele enfrenta limitações, pois em situações em que não se dispõe de uma expressão para a função de distribuição acumulada F , ou casos em que mesmo que se conheça F , não seja possível chegar à uma expressão para a função F^{-1} , isto é, não é possível encontrar uma fórmula fechada para a função inversa da função de distribuição cumulativa de uma variável aleatória, este se torna inviável.

Nesse sentido, para as distribuições beta generalizadas trabalhadas nessa dissertação, tornou-se necessário a utilização do método da aceitação-rejeição, que é uma técnica para gerar amostras de uma variável aleatória com uma distribuição específica, quando a função de distribuição cumulativa ou a FDP não é facilmente invertida, inviabilizando o uso do método de inversão.

O método de aceitação-rejeição, que pode ser visto em trabalhos como [Robert e Casella \(2004\)](#) e [Kroese, Taimre e Botev \(2011\)](#), consiste nos seguintes passos: selecionar uma distribuição proposta com uma FDP $g(x)$ que seja fácil de amostrar e que tenha o mesmo suporte, em que exista uma constante c , tal que $f(x) \leq cg(x)$, para todo x , em que $f(x)$ é a FDP da distribuição alvo; gerar amostras X de acordo com a distribuição proposta $g(x)$; na sequência, para cada amostra X gerada, deve-se gerar uma variável aleatória uniforme U , no intervalo $(0, 1)$; aceitar a amostra X com uma probabilidade igual a: $f(X)/cg(X)$, caso contrário, rejeitar X e voltar ao passo anterior para gerar uma nova amostra, ou seja, X somente é aceito, se: $U \leq f(X)/cg(X)$. No método da aceitação e rejeição, a escolha do parâmetro c é crucial para garantir a eficiência do algoritmo. Deve-se usar sempre o menor valor de c , tal que $f(X) \leq cg(X)$ para que o algoritmo se torne mais eficiente, pois a probabilidade de aceitação é igual a $1/c$. Assim, quanto maior o c , menor é a probabilidade de aceitação, isto é, mais gerações esperamos até a aceitação.

Um problema comum na Inferência Bayesiana é que a forma da distribuição a posteriori pode não resultar em uma distribuição de probabilidade conhecida. Nesses casos, são utilizadas técnicas computacionais para alcançar aproximações da distribuição a posteriori. Na Inferência Bayesiana, uma outra técnica bastante utilizada é a chamada Monte Carlo via Cadeias de Markov (MCMC), que pode ser vista nos trabalhos de [Gamerman e Migon \(1993\)](#).

Há dois principais métodos MCMC, sendo eles o amostrador de Gibbs (Gibbs Sampling) proposto por [Diebolt e Robert \(1994\)](#) e o algoritmo de Metropolis-Hastings, proposto por [Metropolis et al. \(1953\)](#) cujas contribuições foram ampliadas por [Hastings \(1970\)](#). Bastante útil para os casos em que não se conhece a distribuição a posteriori ou quando são demasiadamente complicadas, o método MCMC propicia o processo de simulação que possibilita a geração de amostras de distribuição à posteriori, permitindo o processo inferencial a partir delas.

O método de aceitação-rejeição e o método de Monte Carlo por Cadeias de Markov (MCMC) são usados para gerar amostras aleatórias de distribuições complexas, mas têm características distintas que fazem com que um possa ser mais apropriado que o outro em diferentes situações. Os algoritmos MCMC, como o algoritmo de Metropolis-Hastings e o Gibbs Sampling, são mais complexos e envolvem a construção de cadeias de Markov e a verificação de convergência, o que pode ser mais desafiador.

Para este trabalho, utilizamos o pacote [AcceptReject](#), desenvolvido por [Marinho \(2024\)](#), que fornece a função `accept_reject(.)` no *R*, além de outras funções que implementam o método de aceitação-rejeição de forma otimizada, para gerar observações pseudoaleatórias para variáveis aleatórias discretas ou contínuas. O pacote permite realizar as simulações necessárias para estudar as propriedades das generalizações de beta e aplicar os resultados em uma situação real.

Apesar das várias opções de priori apresentadas nesta dissertação, como já mencionado, iremos analisar de forma detalhada dois desses casos, que são as distribuições Kummer beta e Hipergeométrica de Gauss. Para isso, será apresentado a seguir um quadro comparativo entre as médias teóricas e amostrais obtidas a partir da geração de números aleatórios.

Tamanho da amostra	Distribuição	Média teórica	Média amostral
$m = 100$	Kummer beta	0.3090	0.3355
	H. de Gauss	0.4036	0.4110
$m = 250$	Kummer beta	0.3090	0.3354
	H. de Gauss	0.4036	0.4268
$m = 500$	Kummer beta	0.3090	0.3151
	H. de Gauss	0.4036	0.3969
$m = 1000$	Kummer beta	0.3090	0.3076
	H. de Gauss	0.4036	0.4034

Tabela 2 – Resultado da média teórica e amostral para a distribuição Kummer beta com parâmetros $a = 0.9$, $b = 6$, $c = -7$ e a distribuição Hipergeométrica de Gauss com parâmetros $a = 3$, $b = 4$, $c = 2$, $d = 0.5$, para diferentes tamanhos de amostras.

Pode-se notar que à medida que o tamanho amostral aumenta, a média amostral tende a se aproximar da média teórica. Esse fenômeno é consistente com a Lei dos Grandes Números, que afirma que a média amostral converge para a média teórica à medida que o tamanho da amostra cresce.

Tamanho da amostra	Distribuição	Média teórica	Média amostral
$m = 100$	Kummer beta	0.3893	0.4182
	H. de Gauss	0.6349	0.6623
$m = 250$	Kummer beta	0.3893	0.4036
	H. de Gauss	0.6349	0.6528
$m = 500$	Kummer beta	0.3893	0.3958
	H. de Gauss	0.6349	0.6507
$m = 1000$	Kummer beta	0.3893	0.3859
	H. de Gauss	0.6349	0.6322

Tabela 3 – Resultado da média teórica e amostral para a distribuição Kummer beta com parâmetros $a = 0.7$, $b = 5$, $c = -8$, e a distribuição Hipergeométrica de Gauss com parâmetros $a = 4$, $b = 3$, $c = 1$, $d = -0.9$, para diferentes tamanhos de amostras.

Para amostras pequenas, como $m = 100$, a média amostral pode estar mais distante da média teórica. Para amostras maiores, como $m = 500$ ou $m = 1000$, a média amostral se aproxima cada vez mais da média teórica. A precisão da estimativa da média amostral melhora com o aumento de m , o que é esperado.

Este exemplo ilustra bem o comportamento esperado das médias amostrais para distribuições diferentes, ajudando a comparar a média teórica com a média obtida de amostras aleatórias.

Além disso, a seguir, será apresentada uma análise gráfica, por meio de histograma e da linha da densidade da FDP sobreposta aos dados, construídos no *software R*.

Inicialmente, foram gerados números aleatórios para a distribuição Kummer beta com parâmetros $a = 0.1$, $b = 2$ e $c = -8$, sendo tomadas amostras de tamanhos $m = 100$, 250, 500 e 1000, conforme apresentadas nas figuras 6a, 6b, 6c e 6d, respectivamente.

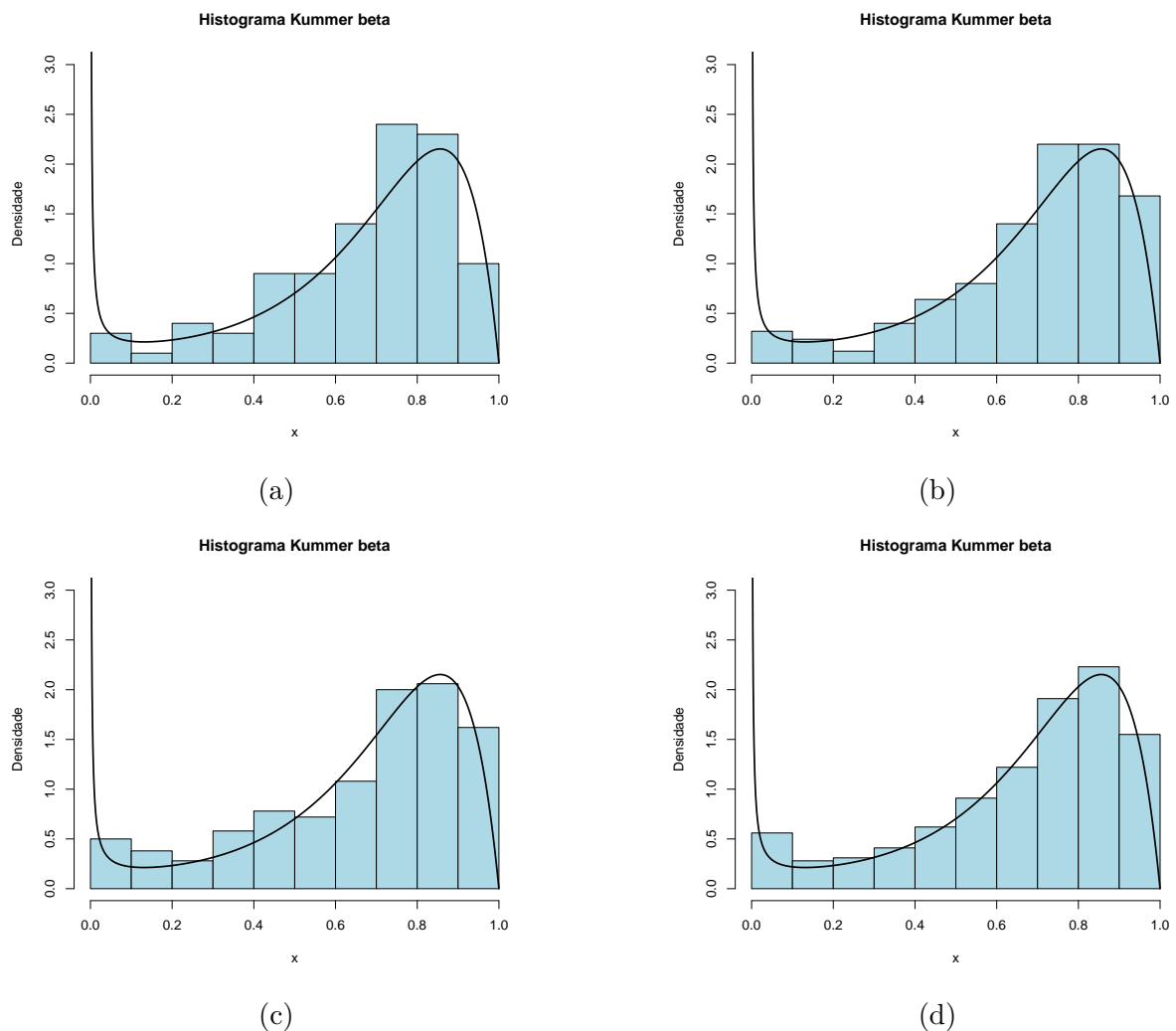


Figura 6 – Histogramas da distribuição Kummer beta com parâmetros $a = 0.1$, $b = 2$ e $c = -8$, para tamanhos de amostra m iguais a: (a) 100, (b) 250, (c) 500, (d) 1000.

Nota-se que à medida que o tamanho da amostra aumenta, a sobreposição da linha da densidade melhor se ajusta aos dados observados, sugerindo que o modelo teórico é um bom ajuste para os dados, pois não há áreas onde a linha da densidade se desvia significativamente do histograma, o que poderia indicar que o modelo não captura bem a verdadeira distribuição dos dados.

Na sequência, foram gerados números aleatórios para a distribuição hipergeométrica de Gauss com parâmetros $a = 2$, $b = 2$ e $c = 2$ e $d = 0.5$, sendo tomadas amostras de

tamanhos $m = 100, 250, 500$ e 1000 , conforme apresentadas nas figuras 7a, 7b, 7c e 7d, respectivamente.

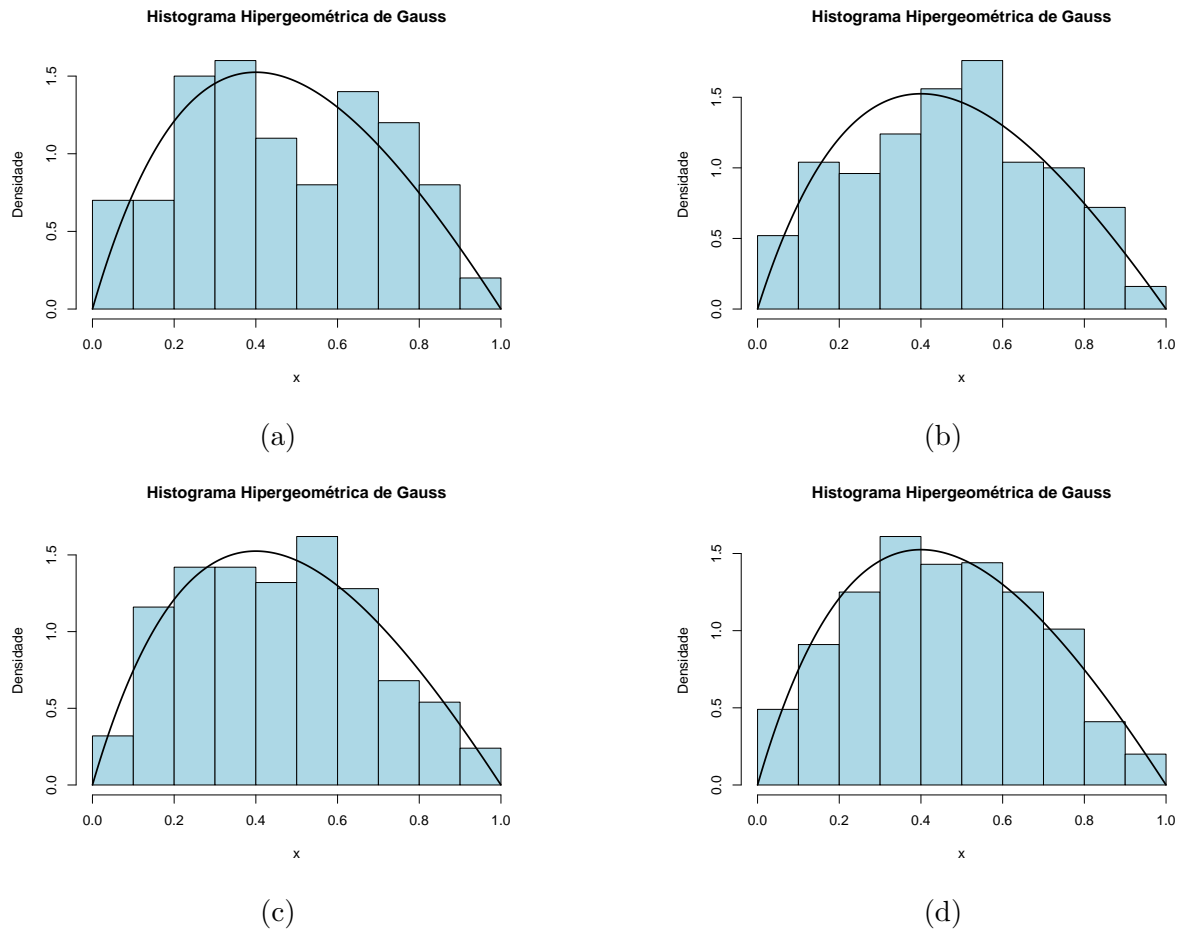


Figura 7 – Histogramas da distribuição Hipergeométrica de Gauss com parâmetros $a = 2, b = 2, c = 2$ e $d = 0.5$, para tamanhos de amostra m iguais a: (a) 100, (b) 250, (c) 500, (d) 1000.

Conforme visto nas figuras acima, a linha da densidade da distribuição Hipergeométrica de Gauss sobrepondo o histograma, proporcionou uma visão clara da distribuição dos dados, adequando-se ao modelo teórico.

A geração de números aleatórios realizadas no R para os casos da Kummer beta e Hipergeométrica de Gauss são tomadas aqui como exemplo, mas podem ser estendidas para as demais distribuições das generalizações de beta.

5 Aplicação com dados simulados

Neste capítulo serão apresentados três métodos, sendo estes o Critério de Informação de Akaike, o Critério de Informação de Bayes, e o Critério de Informação Deviance, que buscam comparar modelos concorrentes e selecionar o melhor entre eles, sendo expostos os resultados obtidos a partir de simulação que busca estimar a probabilidade de sucesso π em um modelo binomial, baseado em dados simulados que quantificam o número de peças defeituosas em um lote de peças.

5.1 Critério de Informação de Akaike (AIC)

O AIC é uma ferramenta estatística amplamente utilizada para comparar e selecionar modelos probabilísticos, foi introduzido por Akaike (1973) e objetiva fornecer uma maneira quantitativa de avaliar a qualidade de um modelo, levando em consideração tanto o ajuste do modelo aos dados quanto a complexidade do modelo.

O AIC é dado pela fórmula:

$$\text{AIC} = -2 \ln(L(\hat{\pi}|y)) + 2k,$$

em que $L(\hat{\pi}|y)$ é a Estimativa de Máxima Verossimilhança do modelo e k é o número de parâmetros. O AIC funciona em uma comparação relativa entre modelos diferentes, em que o objetivo não é calcular um “valor absoluto” de quão bom um modelo é, mas comparar modelos concorrentes e selecionar o melhor entre eles. O modelo com o menor valor de AIC é preferido, pois indica que ele oferece o melhor equilíbrio entre a qualidade do ajuste e a simplicidade do modelo.

5.2 Critério de Informação Bayesiano (BIC)

O BIC, proposto por Schwarz (1978), é um dos critérios mais utilizados de comparação de modelos, semelhante ao AIC, mas com uma penalização mais forte para modelos com muitos parâmetros. Ele ajuda a identificar o modelo mais apropriado entre um conjunto de modelos concorrentes, levando em consideração a qualidade do ajuste, e também, a complexidade do modelo, penalizando o número de parâmetros de acordo com o tamanho da amostra.

O BIC é dado pela fórmula:

$$\text{BIC} = -2 \ln(L(\hat{\pi}|y)) + k \ln(n),$$

em que $L(\hat{\pi}|y)$ é a Estimativa de Máxima Verossimilhança do modelo, k é o número de parâmetros, e n é o número de observações. A utilidade do BIC aparece quando você está comparando vários modelos para os mesmos dados. O modelo com o menor valor de BIC será considerado o melhor.

5.3 Critério de Informação Deviance (DIC)

O DIC, desenvolvido por Spiegelhalter et al. (2002) e colaboradores, é um critério para comparar modelos bayesianos, balanceando qualidade de ajuste com complexidade do modelo. Ele é uma generalização bayesiana de critérios como AIC e BIC, mas leva em conta a incerteza dos parâmetros por meio da distribuição a posteriori.

O DIC é dado pela fórmula:

$$\text{DIC} = 2\bar{D} - D(\bar{\pi}),$$

sendo \bar{D} a média da deviance sobre a distribuição a posteriori dos parâmetros, isto é, $\bar{D} = \mathbb{E}[-2\log(f(y|\pi))]$, e $D(\bar{\theta})$ a deviance no valor médio dos parâmetros, ou seja, $D(\bar{\theta}) = \mathbb{E}[-2\log(f(y|\bar{\pi}))]$. O DIC também busca o equilíbrio entre ajuste e complexidade, sendo o menor valor de DIC o melhor.

5.4 Simulações

As distribuições beta, Kummer beta e Hipergeométrica de Gauss possuem fórmula fechada para a média. Nesse sentido, pode-se realizar o cálculo da média da distribuição a posteriori, que será a estimativa de π (probabilidade de sucesso) após observar os dados.

Seja X uma variável aleatória que expressa o número de peças defeituosas observadas em uma amostra e n o total de peças observadas. Foi realizada uma aplicação com dados simulados para diferentes valores de peças defeituosas em amostras de tamanhos variados.

Sabe-se que a distribuição a priori reflete nosso conhecimento ou suposições iniciais sobre π , antes de observar qualquer dado. A distribuição a posteriori combina essa informação com os dados observacionais (o número de peças defeituosas), resultando em uma estimativa mais refinada e adaptada às novas evidências.

A estimativa bayesiana de π , que é a média da distribuição posterior de π , fornece uma estimativa de probabilidade (a probabilidade de um defeito em uma peça) que é atualizada à medida que novos dados são observados. Esta estimativa pode ser usada para calcular a verossimilhança do modelo, que é uma parte essencial para o cálculo do AIC e BIC.

As simulações permitiram realizar uma análise estatística bayesiana para estimar a probabilidade de defeitos em um conjunto de peças, utilizando uma abordagem com

distribuição a priori Kummer beta e Hipergeométrica de Gauss, e distribuição binomial para modelar os dados, em comparação com a distribuição $Beta(1, 1)$. A seguir, tem-se os dados obtidos pelas simulações realizadas no *R*.

Total de peças	Defeituosas	Distribuição	Estimativa de π	AIC	BIC	DIC
30	8	Beta	0.2812	5.6619	7.0631	3.6619
		Kummer beta	0.2741	5.6385	7.0397	3.6492
		H. de Gauss	0.2615	5.6340	7.0352	3.6340
50	6	Beta	0.1346	5.6246	7.5366	3.6246
		Kummer beta	0.1302	5.5778	7.4898	3.5778
		H. de Gauss	0.1318	5.5930	7.5050	3.5930
70	12	Beta	0.1805	6.1890	8.4375	4.1890
		Kummer beta	0.1743	6.1531	8.4016	4.1531
		H. de Gauss	0.1760	6.1593	8.4078	4.1593
90	9	Beta	0.1086	6.0204	8.5202	4.0204
		Kummer beta	0.1075	6.0027	8.5025	4.0027
		H. de Gauss	0.1082	6.0130	8.5128	4.0130

Tabela 4 – Resultado para a distribuição Beta com parâmetros $\alpha = 1$, $\beta = 1$, Kummer beta com parâmetros $a = 1.5$, $b = 3.5$, $c = 3$ e a distribuição Hipergeométrica de Gauss com parâmetros $a = 1.5$, $b = 3$, $c = 3.5$, $d = 0.9$, para diferentes tamanhos de amostras.

Total de peças	Defeituosas	Distribuição	Estimativa de π	AIC	BIC	DIC
40	2	Beta	0.0714	4.8698	6.5587	2.8698
		Kummer beta	0.0679	4.7847	6.4736	2.7847
		H. de Gauss	0.0708	4.8559	6.5448	2.8559
60	7	Beta	0.1290	5.7678	7.8622	3.7678
		Kummer beta	0.1192	5.6877	7.7820	3.6877
		H. de Gauss	0.1271	5.7451	7.8395	3.7451
80	16	Beta	0.2073	6.4245	8.8065	4.4245
		Kummer beta	0.1925	6.4266	8.8086	4.4266
		H. de Gauss	0.2044	6.4082	8.7902	4.4082
100	7	Beta	0.0784	5.8368	8.4420	3.8368
		Kummer beta	0.0762	5.7913	8.3964	3.7913
		H. de Gauss	0.0780	5.8289	8.4340	3.8289

Tabela 5 – Resultado para a distribuição Beta com parâmetros $\alpha = 1$, $\beta = 1$, Kummer beta com parâmetros $a = 1.5$, $b = 5$, $c = 5.5$ e a distribuição Hipergeométrica de Gauss com parâmetros $a = 1.1$, $b = 2$, $c = 1.8$, $d = 0.4$, para diferentes tamanhos de amostras.

A escolha dos parâmetros apresentados nas duas tabelas acima, pode ser uma parte válida de um processo exploratório na modelagem estatística. Este tipo de abordagem ajuda a identificar rapidamente uma gama de possíveis configurações para os modelos, permitindo uma visão inicial sobre quais parâmetros podem ser mais adequados sem um compromisso profundo com modelos específicos desde o início, com isso, ao testar diferentes combinações de parâmetros, você pode selecionar aqueles que geram resultados equilibrados em termos de ajuste aos dados e desempenho em critérios como o AIC, o BIC e o DIC.

5.5 Resultados e discussões

Inicialmente, após testar generalizações da distribuição beta como priori conjugada para o parâmetro π da distribuição binomial, sendo encontradas diversas distribuições

abordadas neste trabalho, fizemos um estudo mais detalhado para a Kummer beta e Hipergeométrica de Gauss, com intuito de avaliar como essas generalizações de beta se comparam com a distribuição beta usual em termos de ajuste.

Continuando, após a aplicação com dados simulados, nota-se, pelos resultados obtidos (expostos nas tabelas da seção anterior), que as distribuições Kummer beta e Hipergeométrica de Gauss apresentam tanto os valores de AIC e BIC, como também os valores de DIC, de modo geral, menores do que os valores de AIC, BIC e DIC da distribuição beta usual, isso sugere que a distribuição beta, no contexto dessa aplicação, não é a mais adequada para descrever os dados. Em termos da estimativa do parâmetro π , indica que a suposição de que os dados seguem uma distribuição beta pode ser inadequada, e que as outras distribuições fornecem uma explicação mais consistente dos padrões observados nos dados.

Portanto, conclui-se que modelos alternativos, como os casos abordados, devem ser considerados para a descrição mais precisa dos dados e para a estimativa do parâmetro π em questão. Optar por modelos mais apropriados, em contextos de controle de qualidade e gestão de produção, pode ajudar a tomar decisões mais informadas, como determinar a necessidade de ajustes no processo de fabricação, planejar inspeções ou definir limites de aceitação, visto que a estimativa bayesiana oferece uma maneira de quantificar a probabilidade de defeitos levando em conta tanto o conhecimento prévio quanto as evidências empíricas.

5.6 Gráficos da distribuição a posteriori

A seguir, serão apresentados os gráficos da distribuição a posteriori de π (ou p), cuja linha na cor azul representa a densidade de probabilidade e a linha vermelha tracejada indica a estimativa de π , que é a média da distribuição a posteriori.

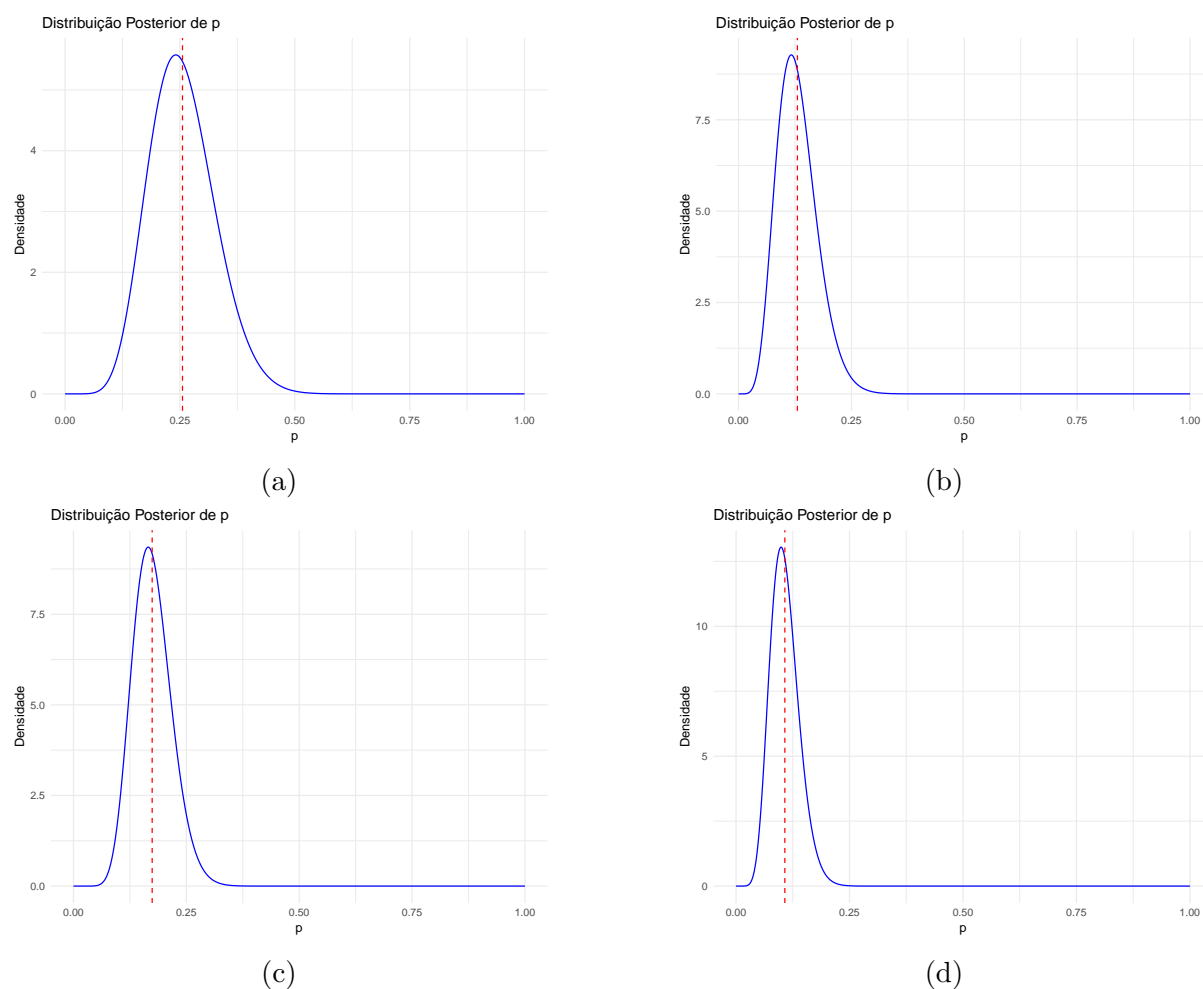


Figura 8 – Gráficos da posteriori Kummer beta com parâmetros $a = 1.5$, $b = 3.5$ e $c = 3$, para total de peças e número de defeitos iguais a: (a) 30 e 8, (b) 50 e 6, (c) 70 e 12, (d) 90 e 9, respectivamente.

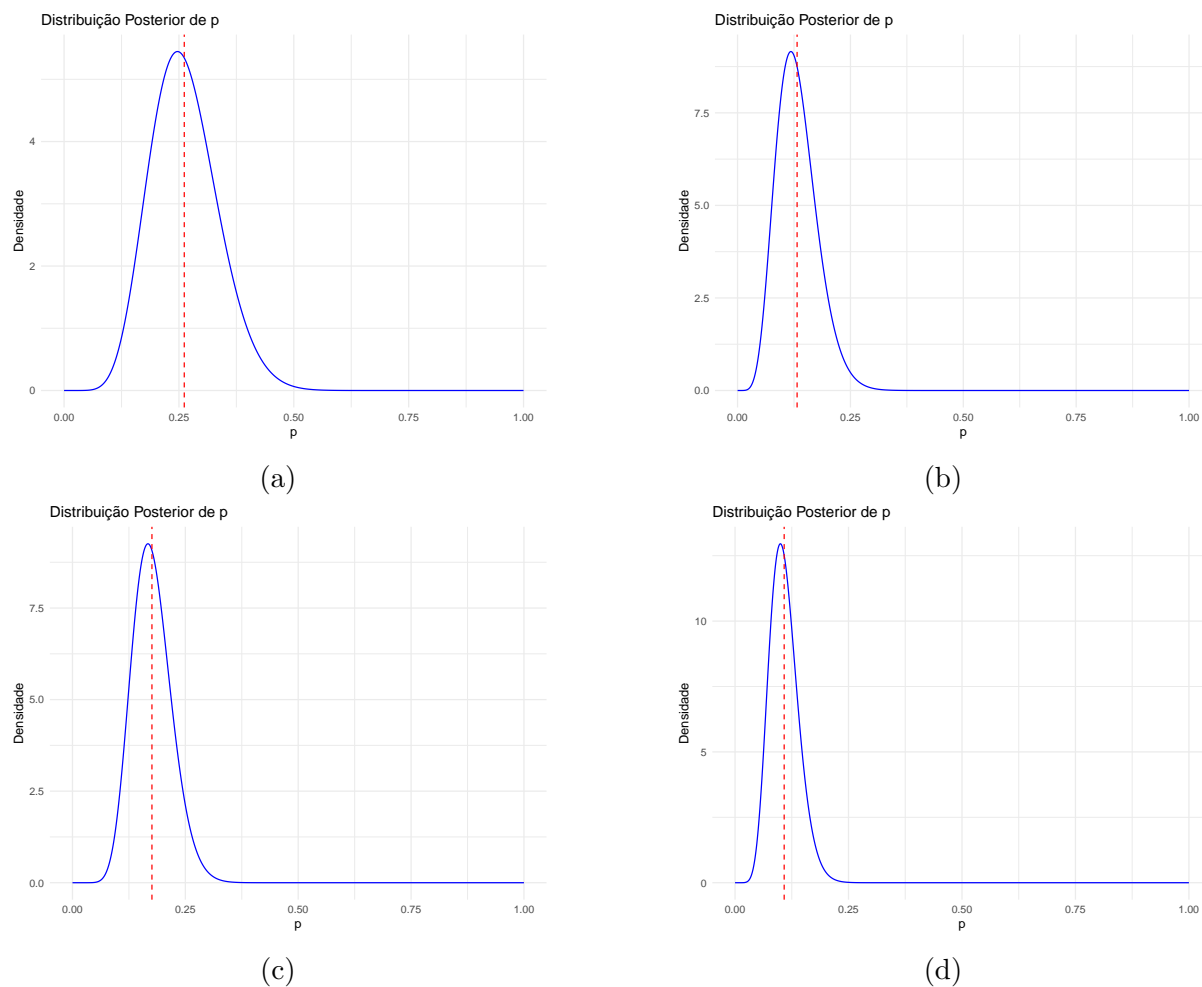


Figura 9 – Gráficos da posteriori Hipergeométrica de Gauss com parâmetros $a = 1.5$, $b = 3$, $c = 3.5$, $d = 0.9$, para total de peças e número de defeitos iguais a: (a) 30 e 8, (b) 50 e 6, (c) 70 e 12, (d) 90 e 9, respectivamente.

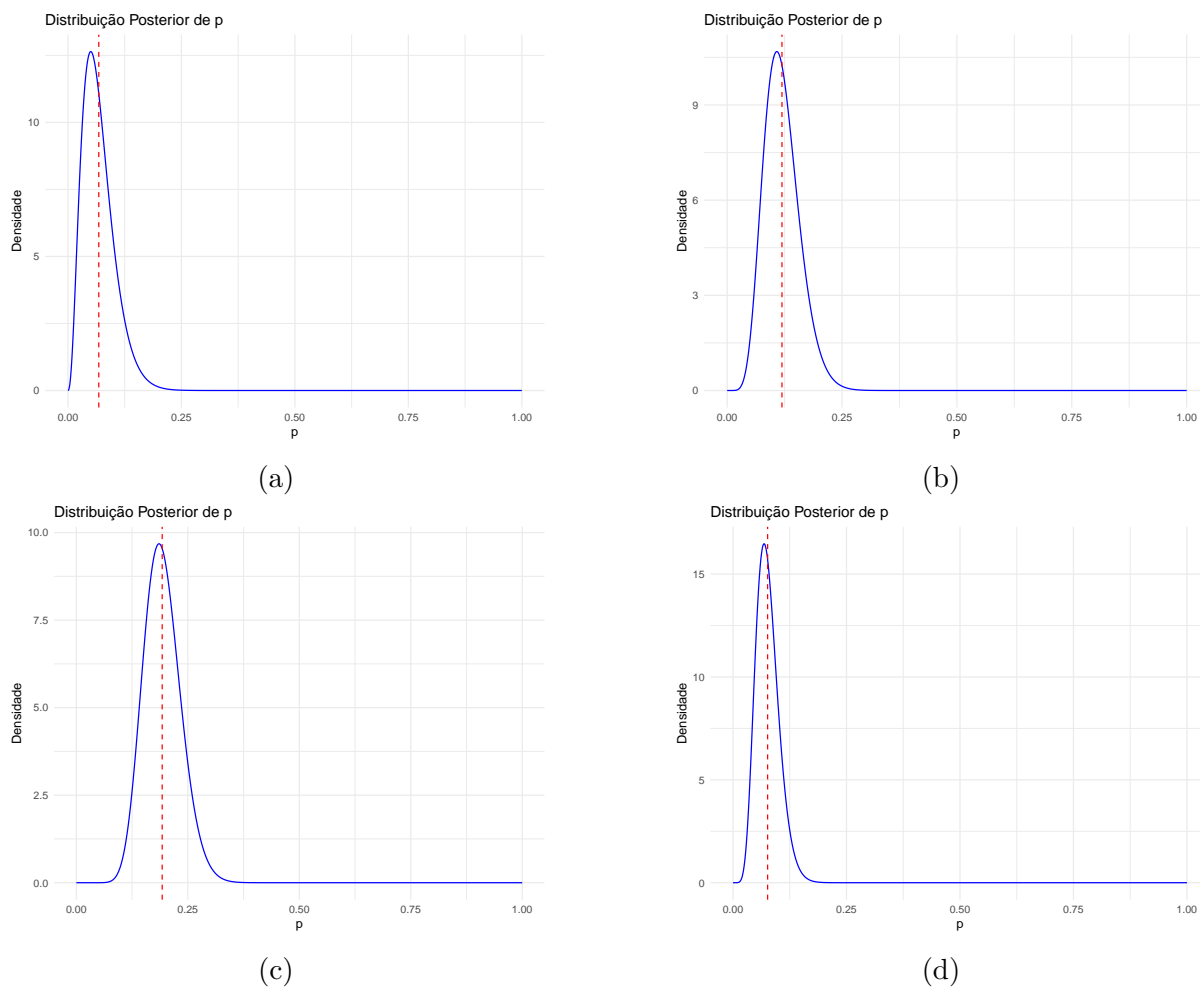
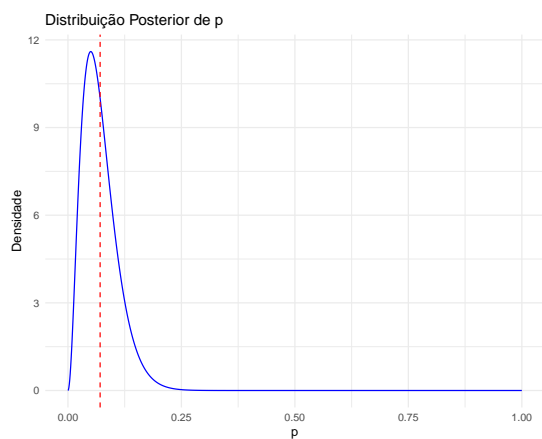
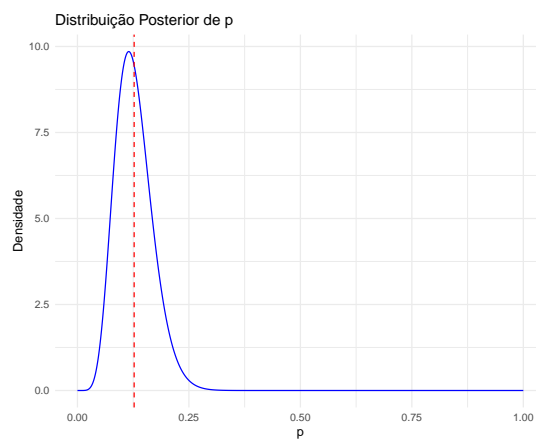


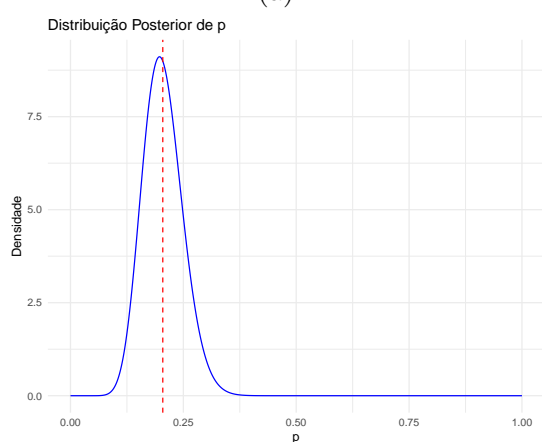
Figura 10 – Gráficos da posteriori Kummer beta com parâmetros $a = 1.5$, $b = 5$ e $c = 5.5$, para total de peças e número de defeitos iguais a: (a) 40 e 2, (b) 60 e 7, (c) 80 e 16, (d) 100 e 7, respectivamente.



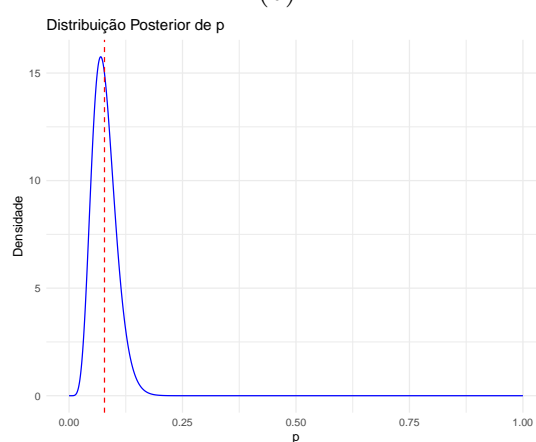
(a)



(b)



(c)



(d)

Figura 11 – Gráficos da posteriori Hipergeométrica de Gauss com parâmetros $a = 1.1$, $b = 2$, $c = 1.8$, $d = 0.4$, para total de peças e número de defeitos iguais a: (a) 40 e 2, (b) 60 e 7, (c) 80 e 16, (d) 100 e 7, respectivamente.

6 Considerações finais

No presente trabalho, a escolha da distribuição beta como priori conjugada para o parâmetro π da distribuição binomial é discutida, destacando suas limitações e a necessidade de alternativas, sendo propostas o uso de generalizações da distribuição beta como priori para o parâmetro π da distribuição binomial. Diante de tal motivação, testamos generalizações de beta, comprovando, via Teorema de Bayes, que tais generalizações são, de fato, conjugadas para o parâmetro π da distribuição binomial.

Após a apresentação das generalizações de beta, propôs-se um estudo mais detalhado da distribuição Kummer beta e da Hipergeométrica de Gauss, podendo ser estendido às demais generalizações de forma equivalente. Para isso, realizamos o processo de geração de números aleatórios, que permitiu estabelecer um quadro comparativo entre as médias teóricas e amostrais obtidas dessas distribuições, além da construção de histogramas que proporcionou uma visão clara da distribuição dos dados adequando-se ao modelo teórico.

Conforme mencionado, com a geração de números aleatórios mostrando-se satisfatória, foi entendido que o modelo se adequava bem aos dados, podendo ser realizada a etapa de aplicação. A partir de dados simulados no contexto de controle de qualidade e gestão de produção de peças, foram realizadas simulações que buscam estimar a probabilidade de sucesso π no modelo binomial, baseado em dados que quantificavam o número de peças defeituosas em um lote de peças.

Além disso, foram apresentados o AIC, o BIC, e o DIC, que permitiram comparar a distribuição Kummer beta e Hipergeométrica de Gauss com a distribuição beta usual. A partir de os critérios testados, a análise das simulações realizadas, nos permite inferir que as distribuições Kummer beta e Hipergeométrica de Gauss mostraram melhores resultados, quando comparadas à distribuição beta, em termos de ajuste.

Portanto, os modelos alternativos, isto é, a distribuição Kummer beta e a Hipergeométrica de Gauss devem ser consideradas para uma descrição mais precisa dos dados e para a estimativa do parâmetro π do modelo binomial. Esse resultado permite inferir que essas generalizações de beta ampliam a flexibilidade da distribuição beta usual, sendo estas mais adequadas para modelar certos conjuntos de dados que apresentam características específicas que não são bem capturadas pela distribuição beta.

Referências Bibliográficas

- ABRAMOWITZ, M.; STEGUN, I. A. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. [S.l.]: Dover Publications, 1968. Citado na página 22.
- AKAIKE, H. Information theory and an extension of the maximum likelihood principle. *Akademiai Kiado*, p. 267–281, 1973. Citado na página 42.
- ARMERO, C.; BAYARRI, M. J. Prior assessments for prediction in queues. *The Statistician*, v. 43, p. 139–153, 1994. Citado 2 vezes nas páginas 23 e 28.
- BAYES, T. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, v. 53, p. 370–418, 1763. Citado 2 vezes nas páginas 13 e 16.
- BERGER, J. O. *Statistical Decision Theory and Bayesian Analysis*. [S.l.]: Springer, 1985. Citado na página 13.
- BOLSTAD, W. M. Introduction to bayesian statistics. In: _____. [S.l.]: John Wiley & Sons, 2004. cap. 8. Citado na página 19.
- BOX, G. E. P.; TIAO, G. *Bayesian inference in statistical analysis*. [S.l.]: Reading Addison-Wesley, 1973. Citado na página 16.
- CASELLA, G.; BERGER, R. L. *Statistical inference*. Duxbury Press, 2002. Citado na página 18.
- CHAMAYOU, J.; LETAC, G. Explicit stationary distributions for compositions of random functions and products of random matrices. *J. Theoret. Probab.*, p. 3–36, 1991. Citado na página 28.
- CHEN, M. H.; SHAO, Q. M.; IBRAHIM, J. G. Monte carlo methods in bayesian computation. *Journal of the American Statistical Association*, 2004. Citado na página 13.
- DAVISON, A. C. *Statistical models*. Cambridge University Press, 2003. Citado na página 13.
- DEGROOT, M. H.; SCHERVISH, M. J. *Probability and Statistics*. 4. ed. [S.l.]: Pearson, 2012. Citado na página 16.
- DIEBOLT, J.; ROBERT, C. P. Estimation of finite mixtures of distributions through bayesian sampling. *Journal of the Royal Statistical Society*, v. 56, p. 363–375, 1994. Citado na página 38.
- GAMERMAN, D.; MIGON, H. S. *Inferencia estatística: Uma abordagem integrada*. Instituto de Matemática, 1993. Citado 2 vezes nas páginas 16 e 37.
- GELMAN, A. et al. *Bayesian data analysis*. CRC Press, p. 17, 2013. Citado na página 20.
- GELMAN, A. et al. *Bayesian data analysis*. [S.l.]: CRC Press, 2014. Citado na página 13.

- GELMAN, A. et al. *Bayesian Data Analysis*. [S.l.]: Chapman and Hall/CRC, 2014. Citado 2 vezes nas páginas 14 e 18.
- GELMAN, A.; HILL, J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. [S.l.]: Cambridge University Press, 2007. Citado na página 30.
- GENTLE, J. E. *Random Number Generation and Monte Carlo Methods*. [S.l.]: Springer, 2003. Citado na página 37.
- GEWEKE, J. *Bayesian Econometrics*. [S.l.]: Oxford University Press, 1999. Citado na página 13.
- GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. Markov chain monte carlo in practice. *Chapman and Hall*, 1996. Citado na página 13.
- GORDY, M. B. Computationally convenient distributional assumptions for common-value auctions. *Comput. Econ.*, v. 12, p. 61–78, 1998. Citado 3 vezes nas páginas 22, 23 e 29.
- GRADSHTEYN, I. S.; RYZHIK, I. M. *Table of integrals, series, and products*. [S.l.: s.n.], 2000. Citado 2 vezes nas páginas 27 e 28.
- GUPTA, A. K.; NADARAJAH, S. Beta bessel distributions. *International Journal of Mathematics and Mathematical Sciences*, v. 2006, p. 1–14, 2006. Citado 2 vezes nas páginas 25 e 26.
- GUPTA, A. K.; NAGAR, D. K. Matrix-variate gauss hypergeometric distribution. *Journal of the Australian Mathematical Society*, p. 335–355, 2012. Citado na página 23.
- HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, v. 57, p. 97–109, 1970. Citado na página 38.
- HOGG, R. V.; TANIS, E. A. Probability and statistical inference. *Pearson*, 2015. Citado na página 19.
- HOOTEN, M. B.; HOBBS, N. T. A guide to bayesian model selection for ecologists. *Ecological Monographs*, 2015. Citado na página 13.
- JAMES, J.; MELVIN, R. N. Bayesian analysis for binomial models with generalized beta prior distributions. *Journal of Educational Statistics*, v. 9, n. 2, p. 163–175, 1984. Citado na página 21.
- JAYNES, E. T. *Probability theory: The logic of science*. [S.l.]: Cambridge University Press, 2003. Citado na página 13.
- JEFFREYS, H. *Theory of Probability*. [S.l.: s.n.], 1939. v. 3. Citado 2 vezes nas páginas 13 e 16.
- KOUDOU, E.; VALLOIS, P. Which distributions have the matsumoto-yor property? *Electron. Comm. Probab.*, p. 556–566, 2011. Citado na página 28.
- KROESE, D. P.; TAIMRE, T.; BOTEV, Z. I. Handbook for monte carlo methods. *Wiley Series in Probability and Statistics*, 2011. Citado na página 37.
- LEE, P. M. *Bayesian Statistics: An Introduction*. 4. ed. [S.l.]: Wiley, 2012. Citado na página 13.

- LIBBY, D. L.; NOVICK, M. R. Multivariate generalized beta-distributions with applicationsto utility assessment. *Journal of Educational Statistics*, p. 271–294, 1982. Citado na página 23.
- MARINHO, P. R. D. *Método de aceitação-rejeição*. 2024. Disponível em: <<https://prdm0.github.io/AcceptReject/reference/index.html>>. Citado na página 38.
- METROPOLIS, N. et al. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, v. 21, p. 1087–1092, 1953. Citado na página 38.
- NADARAJAH, S. The bivariate gauss hypergeometric beta distribution. *Integral Transforms and Special Functions*, p. 859–868, 2008. Citado na página 23.
- NADARAJAH, S.; KOTZ, S. Beta trigonometric distributions. *Portuguese Economic Journal*, v. 5, p. 207–224, 2006. Citado 2 vezes nas páginas 27 e 28.
- NADARAJAH, S.; KOTZ, S. An f_1 beta distribution with bathtub failure rate function. *American Journal of Mathematical and Management Sciences*, v. 26, p. 113–131, 2006. Citado na página 24.
- NADARAJAH, S.; KOTZ, S. Multitude of beta distributions with applications. *Statistics*, v. 41, p. 153–179, 2007. Citado na página 21.
- NADARAJAH, S.; KOTZ, S. Two generalized beta distributions. *Applied Economics*, p. 1743–1751, 2007. Citado na página 29.
- NAGAR, D. K.; BEDOYA-VALENCIA, D.; NADARAJAH, S. Multivariate generalization of the gauss hypergeometric distribution. *Hacettepe Journal of Mathematics and Statistics*, p. 933–948, 2015. Citado na página 23.
- NAGAR, D. K.; VALENCIA, D. B. Product and quotient of independent gauss hypergeometric variables. *Ingeniería y Ciencia*, p. 29–48, 2011. Citado na página 23.
- NASCIMENTO, F. Modelos probabilísticos para dados extremos: Teoria e aplicações. 2012. Citado na página 17.
- NEUMANN, J. von. Various techniques used in connection with random digits. *Journal of Research of the National Institute of Standards and Technology*, v. 3, p. 36–38, 1951. Citado na página 37.
- NG, K. W.; KOTZ, S. Kummer-gamma and kummer-beta univariate and multivariate distributions. *Research Report*, v. 84, 1995. Citado 2 vezes nas páginas 22 e 29.
- O’HAGAN, A. *Advanced Theory of Statistics: Bayesian Inference*. [S.l.: s.n.], 1994. v. 2. Citado na página 18.
- PRESS, W. H. et al. *Numerical Recipes: The Art of Scientific Computing*. [S.l.]: Cambridge University Press, 2007. Citado na página 37.
- PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O. I. Integrals and series. *Gordon and Breach Science Publishers*, 1986. Citado na página 23.
- ROBERT, C. P. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. [S.l.]: Springer, 2007. Citado na página 13.

ROBERT, C. P.; CASELLA, G. *Monte Carlo statistical methods*. [S.l.]: Springer, 2004. Citado na página 37.

RÉNYI, A. On measures of information and entropy. *University of California Press*, p. 547–561, 1960. Citado na página 23.

SCHWARZ, G. Estimating the dimension of a model. *Annals of Statistics*, v. 6, p. 461–464, 1978. Citado na página 42.

SHANNON, C. E. A mathematical theory of communication. *Bell System Technical Journal*, p. 379–423, 1948. Citado na página 23.

SPIEGELHALTER, D. J.; ABRAMS, K. R.; MYLES, J. P. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. [S.l.]: Wiley, 1999. Citado na página 13.

SPIEGELHALTER, D. J. et al. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, v. 64, p. 583–639, 2002. Citado na página 43.

STIGLER, S. M. *The history of statistics: The measurement of uncertainty before 1900*. *Harvard University Press*, 1986. Citado na página 13.

ZELLNER, A. *An Introduction to Bayesian Inference in Econometrics*. [S.l.]: John Wiley Sons, 1971. Citado na página 13.

Apêndices

APÊNDICE A – Cálculo para obtenção da posteriori

Segue as generalizações da distribuição beta como priori conjugada para o modelo binomial.

A.0.1 Priori Kummer beta

Seja a priori Kummer beta que possui FDP dada por:

$$g(\pi) = K(a, b, c) \pi^{a-1} (1 - \pi)^{b-1} \exp(-c \pi).$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = K(a, b, c) \pi^{a-1} (1 - \pi)^{b-1} \exp(-c \pi) \times \binom{n}{y} \pi^y (1 - \pi)^{n-y} \\ &= \binom{n}{y} K(a, b, c) \pi^{a+y-1} (1 - \pi)^{n-y+b-1} \exp(-c \pi) \\ &\propto \pi^{a+y-1} (1 - \pi)^{n-y+b-1} \exp(-c \pi), \end{aligned}$$

é a posteriori, a qual é outra distribuição Kummer beta, com parâmetros $a' = a + y$, $b' = n + b - y$ e c .

A.0.2 Priori Hipergeométrica de Gauss

Seja a priori Hipergeométrica de Gauss com FDP dada por:

$$g(\pi) = \frac{K \pi^{a-1} (1 - \pi)^{b-1}}{(1 + d\pi)^c}.$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = \frac{K \pi^{a-1} (1 - \pi)^{b-1}}{(1 + d\pi)^c} \times \binom{n}{y} \pi^y (1 - \pi)^{n-y} \\ &= \binom{n}{y} \frac{K \pi^{y+a-1} (1 - \pi)^{n-y+b-1}}{(1 + d\pi)^c} \\ &\propto \frac{\pi^{y+a-1} (1 - \pi)^{n-y+b-1}}{(1 + d\pi)^c}, \end{aligned}$$

é a posteriori, a qual é outra distribuição Hipergeométrica de Gauss, com parâmetros $a' = a + y$, $b' = n + b - y$, c e d .

A.0.3 Priori Apell beta

Seja a priori Apell beta com FDP dada por:

$$g(\pi) = \frac{C\pi^{\alpha-1}(1-\pi)^{\beta-1}}{(1-u\pi)^\rho(1-v\pi)^\lambda}.$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = \frac{C\pi^{\alpha-1}(1-\pi)^{\beta-1}}{(1-u\pi)^\rho(1-v\pi)^\lambda} \times \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \binom{n}{y} \frac{C\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}}{(1-u\pi)^\rho(1-v\pi)^\lambda} \\ &\propto \frac{\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}}{(1-u\pi)^\rho(1-v\pi)^\lambda}, \end{aligned}$$

é a posteriori, a qual é outra distribuição Apell beta, com parâmetros $\alpha' = \alpha + y$, $\beta' = n + \beta - y$, ρ , λ , u e v .

A.0.4 Priori BB I

Seja a priori BB I com FDP dada por:

$$g(\pi) = C\pi^{\alpha-1}(1-\pi)^{\beta-1}I_\nu(c\pi).$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = C\pi^{\alpha-1}(1-\pi)^{\beta-1}I_\nu(c\pi) \times \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \binom{n}{y} C\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}I_\nu(c\pi) \\ &\propto \pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}I_\nu(c\pi), \end{aligned}$$

é a posteriori, a qual é outra distribuição BB I, com parâmetros $\alpha' = \alpha + y$, $\beta' = n + \beta - y$, c e ν .

A.0.5 Priori BB II

Seja a priori BB II com FDP dada por:

$$f(\pi) = C\pi^{\alpha-1}(1-\pi)^{\beta-1} \exp(c\pi)I_\nu(c\pi).$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = C\pi^{\alpha-1}(1-\pi)^{\beta-1} \exp(c\pi)I_\nu(c\pi) \times \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \binom{n}{y} C\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1} \exp(c\pi)I_\nu(c\pi) \\ &\propto \pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1} \exp(c\pi)I_\nu(c\pi), \end{aligned}$$

é a posteriori, a qual é outra distribuição BB II, com parâmetros $\alpha' = \alpha + y$, $\beta' = n + \beta - y$, c e ν .

A.0.6 Priori BB III

Seja a priori BB III com FDP dada por:

$$g(\pi) = C\pi^{\alpha-1}(1-\pi)^{\beta-1}\exp(-c\pi)I_{\nu}(c\pi).$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = C\pi^{\alpha-1}(1-\pi)^{\beta-1}\exp(-c\pi)I_{\nu}(c\pi) \times \binom{n}{y}\pi^y(1-\pi)^{n-y} \\ &= \binom{n}{y}C\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}\exp(-c\pi)I_{\nu}(c\pi) \\ &\propto \pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}\exp(c\pi)I_{\nu}(-c\pi), \end{aligned}$$

é a posteriori, a qual é outra distribuição BB III, com parâmetros $\alpha' = \alpha + y$, $\beta' = n + \beta - y$, c e ν .

A.0.7 Priori BT Cos I

Seja a priori BT Cos I com FDP dada por:

$$g(\pi) = C\pi^{\nu-1}(1-\pi)^{\mu-1}\cos(a\pi).$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = C\pi^{\nu-1}(1-\pi)^{\mu-1}\cos(a\pi) \times \binom{n}{y}\pi^y(1-\pi)^{n-y} \\ &= \binom{n}{y}C\pi^{y+\nu-1}(1-\pi)^{n-y+\mu-1}\cos(a\pi) \\ &\propto \pi^{y+\nu-1}(1-\pi)^{n-y+\mu-1}\cos(a\pi), \end{aligned}$$

é a posteriori, a qual é outra distribuição BT Cos I, com os parâmetros $\nu' = \nu + y$, $\mu' = n + \mu - y$ and a .

A.0.8 Priori BT Sen I

Seja a priori BT Sen I com FDP dada por:

$$g(\pi) = C\pi^{\nu-1}(1-\pi)^{\mu-1}\text{sen}(a\pi).$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = C\pi^{\nu-1}(1-\pi)^{\mu-1} \text{sen}(a\pi) \times \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \binom{n}{y} C\pi^{y+\nu-1}(1-\pi)^{n-y+\mu-1} \text{sen}(a\pi) \\ &\propto \pi^{y+\nu-1}(1-\pi)^{n-y+\mu-1} \text{sen}(a\pi), \end{aligned}$$

é a posteriori, a qual é outra distribuição BT Sen I, com os parâmetros $\nu' = y + \nu$, $\mu' = n - y + \mu$ e a .

A.0.9 Priori Kummer do tipo II

Seja a priori Kummer do tipo II com FDP dada por:

$$g(\pi) = N_{\delta}(a, b, c)\pi^{a-1}(1-\pi)^{b-1}(1+(\delta-1)\pi)^c I_{(0,1)}(\pi).$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) \\ &= N_{\delta}(a, b, c)\pi^{a-1}(1-\pi)^{b-1}(1+(\delta-1)\pi)^c I_{(0,1)}(\pi) \times \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \binom{n}{y} N_{\delta}(a, b, c)\pi^{y+a-1}(1-\pi)^{n-y+b-1}(1+(\delta-1)\pi)^c I_{(0,1)}(\pi) \\ &\propto \pi^{y+a-1}(1-\pi)^{n-y+b-1}(1+(\delta-1)\pi)^c I_{(0,1)}(\pi), \end{aligned}$$

é a posteriori, a qual é outra distribuição Kummer do tipo II, com parâmetros $a' = a + y$ e $b' = n - y + b$.

A.0.10 Priori Beta Generalidaza do tipo II

Seja a priori Beta Generalidaza do tipo II com FDP dada por:

$$g(\pi) = \frac{D\pi^{\alpha-1}(1-\pi)^{\beta-1} \exp(-p\pi)}{(\pi+z)^{\rho}}.$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = \frac{D\pi^{\alpha-1}(1-\pi)^{\beta-1} \exp(-p\pi)}{(\pi+z)^{\rho}} \times \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \binom{n}{y} \frac{D\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1} \exp(-p\pi)}{(\pi+z)^{\rho}} \\ &\propto \frac{\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1} \exp(-p\pi)}{(\pi+z)^{\rho}}, \end{aligned}$$

é a posteriori, a qual é outra distribuição Beta Generalidaza do tipo II, com parâmetros $\alpha' = a + y$, $\beta' = n + \beta - y$, ρ , p e z .

A.0.11 Priori Kummer beta II

Seja a priori Kummer beta II com FDP dada por:

$$g(\pi) = \frac{\pi^{\alpha-1}(1-\pi)^{\beta-1}e^{\lambda\pi}}{B(\alpha, \beta)_1 F_1(\alpha, \alpha + \beta, \lambda)}.$$

Pelo Teorema de Bayes, segue que:

$$\begin{aligned} g(\pi|y) &\propto g(\pi) \times f(y|\pi) = \frac{\pi^{\alpha-1}(1-\pi)^{\beta-1}e^{\lambda\pi}}{B(\alpha, \beta)_1 F_1(\alpha, \alpha + \beta, \lambda)} \times \binom{n}{y} \pi^y (1-\pi)^{n-y} \\ &= \binom{n}{y} \frac{\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}e^{\lambda\pi}}{B(\alpha, \beta)_1 F_1(\alpha, \alpha + \beta, \lambda)} \\ &\propto \pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}e^{\lambda\pi}, \end{aligned}$$

é a posteriori, a qual é outra distribuição Kummer beta II, com parâmetros $\alpha' = \alpha + y$ e $\beta' = n + \beta - y$ e λ .

Portanto, com a realização dos cálculos acima, constatou-se que as generalizações de Beta usadas neste trabalho são todas conjugadas para o modelo binomial.